The validation of brief cognitive assessment tools to evaluate suspected dementia in Chinese-speaking populations

Ruan-Ching Yu

Prepared under the supervision of
Professor Gill Livingston
Doctor Naaheed Mukadam

(Supervisory panel)
Professor Narinder Kapur
Doctor Joshua Stott

A thesis submitted for the degree of Doctor of Philosophy
at the Division of Psychiatry, University College London

March, 2023

Declaration

I, Ruan-Ching Yu, confirm that the work presented in the thesis is my own. Where information has been derived from other sources, I confirmed that this had been indicated in the thesis.

7th March 2023

Ruan-Ching Yu

Date

Acknowledgements

My supervisors, Professor Gill Livingston, Dr Naaheed Mukadam, and my supervisory panel, Professor Narinder Kapur, and Dr Joshua Stott, have been tremendously encouraging and supportive throughout the duration of my PhD. I am so thankful to have had four of you as my supervisors and my supervisory panel to direct my work with your clinical, statistical, and academic expertise. Dr Kathy Liu is thanked for practising the administration of ACE-III in Chinese with me.

Thank you to the Ministry of Education in Taiwan for the funding which allowed me to complete a PhD at UCL. Thank you to the neurologists and postdoc researcher, Drs Chaur-Jong Hu, Chien-Tai Hong, Lung Chan, Li-Kai Huang, and Cheng-Chang Yang, for assisting me with adapting the tests and recruiting the participants. Thank you to the psychologist, Pei-Jung Wu, for our discussion regarding the administration of the tests. Thank you to the nurses, particularly Yun-Zheng Li, for hunting for suitable clinics for my interviews. In addition, I am extremely grateful to all those who participated in my study. Special thanks to my friend Jen-Chieh Lai, a psychologist, for answering my questions regarding clinical practise in Taiwan and participating in my systematic review study as a second rater. Huge thanks to my friend, R specialist Hung-Ling Chang, for designing a customized script for my meta-analysis.

I am grateful to my colleagues from the Division of Psychiatry and friends from Taiwan; I cannot name each of you and explain why (yes, you know I am referring to you), but Ping-Jen, many thanks for connecting me with the collaborative networks and showing me around UCL when we first met; And, Kai-Yuan, many thanks for providing me with a wealth of helpful advice on how to survive through some extremely difficult times encountered during my PhD. Moise, Merve, and Inayah, thank you for being my dinner partners and for always being amusing. Wan-Ting, thank you for offering me accommodation throughout my time in Taiwan.

To my Dad, Mum, and Brother, thank you for taking the time to comfort me by text message or phone call during the most difficult times of COVID-19, and for

your unwavering faith in my ability to overcome any obstacle overseas. Nini and Adin, despite the fact that one of you appeared to have damaged my laptop, I am still glad to have enjoyed the company of your most adorable pets during my stay at home.

Thank you, Grandma, for remembering who I am for so long despite your dementia. Our talk motivated me to pursue a PhD in the field of dementia.

Abstract

Background

Around 20% of individuals living with dementia are from Chinese-speaking populations. Brief cognitive assessments can assist in the diagnosis of dementia, but it is unclear which are best for Chinese-speaking populations. Additionally, there is a pressing need to validate brief cognitive assessments in Taiwan.

Aims

I aimed, firstly, to critically examine evidence about brief (< 20mins) cognitive assessment tools used to assist in the diagnosis of dementia in Chinese-speaking populations. Secondly, I aimed to identify gaps in such instrument validation for Taiwanese populations and validate a brief cognitive test for use with all patients presenting with suspected dementia. An additional aim was to culturally adapt and test a cognitive assessment battery for acceptability in this group.

Methods

I conducted a systematic review and meta-analysis of brief cognitive assessments used for suspected dementia and mild cognitive impairment (MCI) among Chinese-speaking populations. I used quality criteria for instruments as > 75% for sensitivity and specificity and < 75% for heterogeneity. I validated the Taiwanese version of the Addenbrooke's Cognitive Examination, ACE-III (T-ACE-III) with 90 participants in a Taiwanese memory clinic. I culturally adapted a brief cognitive assessment battery and determined its acceptability in Taiwan.

Results

My meta-analysis results found that the ACE-III and ACE-R were the only tests that met quality criteria across clinical populations for dementia and MCI. My

empirical study showed that the T-ACE-III detected dementia in a Taiwanese population with a specificity of 100.0%, a sensitivity of 89.5%, and a Youden Index of 0.895, using an optimal cut-off value of 73/74.

Conclusions

The ACE-III and ACE-R demonstrated satisfactory psychometric properties for suspected dementia and MCI in Chinese-speaking populations. The T-ACE-III is an effective tool to help diagnose dementia in a Taiwanese clinical population.

Further Research

Building on my thesis publication, further research could examine T-ACE-III in dementia subtypes.

Impact Statement

The work in my thesis explores the brief cognitive instruments that have been validated in Chinese-speaking populations and examines their cultural adaptation processes. I also conducted a validation study in Taiwan, a Chinese-speaking culture. The results of my work have potential impact on the choice of cognitive assessment measures used in clinical settings, procedures used for adaptation and translation, and may help to inform those implementing policies related to dementia diagnosis.

1. Clinical impact

The main aims of my thesis were to identify the optimal brief cognitive assessment tools that have been used in Chinese-speaking populations and validate a specific tool in one population – Taiwan. Clinical time constraints mean that clinicians require validated short cognitive instruments or fewer patients will be seen. I found that there are relatively few options for cognitive assessments used in assisting diagnosis in Taiwan. One of the most commonly used diagnostic routine tests is the Mini-Mental State Examination (MMSE), where performance is affected greatly by education, thus underestimating impairment in the most educated and overestimating it in those with little education.

I culturally adapted the Taiwanese version of ACE-III (T-ACE-III) and found it to be reliable, valid, and quick to administer when used as a cognitive test to inform the diagnosis of people with suspected dementia. I have published the findings and made it available on the website of the Brain and Mind Centre at the University of Sydney, providing psychologists, psychiatrists, neurologists, and other clinicians in Taiwan with more alternatives. The T-ACE-III could be used to help with early diagnosis and potentially in monitoring the progress of dementia or examining the effectiveness of a cognitive intervention or treatment.

Additionally, in my meta-analysis and narrative review research, I was able to summarise in detail information about each brief cognitive assessment (e.g., validity measures, participants studied (types of dementia and MCI)). This

information allows clinicians to select the most suitable tools taking into account particular individuals and settings.

2. Academic impact

I have disseminated most of the results of my PhD at international conferences (see *Chapter 11*) and in peer-reviewed journals. I have published the findings of my validation study in a neuropsychology journal (Yu et al., 2021), and I have submitted the findings of my systematic review and meta-analysis to a dementia-specific journal.

3. Policy and public health impact

In terms of policy implication, timely diagnosis of dementia has been a key policy goal in many countries, including Taiwan. The Ministry of Health and Welfare in Taiwan (2020) has set a target to increase the dementia diagnosis rate from 30% to 70% by 2025. As demonstrated by my PhD research, the Taiwanese version of the ACE-III performs exceptionally well, highlighting its potential as an effective tool for accurate dementia diagnosis.

In summary, the outcome of my work benefits clinical practice by providing evidence regarding cognitive tests for clinicians to choose in Taiwan, and in other Chinese-speaking populations.

UCL Research Paper Declaration Form

referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- *have been uploaded to a preprint server*
- *are in submission to a peer-reviewed publication*
- *have been published in a peer-reviewed publication, e.g. journal, textbook.*

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)	
a) What is the title of the manuscript?	Validation of the Taiwanese version of ACE-III (T-ACE-III) to detect dementia in a memory clinic
b) Please include a link to or doi for the work	https://doi.org/10.1093/arclin/acab089
c) Where was the work published?	Archives of Clinical Neuropsychology
d) Who published the work?	OUP
e) When was the work published?	February 2022
f) List the manuscript's authors in the order they appear on the publication	Ruan-Ching Yu, Naaheed Mukadam, Narinder Kapur, Joshua Stott, Chaur-Jong Hu, Chien-Tai Hong, Cheng-Chang Yang, Lung Chan, Li-Kai Huang, Gill Livingston
g) Was the work peer reviewed?	Yes
h) Have you retained the copyright?	Yes If 'No', please seek permission from the relevant publisher and check the box next to the below statement: ☐ *I acknowledge permission of the publisher named under 1d to include in this thesis portions of the publication named as included in 1c.*
i) Was an earlier form of the manuscript uploaded to a preprint server?	No
2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)	
a) What is the current title of the manuscript?	Systematic review and meta-analysis of brief cognitive instruments to evaluate suspected dementia in Chinese-speaking populations
b) Has the manuscript been uploaded to a preprint server?	Yes. http://dx.doi.org/10.2139/ssrn.4216761
c) Where is the work intended to be published?	Journal of Alzheimer's Disease Report
d) List the manuscript's authors in the intended authorship order	Ruan-Ching Yu, Jen-Chieh Lai, Esther K. Hui, Naaheed Mukadam, Narinder Kapur, Joshua Stott, Gill Livingston
e) Stage of publication	Accepted

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

For 1. a): Ruan-Ching Yu, Naaheed Mukadam, Narinder Kapur, Joshua Stott, Gill Livingston did the conceptualization and design, analysis and interpretation of data, and preparation of manuscript. Chaur-Jong Hu, Chien-Tai Hong, Cheng-Chang Yang, LungChan, Li-Kai Huang did the reading and commenting on manuscript, provision of suggestion on literature review and discussion.

For 2. a): Gill Livingston and Ruan-Ching Yu conceptualized and designed the study. Gill Livingston and Naaheed supervised the research activities. Ruan-Ching Yu prepared (searching, screening, retrieving, and maintaining the research data), analysed the data (using R software to conduct the meta-analysis), visualised the results, scored all studies of the selected articles, interpreted the results, and wrote up the original draft. Gill Livingston, Naaheed Mukadam, Narinder Kapur, and Joshua Stott interpreted the results, reviewed, and edited the manuscript. Jen-Chieh Lai prepared (screening and retrieving the research data) and scored all studies with clinical controls. Esther K. Hui scored of 20% studies with community-based controls and formatted the manuscript.

4. In which chapter(s) of your thesis can this material be found?

For 1. a): Chapters 7 and 8
For 2. a): Chapters 4, 5, and 6

5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

	Ruan-Ching Yu		
Candidate:	_____	*Date:*	07/03/2022
	Gill Livingston		
Supervisor/ Senior Author (where appropriate):	_____	*Date:*	07/03/2022

Table of Contents

Declaration 2

Acknowledgements 3

Abstract 5

Impact Statement 7

Statement of personal contribution 17

Abbreviations 18

Chapter 1. The current project 20

1.1 Personal interests 20

1.2 The Taiwanese context 21

1.3 Start of the project 23

1.4 Outline of my thesis chapters 23

Chapter 2. Background 25

2.1 Dementia 25

2.2 Chinese-speaking culture 28

2.3 Brief cognitive assessments 32

2.4 Summary 36

Chapter 3. Aims, objectives, and research questions 37

3.1 Systematic review and meta-analysis of brief cognitive tests (addressed in *Chapter 4*) 37

3.2 Narrative review of the brief cognitive tests excluded from the meta-analysis (addressed in *Chapter 5*) 37

3.3 The reported translation and cultural adaptation of all brief cognitive instruments (addressed in *Chapter 6*) 38

3.4 The T-ACE-III study (addressed in *Chapter 7*) 38

3.5 The supplementary tests of the T-ACE-III for individual cognitive domains and assessment of malingering (addressed in *Chapter 8*) 39

Chapter 4. Systematic review and meta-analysis of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia 40

4.1 Introduction 40

4.2 Methods 43

4.3 Results 50

4.4 Discussion 65
4.5 Conclusions 69

Chapter 5. Narrative review of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia 70
5.1 Research aims and questions 70
5.2 Methods 70
5.3 Results 71
5.4 Discussion 78

Chapter 6. The translation and cultural adaptation procedures employed for brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia 81
6.1 Research aims and questions 81
6.2 Methods 81
6.3 Results 84
6.4 Summary of findings 88

Chapter 7. Validation of the Taiwanese version of ACE-III for detecting suspected dementia in Taiwan 89
7.1 Introduction 89
7.2 Aims 94
7.3 Methods 95
7.4 Results 106
7.5 Discussion 114
7.6 Conclusions 118

Chapter 8. Acceptability of tests of executive function, speed of processing and performance validity 119
8.1 Background 119
8.2 Methods 124
8.3 Results 128
8.4 Discussion 132
8.5 Conclusions 133

Chapter 9. General Discussion 134
9.1 Summary of findings 134
9.2 Comparison with other research 137
9.3 Strengths and limitations of my PhD thesis 154
9.4 Clinical implications 156

9.5 Policy implications 158
9.6 Future research 158
9.7 Dissemination 161
9.8 The potential impact of the COVID-19 pandemic on the administration of cognitive assessments..... 161

Chapter 10. Conclusions 163
10.1 Chinese brief cognitive assessments in clinical settings 163
10.2 Brief cognitive assessments in Taiwan 163
10.3 Conclusions 164

Chapter 11. Other related academic achievements 165
11.1 Papers from my PhD 165
11.2 Conferences and other presentations 165
11.3 Other achievements 166

References 167

Appendices 186
Appendix 1: Registered protocol on PROSPERO 186
Appendix 2: Operational criteria for the scoring of the Centre for Evidence-based Medicine (CEBM) diagnostics criteria. 187
Appendix 3: PRISMA-DTA Checklist 188
Appendix 4: Brief cognitive tests for dementia: Diagnostic performance (in alphabetical order) 191
Appendix 5: Brief cognitive tests for MCI: Diagnostic performance (in alphabetical order) 195
Appendix 6: 2x2 data table 198
Appendix 7: Random-effect bivariate model analysis 199
Appendix 8: The HSROC / SROC curves of the eight tests 200
Appendix 9: Likelihood ratio regression based on bivariate models (clinical context) 202
Appendix 10: Likelihood ratio regression based on bivariate models (population) 203
Appendix 11: Likelihood ratio regression based on bivariate models (subtype) 204
Appendix 12: Likelihood ratio regression based on bivariate models (reference standard) 205
Appendix 13: Likelihood ratio regression based on bivariate models (scoring system) 206
Appendix 14: Deek’s funnel plots 207

Appendix 15: The MTRQ and MCAR Criteria 208
Appendix 16: Published paper from the T-ACE-III study 209
Appendix 17: T-ACE-III study - UCL ethical approval letter 221
Appendix 18: T-ACE-III study - Taiwan ethical approval letter (in English & traditional Chinese) 223
Appendix 19: T-ACE-III study - The use of ACE-III approval letter from University of Sydney 225
Appendix 20: T-ACE-III study - The use of Brixton Spatial Anticipation test approval letter from the original author(s) 226
Appendix 21: T-ACE-III study - The use of The-coin-in-the-hand test approval letter from the original author(s) 227
Appendix 22: Discussion on T-ACE-III adaptation 228
Appendix 23: T-ACE-III study - poster (in English & traditional Chinese) 230
Appendix 24: T-ACE-III study - participant information sheet (in English & traditional Chinese) 232
Appendix 25: T-ACE-III study - proxy information sheet (in English & traditional Chinese) 238
Appendix 26: T-ACE-III study informed consent form for participants (in English & traditional Chinese) 244
Appendix 27: T-ACE-III study informed consent form for interviewees and proxy (in English & traditional Chinese) 251
Appendix 28: T-ACE-III study- T-ACE-III questionnaire (UK & Taiwanese versions) 258
Appendix 29: ROC curves of T-ACE-III, CASI, and MMSE 270
Appendix 30: T-ACE-III study- Brixton test questionnaire (UK & Taiwanese versions) 271
Appendix 31: T-ACE-III study- The-coin-in-the-hand test (UK & Taiwanese versions) 273

List of tables

Table 1: The contrast between written forms of simplified and traditional Chinese 30

Table 2: Possible scenarios where some form of cross-cultural adaptation is required 31

Table 3: Definition of common measures of test accuracy 47

Table 4: Characteristics of included studies and quality score (Arranged by country, setting, test name) 53

Table 5: The administration time and cognitive domains of all tests (including the studies included in the meta-analysis) 77

Table 6: Translation steps undertaken and the individual professionals involved in the papers included in the review 86

Table 7: Domains and items of ACE versions & MMSE 91

Table 8: Reported translations of ACE, ACE-R, ACE-III, & M-ACE in East Asia 94

Table 9: The modified items in the Simplified and Traditional Chinese versions of ACE-III compared to the English (UK) version 100

Table 10: Schedule of the recruitment 104

Table 11: Demographic and neuropsychological characteristics of participant groups 107

Table 12: Cut-off scores for the T-ACE-III in differentiating dementia from non-dementia, sensitivity values, specificity values, and Youden index 109

Table 13: Cut-off scores for the T-ACE-III (adjusted) in differentiating dementia from non-dementia 110

Table 14: Association of demographic factors with T-ACE-III score and diagnosis of dementia 113

Table 15: The executive domains covered by the tests used in this thesis 123

Table 16: The completion rate of dementia and non-dementia groups 129

Table 17: The performance on supplementary tests by dementia and non-dementia groups in Taiwan (with numbers completing each test) 131

Table 18: Systematic reviews of dementia diagnostic and screening tools 141

List of figures

Figure 1: PRISMA flow diagram of references identified and included in the review..... 51

Figure 2: Meta-analyses for diagnostic test accuracy on diagnosing Dementia 61

Figure 3: Meta-analyses for diagnostic test accuracy on diagnosing MCI 63

Figure 4: The adaptation processes for the T-ACE-III 97

Figure 5: ROC curves of T-ACE-III. 109

Figure 6: A set of responses from the Brixton test (from van den Berg et al., 2009, p. 697)..... 125

Statement of personal contribution

My contributions to my PhD are listed as follows:

- For my systematic review and meta-analysis, I wrote the protocol and finalised it after discussing and revising it together with my supervisors, registered the protocol on PROSPERO. I searched the literature in the specified databases, screened the abstracts and titles, extracted the data, selected suitable criteria for quality assessments, and collaborated with two raters to rate quality of the papers. I also conducted updated searches on 1 December 2020 and 12 October 2021 and updated PROSPERO. I learnt to use R and Stata software to conduct the meta-analysis. I completed the first draft of a paper in preparation for publication after finishing the meta-analysis and narrative synthesis. I updated the paper after comments from co-authors.

- For my validation study, I wrote the validation research protocol and finalised it after discussing and revising it together with my supervisors. I then contacted three neurologists in Taiwan's eastern, southern, and northern regions as potential collaborators. I applied for and gained ethical approval from the UCL Research Ethics Committee (UCL-REC) and Taipei Medical University-Joint Institutional Review Board (TMU-JIRB). I then culturally adapted and translated the instruments after discussing with other professionals. I collaborated with four neurologists. I invited the participants and obtained informed consent. I conducted all the interviews with patients in Taiwan, computed test scores, and analysed the data. I drafted the initial version of the published paper and updated it after comments from my co-authors.

- I wrote the entire thesis and revised it after comments from my supervisors.

Abbreviations

ACE-III	Addenbrooke's Cognitive Examination - III
ACE-R	Addenbrooke's Cognitive Examination - Revised
AD	Alzheimer's disease
aMCI	Amnestic mild cognitive impairment
amMCI	Amnestic mild cognitive impairment - multiple domains
asMCI	Amnestic mild cognitive impairment - single domain
AUC	Area under the curve
CASI	Cognitive Abilities Screening Instrument
CDR	Clinical Dementia Rating
CIH	Coin-in-the Hand test
CI	Cognitive impairment
CIs	Confidence intervals
DLB	Dementia with Lewy bodies
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTA	Diagnostic test accuracy
EH	Esther. K Hui
FTD	Frontotemporal dementia
GL	Gill Livingston
ICD	International Classification of Diseases
ICF	International Classification of Functioning, Disability and Health
JCL	Jen-Chieh Lai
JS	Joshua Stott
LR	Likelihood ratio
MCAR	Manchester Cultural Adaptation Reporting Questionnaire
MCI	Mild cognitive impairment
MMSE	Mini-Mental State Examination
MoCA	Montreal Cognitive Assessment
MTRQ	Manchester Translation Reporting Questionnaire
naMCI	Non-amnestic mild cognitive impairment
NCD	Neurocognitive disorder
NK	Narinder Kapur
NM	Naaheed Mukadam

PDD	Parkinson’s disease dementia
ROC	Receiver operating characteristic
SN	Sensitivity
SP	Specificity
T-ACE-III	Taiwanese version of the Addenbrooke’s Cognitive Examination III
VaD	Vascular dementia
vaMCI	Vascular mild cognitive impairment
WAIS-IV-PSI	Processing Speed Index from the Wechsler Adult Intelligence Scale IV
WHO	World Health Organization

Chapter 1. The current project

“Speaking in the same language does not equal communication, especially when there is a cultural divide.”
(Gerry Abbey, in *Cheers, Beers, and Eastern Promise*)

"Cultural differences should not separate us from each other, but rather cultural diversity brings a collective strength that can benefit all of humanity."
(Robert Alan)

1.1 Personal interests

I was first inspired in the field of assessment and statistics whilst working as a research assistant at the Science Education Centre at a university in Taiwan for two years, where I gained knowledge of different types of assessments, and learnt advanced statistical skills from exploring the factors that may influence test performance.

I expanded this passion to the clinical field of dementia whilst working as a research assistant in the neurology department at a hospital in Taiwan. My fluency using the local dialect (Taiwanese Hokkien, please *see 1.2.2 The multi-ethnicities living in* Taiwan) improved after practising Hokkien with the psychologists and my cousins. During that time, I first used cognitive assessments with patients who had dementia and MCI, and I came across some language barriers when talking to Hokkien users. For example, there was no formal guidance on switching between Chinese and Taiwanese Hokkien or using adapted phrases on certain items, such as “repeat this sentence after me”. Also, the cognitive tests usually took an hour or more to administer, and patients often refused to continue or were abrupt with the interviewers. This resulted in a waste of resources and time, or in the worst-case scenario led to a late diagnosis (please see as *1.2.3 Diagnostic methods in Taiwan*).

At a personal level, I clearly understand how a late diagnosis can cause problems to an entire family because my grandmother has dementia, in the context of a

late diagnosis and a lack of awareness of information about the disease. After finding out she had multiple strokes, she started to experience an abrupt decline in cognitive function. Three years before the diagnosis was formally made, she displayed unusual behaviours from time to time. Following the diagnosis, my father quit his job and devoted himself to taking care of her for nine years. As the disease progressed, the whole family had to bear financial and psychological distress. Some of this could have been avoided if an earlier, accurate diagnosis had been made.

1.2 The Taiwanese context

Dementia is perceived differently by various Chinese-speaking populations. For example, dementia may be seen as an unanticipated tragedy for Taiwanese people, but it is considered as an expected life occurrence due to natural ageing for people from China and Hong Kong (Ramsay, 2017). In the following sections, I present dementia in a Taiwanese context in terms of an ageing society, a multi-ethnic population, and diagnostic processes.

1.2.1 The ageing society in Taiwan

Fortunately, fewer people are dying young now and this has led to an increase in the number of older people, which in turn means there are more individuals living with dementia. Taiwan is not exempt from this scenario. According to epidemiological studies conducted in Taiwan between the 1980s and 1990s, the prevalence of dementia ranged from 1.7% to 4.3% among individuals aged more than 65, which is much lower than the rate of 5% to 10% reported in Europe and the US (Yang, 2016). This could be due to low reporting and diagnosis rates, higher mortality rates, or differences in race (Yang, 2016). A more recent national study (Chen et al., 2014) conducted between 2011 and 2013 used the Ascertainment of Dementia 8 (AD8) for screening and found the numbers for very mild dementia were under-reported in some early studies.

1.2.2 The multi-ethnicities living in Taiwan

Taiwan, where I am from, has a population made up of multiple ethnic groups that emigrated from other countries (mainly China) at different times, and different dialects are spoken locally. The four main ethnic groups of Taiwanese people are Southern Min, Hakka, Mainlanders, and Taiwanese Aboriginals. The most commonly spoken Chinese dialect is Taiwanese Hokkien (also called Taiwanese, spoken by around 70% of the entire Taiwanese population) although Mandarin Chinese is the official language in Taiwan. Therefore, cognitive assessments from other Chinese-speaking populations (e.g., China or Singapore) are unsuitable and it is essential to culturally adapt and validate the cognitive assessments for Taiwan for its local inhabitants, so as to create an optimal cognitive battery that can be used to assist timely diagnosis of dementia.

1.2.3 Diagnostic methods in Taiwan

Whilst working at the department of neurology in Taiwan, I noticed that the most frequently used cognitive assessment tools for assisting diagnosis in clinics were the Mini-Mental State Examination (MMSE) (Folstein et al., 1975), Alzheimer's Disease Assessment Scale (ADAS-Cog) (Rosen et al., 2006) and the Cognitive Abilities Screening Instrument (CASI) (Lin et al., 2002). The Clinical Dementia Rating scale (CDR) (Morris, 1993) was used for both making a dementia diagnosis and for staging dementia. In recent years, there have been many newly developed cognitive assessment tools in western countries, but they have not yet been validated in every Chinese-speaking population. Clinicians cannot easily access information about which test to use in diagnosing dementia as instruments have not been systematically reviewed and subject to a meta-analysis. Therefore, the options for diagnostic assessments for clinicians are relatively few. There is a pressing need for this information and for validation of brief cognitive tests that harmonise with Taiwanese culture, so as to allow timely detection of possible dementia in Taiwanese populations.

1.3 Start of the project

For this research project, I aimed to review the diagnostic performance of brief cognitive assessments used in detecting dementia among Chinese-speaking populations, and culturally adapt and validate the T-ACE-III. I also aimed to establish the acceptability of a cognitive battery for detecting dementia with patients in Taiwan. I received funding from the Ministry of Education in Taiwan for my PhD, which I started in April 2019. My supervisors were Professor Gill Livingston (GL; primary supervisor, Department of Mental Health of Older People UCL), Doctor Naaheed Mukadam (NM; second supervisor, Department of Mental Health of Older People, UCL), and my wider supervisory panel comprised Professor Narinder Kapur and Doctor Joshua Stott (NK, JS; Department of Clinical, Educational and Health Psychology, UCL).

1.4 Outline of my thesis chapters

In *Chapter 2*, I introduce the terms I use extensively in my thesis. I discuss the prevalence of dementia globally and in Taiwan and outline the latest diagnostic criteria for dementia. I then describe Chinese-speaking cultures, the differences between simplified and traditional Chinese, and procedures for cultural adaptation and translation of cognitive tests. I next highlight the key features of brief cognitive assessments and examine the main findings of past systematic reviews on this topic.

In *Chapter 3*, I list the aims and objectives of my PhD.

In *Chapter 4*, I provide the background, methods, findings, and discussion of the systematic review and meta-analysis which I carried out on the brief cognitive assessment tools used in detecting suspected dementia among Chinese-speaking populations.

In *Chapter 5*, I describe the results of narrative review of the brief cognitive assessment tools that were excluded from the meta-analysis due to an insufficient number of studies (e.g., they were only used in a particular Chinese-

speaking population). In conjunction with the findings of *Chapter 4*, I explain how findings linked to the next stage of my PhD.

In *Chapter 6*, I describe the published translation and cultural adaptation processes for brief cognitive assessments used in each study (including the studies in the meta-analysis and the narrative review).

In *Chapter 7*, I detail how I conducted the validation study of the ACE in Taiwan step-by-step, including adapting the instruments to construct the T-ACE-III. I also outline how I invited participants to take part in the study, and how I gathered, and analysed the data.

In *Chapter 8*, I discuss the acceptability and performance of the supplementary tests (Brixton test, WAIS-IV-PSI, and the Coin-in-the-Hand test) in Taiwanese patients with dementia.

In *Chapter 9*, I discuss the overall results of my thesis. I compare my results with findings from past studies and consider the strengths and limitations of my work. I also discuss implications for future research and practice.

In *Chapter 10*, I outline the conclusions of my PhD research.

Finally, in *Chapter 11*, I list other academic achievements related to my PhD.

Chapter 2. Background

In this background chapter, I first report the latest diagnostic criteria for dementia. Then I report figures for its worldwide prevalence, prevalence in Chinese-speaking populations, and specifically the prevalence in Taiwan. I explain why historical, economic, and societal contexts differ among different Chinese-speaking populations and the differences between written forms of simplified and traditional Chinese. I introduce five possible scenarios when cross-cultural adaptation is needed, and the steps recommended for such adaptation from existing guidelines.

Finally, I summarise the features of brief cognitive assessments, the results from past systematic reviews, and possible issues while applying brief cognitive assessment tools in dementia services.

2.1 Dementia

Dementia diagnosis, treatment, care, and support is included in the framework of the global action plan of the World Health Organization (WHO) for the period from 2017 to 2025 (WHO, 2012, 2017). On account of its influence on the individual, their family, and the society, the WHO has designated dementia as a priority for public health.

2.1.1 Diagnostic criteria for dementia

In 2011 and 2013, the National Institute on Aging and Alzheimer's Association (NIA-AA) and American Psychiatric Association (APA) updated the diagnostic criteria for dementia in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). At the end of 2022, the text revision of DSM-5 (DSM-5-TR) was released (APA, 2022). Memory impairment is no longer necessarily considered as the main symptom for the diagnosis of dementia. Instead, symptoms include personality and neurobehavioral change in the cognitive and

behaviour domains and at least two instances of cognitive impairment (Jack et al., 2018). Mild cognitive impairment (MCI) or mild neurocognitive disorder (mild NCD) is considered as an at-risk condition rather than a diagnosis of illness by many people. It is characterised by a decline in cognitive abilities beyond normal ageing, without impacting on daily life. The six cognitive domains which may, compared to a previous level, be affected in MCI (or mild NCD: at least one deficit) and dementia (major NCD: at least two deficits) are listed in the newest edition of DSM (DSM-5-TR) (APA, 2022). They are:

- **Complex attention,** which includes sustained attention, divided attention, selective attention, and information processing speed
- **Executive function,** which includes planning, decision making, working memory, responding to feedback, inhibition, and mental flexibility
- **Learning and memory,** which includes free recall, cued recall, recognition memory, semantic and autobiographical long-term memory, and implicit learning
- **Language,** which includes object naming, word finding, fluency, grammar and syntax, and receptive language
- **Perceptual-motor function**, which includes visual perception, visuo-constructional reasoning, and perceptual-motor coordination
- **Social cognition**, which includes recognition of emotions, theory of mind and insight. Symptoms of social cognition may be particular noticeable in assessing patients with non-Alzheimer's dementias, such as frontotemporal dementia (Simpson, 2014).

The DSM-5 (or DSM-5TR) offers criteria and definitions for mental health practitioners to classify diseases using a common language, and it is specifically designed for mental disorders. As opposed to the DSM-5 (or DSM-5TR), which covers mental illnesses in extensive detail, the international statistical classification of diseases and related health problems: tenth revision (ICD-10) encompasses conditions and diseases related to every aspect of human physiology. ICD-10 states that the manifestation of dementia (F00-F03) is disturbance of multiple higher cortical functions, including memory, thinking,

orientation, comprehension, calculation, learning capacity, language, and judgement. The impairments are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation (WHO, 2004).

The framework of International Classification of Functioning, Disability and Health (ICF) endorsed by WHO in 2000 (WHO, 2001), is a third set of diagnostic criteria revised from the International Classification of Impairment, Disability, and Handicap (ICIDH) that emphasises functioning as a component of health and moves beyond the approach of treating diseases as a result of their symptoms (i.e., impairment, disability, and handicap) (Badarunisa et al., 2015). ICF does not provide a definition for dementia but proposed a generic qualifier describing the frequency and intensity of a problem occurring over the previous 30 days in terms of three components: Body functions and structures, activities, and participation (at individual and societal levels), and personal and environmental factors (at a contextual level) (WHO, 2003), in order to understand the functional outcomes of dementia (Badarunisa et al., 2015). The primary diagnostic feature indicated by DSM and ICD is "cognitive impairment sufficient to impair personal activities of daily living or which causes significant social and occupational impairments." The first component of ICD (body functions and structures) reflects "cognitive impairment," while the second component (activities and participation) can be referred to as "personal activities of daily living," and the third component (environmental factors) measures the barriers or facilitators in the environment of people with dementia (Badarunisa et al., 2015).

There are global (e.g., consciousness, orientation, global psychosocial functions, temperament and personality, sleep functions) and specific (e.g., attention, memory psychomotor, emotion, perception) mental functions under the system of ICF, which is distinct from DSM-5 (or DSM-5TR) and ICD-10. In general, ICF is more holistic in disease classification than DSM-5 (or DSM-5TR) and ICD-10, with the environmental component exploring the physical, social, and attitudinal factors that influence how people live their lives.

My study focuses mainly on dementia and the evaluation of cognitive functions; hence I use the DSM-5TR, which includes dementia within its extensive and in-depth coverage of mental disorders.

2.1.2 Prevalence

There are over 55 million people living with dementia globally (Gauthier et al., 2021) of whom over 11 million people are from Chinese-speaking populations (approximately 10 million in China, 226,000 in Taiwan, 123,000 in Malaysia, 115,000 in Hong Kong, and 45,000 in Singapore), and this number is estimated to exceed 19 million by 2030 (ADI, 2014). According to the latest prevalence investigation conducted in 2019 by the Ministry of the Interior in Taiwan, dementia affects 280,000 (7.8%) out of 3.6 million individuals over the age of 65, which means there is one person living with dementia for every 12 people. The ageing population (> 65 years old) had already reached 14% of the total population in March 2018 and Taiwan is predicted to become a ‘super-aged society’ by 2025 where at least 20% of the population is over 65 years old (Ministry of Health and Welfare in Taiwan, 2020). In 2017, the WHO proposed the “Global action plan on the public health response to dementia 2017-2025” (WHO, 2017). One of its aims is improving diagnosis, treatment, and care. Failing to diagnose dementia can delay medical and social treatment and support, whereas erroneous diagnosis of dementia can lead to a waste of healthcare systems and needless distress for the individual and their family.

2.2 Chinese-speaking culture

2.2.1 Chinese-speaking populations

The Chinese languages (and all its varieties and dialects, such as Mandarin and Cantonese) are one of the most commonly spoken languages in the world, with a growing number of speakers, reaching 1.31 billion in 2020 (Statista, 2021). Legally, Mandarin Chinese is the national language in The People’s Republic of

China (including Hong Kong and Macau), Taiwan, Malaysia, and Singapore (Williams, 2016). The spoken form of Mandarin Chinese in Taiwan and China can be compared to English in the UK and the US, where the majority of the language remains the same, but vocabulary, intonation, and accent vary. For example, in Taiwan, “tǔ dòu” refers to peanuts, but in China it refers to potatoes. In addition, Mandarin and Hokkien are widely intermingled in Taiwan. These variations may represent the diversity of Chinese-speaking populations. They share very different historical, economic, and societal contexts as well as differing trajectories of life expectancy and health status.

Despite the difference in varieties of spoken “Chinese” language (Cantonese, Hakka, Wu and Min language) across different populations, technically, they are often considered as being in the same language group (Williams, 2016).

2.2.2 Written forms of Chinese

In Taiwan, Hong Kong, and Macau, people use Traditional Chinese in writing, whereas, in China, and Singapore they use Simplified Chinese. The most significant difference between the written form of Simplified and Traditional Chinese is that they use different forms of the individual characters. Simplified Chinese was implemented by the Chinese government between 1956 and 1964 in order to improve literacy. There are many different methods of simplification, but some of the most commonly used methods, in which a reader of one may not be able to read the other, are shown in *table 1*. Taiwanese people who never learnt Simplified Chinese are unable to recognise the simplified form. This is particularly the case for older people. However, the younger generation are more likely to have encountered the simplified form on the internet and so can usually read it. Furthermore, the Chinese Pinyin only exists in the system of simplified Chinese in China and Singapore, reflecting the Chinese phonemic system which is totally different to traditional Chinese. Different Chinese-speaking populations may create their own characters or words according to the grammatical rules. To other Chinese users, the writing in the local press of Hong Kong and Taiwan is frequently unintelligible (Wong & Xiao, 2010).

Table 1: The contrast between written forms of simplified and traditional Chinese

Method of simplification	Traditional Chinese	Simplified Chinese	English	Problems in recognising
Replacing a character with a less-commonly used and phonetically similar character	穀	谷	cereal; grain / valley; gorge	No relation between each other, but phonetically similar
	醜	丑	ugly / clown	
Simplifying the character based on handwriting or calligraphy style of the same character	書	书	book	Not familiar with handwriting or calligraphy
	興	兴	rise; prosper	
Omitting character components	廣	广	broad	Little indication as to which characters are represented
	廠	厂	factory	
Creating a new character of elements chosen for their phonetic and semantic values	歷	历	experience; to experience	
	曆	历	diary	
	賓	宾	guest	As above

Note. This table contains some information derived from a Wikipedia page (https://zh.wikipedia.org/zh-sg/简化字的简化方法)

2.2.3 Guidelines for cross-cultural adaptation

The assessments validated in one Chinese-speaking population still need cultural adjustments in order to be used in another Chinese-speaking population. A systematic review, with a focus on the assessments used in Asia (China, Hong Kong, India, Japan, Malaysia, Philippines, Singapore, South Korea, Sri Lanka, and Taiwan), reported that the majority of screening tools used have been adapted and translated from English and were originally designed to be applied in European or North American populations. It is therefore necessary to validate assessments in individual Asian cohorts with differing cultures and languages so as to produce robust, culturally and educationally sensitive instruments (Rosli et al., 2016).

There are guidelines for cross-cultural adaptation of individual items, instructions for the test, and the response sheet that were designed, based on a review of such adaptations in the fields of medicine, sociology, and psychology (Beaton et al., 2000). The term ‘cross-cultural adaptation’ refers to not only the process of

translation from a source language to a target language, but also issues such as preparing tests for use in different settings (Beaton et al., 2000). As set out in *Table 2*, cultural adaptation should be considered for five possible scenarios (see *Table 2*), and the recommended processes include five stages: Translation, synthesis, back translation, expert committee review, and pretesting (Beaton et al., 2000).

Table 2: Possible scenarios where some form of cross-cultural adaptation is required

	Using a questionnaire in a new population	**Results in a change in…**			**Adaptation required**	
		Culture	Language	Country of use	Translation	Cultural adaptation
A	*Use with same population. No change in culture, language, or country from source*	-	-	-	-	-
B	*Use with established immigrants in source country*	✓	-	-	-	✓
C	*Use in other country, same language*	✓	-	✓	-	✓
D	*Use with new immigrants, not English-speaking, but in same source country*	✓	✓	-	✓	✓
E	*Use in another country and another language*	✓	✓	✓	✓	✓

Note. Adapted from Guillemin et al. (1993), retrieved from Beaton et al. (2000)

In 2017, a systematic review scored the procedures for translation and cultural adaptation reported in studies of the ACE – ACE, ACE-R, ACE-III (Mirza et al., 2017). They used two scales - the Manchester Translation Reporting Questionnaire (MTRQ) and Manchester Cultural Adaptation Reporting Questionnaire (MCAR), to ascertain whether the procedures reported in each study can be successfully replicated by other researchers who intend to apply the ACE instruments. Only six out of 32 studies reached the highest standard in MTRQ and MCAR (Mirza et al., 2017), indicating that the reported procedures of translation and cultural adaptation were neglected. In 2020, there was a study aiming to develop guidelines specifically for the Urdu version of ACE-III (Waheed et al., 2020).

2.3 Brief cognitive assessments

The aims of the WHO “Global action plan on the public health response to dementia 2017-2025” (WHO, 2017) include improving diagnosis, treatment, and care. Failing to diagnose dementia can delay medical treatment and social care and support, whereas incorrectly diagnosing the presence of dementia can lead to a waste of healthcare resources and needless distress for the individual. The diagnosis of dementia can be divided into two main stages, detection of dementia (confirmation of initial diagnostic suspicion) and the aetiological diagnosis of the pathology causing the dementia (Riello et al., 2021). Various tests are used to assess cognitive functioning and everyday adjustment as part of the process of making a diagnosis. Cognitive assessment is usually based on paper-and-pencil neuropsychological tests that cover a wide range of cognitive functions, so as to provide an overview of the person’s general mental status (Arevalo-Rodriguez et al., 2015).

2.3.1 Features

An International Psychogeriatric Association survey found that the characteristics of brief cognitive instruments which can predict their frequency of use were “effectiveness”, “ease of administration,” and clinicians’ familiarity with the instruments (Shulman et al., 2006). However, frequency of use is not synonymous with accuracy; misclassification of suspected dementia may occur following brief assessments, and this may arise from biases which occur because of factors such as age, education, and ethnicity (Ranson et al., 2019).

2.3.2 Previous systematic reviews and meta-analyses of dementia cognitive diagnostic and screening tools

Several systematic review and meta-analysis studies, as well as narrative reviews, have been conducted on the effectiveness of cognitive diagnostic and screening measures for MCI or dementia. Most of these concentrated on primary

care or community settings (Brodaty et al., 2006; Cullen et al., 2007; Karimi et al., 2022; Lin et al., 2013; Lonie et al., 2009; Milne et al., 2008), or specified the test's requirements and features, such as tests assessing just a single cognitive domain (Mitchell & Malladi, 2010b), tests evaluating more than two cognitive domains but in less than 10 minutes (Mitchell & Malladi, 2010a), and tests requiring 10 to 45 minutes to measure (Appels & Scherder, 2010).

2.3.2.1 Systematic reviews of brief cognitive assessment tools

In more recent years, systematic reviews focusing on brief forms of cognitive assessments used in dementia have been completed for English (Velayudhan et al., 2014*, see 2.3.2.1.1 English-speaking populations*) and Spanish-speaking populations (Custodio et al., 2020, *see 2.3.2.1.2 Spanish-speaking populations*). Additionally, one systematic review of all dementia screening tools regardless of languages has also been completed (Tsoi et al., 2015, see *2.3.2.1.3 Asian populations*).

2.3.2.1.1 English-speaking populations

In the English-speaking study, the tests that demonstrated good psychometric properties were: The Test for Early Detection of Dementia (TE4D-cognitive part) for detecting AD in a community-based setting, Test Your Memory (TYM) and Addenbrooke's Cognitive Examination-III (ACE-III) for identifying dementia as well as AD and FTD in the clinical setting, Cognitive Assessment Screening Test (CAST) for recognising dementia in the general medical clinic. The study identified a need for validation in different settings, populations, and dementia subtypes. They recommended that future research should focus on more validations of tests that already exist rather than developing new ones (Velayudhan et al., 2014).

2.3.2.1.2 Spanish-speaking populations

In the study of instruments for Spanish-speaking populations, the MoCA was found to be the most widely used test, but it required cultural adaptation and specific cut-off scores for individuals with different educational levels; the Memory Alteration Test (M@T) was the only test that was validated in a group with a low education. The authors concluded there was an urgent need to validate appropriate brief cognitive assessment instruments in populations with a low educational level (Custodio et al., 2020).

2.3.2.1.3 Asian populations

There are also systematic reviews of studies completed in Asian and Chinese-speaking populations that examined not only brief forms of cognitive tests but also ones that last up to 20 - 45 minutes. One Asian study investigated cognitive assessment tools and included studies validated in many languages, such as Chinese (simplified and traditional forms), English, Japanese, Korean, Malay, Malayalam, Singhalese, and Tagalog. They reported that the MMSE and MoCA were the most widely validated and translated assessment tools in South Korea, China, Hong Kong, Japan, Taiwan, Malaysia, Philippines, Singapore, India, Sri Lanka; they noted that education bias existed, due to a low literacy rate among elderly populations in many Asian countries. This may have resulted in an overdiagnosis of dementia. Recommended cut-off scores of MoCA between different populations varied as items were adapted to suit their local culture, and this can result in differences in test scores between countries. The authors also pointed out that the majority of diagnostic and screening tools were designed for use with European or North American populations and few assessments were free from educational and cultural biases. They concluded that new assessment tools should be developed within large Asian cohorts in order to produce a sensitive instrument without such biases (Rosli et al., 2016).

2.3.2.2 Meta-analysis of brief cognitive assessment tools

Since the time that the MMSE's intellectual property rights were transferred to Psychological Assessment Resources in 2001, its utilisation has declined (Folstein et al., 1975; Powsner & Powsner, 2005). To discover MMSE alternatives, a meta-analysis and systematic review study on all dementia screening tools without language restraints was conducted. Mini cognitive scale (Mini-Cog) and Addenbrooke's Cognitive Examination-Revised (ACE-R) were the best in detecting dementia, whereas the Montreal Cognitive Assessment (MoCA) performed the best in identifying MCI (Tsoi et al., 2015).

In addition, a systematic review and meta-analysis in Chinese-speaking populations, that included diagnostic and screening tools taking less than 35 minutes, for all-phase, all-cause dementia, and in all settings (e.g., community, primary care, specialist care, hospital wards), identified 81 assessments from 167 studies. This review confirmed that the ACE-R and MoCA also had the best diagnostic accuracy in these populations (Huo et al., 2021). However, they included neuropsychological and cognitive tests that were used as a dementia screening tool across all contexts (e.g., day care centres, clinics, community) and not just as a diagnostic tool. In addition, the heterogeneity between studies was high for most screening tools, which may impact the generalisability of their findings. The study by Huo et al performed a meta-analysis of studies in all settings that provided an overview of test performances. Since they did not distinguish the settings (community or clinic) where the tests were employed, their findings cannot be directly applied to a specific environment.

Comprehensive neuropsychological testing is not only time-consuming but also not accessible to every clinician and may be burdensome for patients. There have been a regular stream of brief cognitive assessments or batteries developed to help diagnose suspected dementia; therefore, it is important to select the most appropriate brief assessment tool for individuals in order to achieve accurate early detection of dementia and MCI in clinical settings for a range of populations. It is also important to keep in mind the clinical history and demographic features of individuals with suspected dementia, noting that cognitive tests are only part of the diagnostic process.

2.4 Summary

In this chapter, I have reported on the prevalence of dementia, and that one-fifth of people with dementia are from Chinese-speaking populations. I have summarised the latest diagnostic criteria for dementia updated by NIA-AA and APA, ICD-10, and the ICF criteria proposed by WHO. I have pointed to the importance of culturally adapting an assessment from one Chinese-speaking population to another by outlining the differences between such populations, not only in historical, economic, and societal contexts but also in written forms of simplified and traditional Chinese. I have introduced the cross-cultural guidelines I used in my PhD project.

I summarised the key features (including advantages and disadvantages) of brief cognitive assessment tools used in dementia, findings from past systematic reviews of such tools, and limitations of systematic reviews. There appears to be no systematic review and meta-analysis study equivalent to mine that focuses on brief dementia assessment tools used in clinics for Chinese-speaking individuals with dementia. With an appreciation that there is a growing number of people living with dementia, the benefits of validating brief cognitive assessments for one Chinese-speaking population, namely Taiwan, would help in the early detection and diagnosis of dementia and provide more options for clinicians who choose to use brief cognitive assessments.

In the next chapter, I discuss the overall aims of my PhD and the objectives of the research studies I carried out.

Chapter 3. Aims, objectives, and research questions

The overall aim of my PhD was to investigate the brief cognitive assessment tools used in detecting suspected dementia in Chinese-speaking populations and to culturally adapt a brief cognitive assessment tool for use in Taiwan. The specific objectives and research questions for each part of my PhD work are listed below.

3.1 Systematic review and meta-analysis of brief cognitive tests (addressed in *Chapter 4*)

For my systematic review and meta-analysis, the objectives were:

1. To systematically review and meta-analyse the validation data relating to brief cognitive assessment tools used to detect suspected dementia (including MCI) in clinical settings in Chinese-speaking populations.
2. To use the results to inform my validation study of a brief cognitive assessment tool in Taiwan.

My research questions were:

1. What are the brief cognitive assessments with the best psychometric properties (validity and heterogeneity) in detecting dementia and MCI in clinical settings across different Chinese-speaking populations?
2. What is the quality of each study across different Chinese-speaking populations?

3.2 Narrative review of the brief cognitive tests excluded from the meta-analysis (addressed in *Chapter 5*)

For my narrative review of tests excluded from the meta-analysis as there was only one study, my objectives were:

1. To report the diagnostic performances of each brief cognitive assessment tools used in detecting dementia (including MCI) in Chinese-speaking populations.
2. To use the results to inform my validation study carried out in Taiwan.

My research question was:

1. What are the diagnostic performances, cognitive domains measured, administration time of brief cognitive assessment tools used in detecting suspected dementia (dementia, MCI) across different Chinese-speaking populations?

3.3 The reported translation and cultural adaptation of all brief cognitive instruments (addressed in *Chapter 6*)

For my analysis of the reported translation and cultural adaption, my objective was:

1. To document the procedures of translation and cultural adaptation reported in each study (including both in the narrative review and meta-analysis).

My research question was:

1. What are the procedures of translation and cultural adaptation reported in different Chinese-speaking populations?

3.4 The T-ACE-III study (addressed in *Chapter 7*)

For my empirical study, my main objective was:

1. To culturally adapt the Taiwanese version of ACE-III (T-ACE-III) and to gather validation data in Taiwan.

My research questions were:

1. What is the sensitivity, specificity, and the best cut-off scores of applying the T-ACE-III scale in a clinical context in Taiwan?
2. What is the diagnostic accuracy of T-ACE-III in differentiating dementia from non-dementia individuals in Taiwan?
3. What is the diagnostic accuracy of T-ACE-III in differentiating AD and non-AD from non-dementia individuals?

3.5 The supplementary tests of the T-ACE-III for individual cognitive domains and assessment of malingering (addressed in *Chapter 8*)

For my study on the supplementary tests, my objective was:

1. To explore the acceptability and performance of supplementary cognitive tests in Taiwan - The Brixton Spatial Anticipation Test, tests that made up the Processing Speed Index from the Wechsler Adult Intelligence Scale IV (WAIS-IV-PSI), and The Coin-in-the Hand (CIH) test.

My research questions were:

1. How acceptable are these tests to people who came to a neurology clinic in North Taiwan?
2. What is the level of performance on these tests among people in dementia and the non-dementia groups?

Chapter 4. Systematic review and meta-analysis of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia

In this chapter, I will describe how I carried out my systematic review and meta-analysis of brief (taking ≤ 20 minutes) cognitive assessments used in detecting suspected dementia in Chinese-speaking populations, and the findings. The results of my narrative review are discussed in *Chapter 5*.

4.1 Introduction

There is no “best test” which can be used in all situations for a diagnosis of dementia, as the psychometric properties including cultural adaptation and acceptability vary depending on the clinical setting, the population, and the individual (e.g., educational level, age). Selecting the most appropriate tests to determine the presence or absence of dementia has an impact on the likelihood and accuracy of diagnosis and the gateway to intervention. Chinese-speaking populations have distinct education, ethnicity, and culture compared to western countries, where the majority of tests were initially validated and implemented (Li et al., 2016; Phillips et al., 2015; Shanhu et al., 2019; Shim et al., 2017). Therefore, it is essential to determine which test is most appropriate for the population. To my knowledge, neither a systematic review nor an analytic study has been previously undertaken on the topic of the best brief cognitive tests for Chinese-speaking individuals with suspected dementia in clinical settings where the focus is on detection of dementia and MCI.

4.1.1 A lack of systematic reviews of brief cognitive assessments to be used with dementia and MCI in Chinese-speaking populations

When I carried out a literature search for my systematic review on the topic of brief cognitive assessments used in diagnosing patients with suspected dementia

and MCI in April 2019, I found out that the only one systematic review (Velayudhan et al., 2014) and one meta-analysis (Tsoi et al., 2015) included tests that could be useful in clinical settings, and which took less than 20 minutes. The previous study (Velayudhan et al., 2014) focused on English-speaking individuals and concluded that conducting validation studies on existing tests is more important than designing new ones. The study also found that the ACE-R was useful in detecting all forms of dementia while the ACE-III showed promise but needed more validation in a range of settings (e.g., general neurology or memory clinics). The latter study (Tsoi et al., 2015) confirmed that Mini-Cog and ACE-R performed the best in cases of dementia, whereas the Mini-Cog performed the best for MCI.

After I started my systematic review study, another systematic review, carried out with Spanish speaking individuals on the same topic, concluded that diagnostic accuracy of brief cognitive screening tools should be explored in different settings (e.g., community, memory clinic), and for different stages and types of dementia (Custodio et al., 2020). One year later, there was a further systematic review and meta-analysis (Huo et al., 2021). In Chinese-speaking individuals, Huo et al. examined dementia screening tools (including neuropsychological and cognitive tests) which took a longer administration time than the 20 minutes time limit that I adopted (≤ 35 minutes). However, in contrast to my study, the review conducted by Huo et al. did not include MCI, but only studied tests used for screening purposes, and covered all settings, rather than separating out what were the best instruments in different settings (e.g., hospital, community, care home). My own review considered brief instruments which were used as part of the diagnosis process where the people with dementia presented to secondary care, and those without dementia were from either secondary care or community settings.

These systematic review studies shed slight on the issue of clinical context. Study participants were selected from communities, residential care homes, hospital wards, or from secondary care settings – e.g., memory clinics, dementia clinics, neurology clinics. The prevalence of dementia should be higher in secondary care than in other clinical settings; hence, cognitive tests administered in secondary

care should serve different purposes than those administered in other clinical settings, which should be investigated individually.

4.1.2 Differences between screening tests and diagnostic tests

By using a screening test with a high sensitivity and a cut-off point in the initial stage, only a few cases are likely to be missed, despite a large number of false positives. This approach results in a later screened population with a prevalence of the condition that is closer to the desired 50%. Therefore, a diagnostic test with a high specificity and a cut-off point to eliminate false positives is more effective in the second stage (Wang et al., 2003). However, it is essential to use these tests correctly as both are used to determine the presence of a disease. The distinction between the two is based on how they are employed, not how they were created, namely the purpose of the test (Wang et al., 2003).

Typically, diagnostic tests are administered when a person is suspected of having a certain disease, and their purpose is to confirm or refute this diagnosis (Wang et al., 2003). In contrast, screening tests are administered to a larger, usually asymptomatic population with a higher risk to determine which of them could have the disease in question (Wang et al., 2003). Those tests should be simple to conduct, acceptable to patients, sensitive, and capable of detecting a curable condition or a disorder for which treatment improves outcome (Wilson et al., 1968). In an effort to discover true positives and prevent missing cases, physicians may seek to maximise sensitivity of screening tools (Gilbert et al., 2001).

4.1.3 Aims

I therefore set out to conduct the first systematic review and meta-analysis of studies found in English databases that used brief cognitive assessments as cognitive tests for diagnosis with people suspected of having dementia and presenting to secondary care in Chinese-speaking populations. Subsequently, I

also aimed to apply the findings from the meta-analysis (present chapter, *Chapter 4*) and narrative review (following chapter, *Chapter 5*) to choose a test to validate and adapt an optimal cognitive battery for use with Taiwanese populations, and this is described in *Chapter 6*. My research question was:

- What are the brief cognitive assessment tools with the best sensitivity and specificity in detecting suspected dementia (dementia, MCI) across different Chinese-speaking populations?

4.2 Methods

4.2.1 Research protocol

I registered the protocol on PROSPERO (registration number: CRD42019134092, see *Appendix 1* for a copy of the submitted protocol) and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). I also followed the PRISMA-diagnostic test accuracy (PRISMA-DTA) checklist (McInnes et al., 2018) and the guidelines from the Cochrane Diagnostic Test Accuracy Working Group (Deeks et al., 2022).

4.2.2 Data sources

I searched seven databases [Ovid MEDLINE®, PsychInfo, PsycTESTS, and Web of Science core collection, The Cochrane Library (Cochrane Library of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), and Cochrane Methodology Register)] from inception to July 3rd, 2019. I updated my search until December 1, 2020, and then until October 12, 2021.

4.2.3 Search strategy

I searched abstract, keywords and titles using the search terms: (“Chinese” OR “Mandarin” OR “Hokkien” OR “Hoklo” OR “Cantonese” OR “Hakka” OR “Taiwan*” OR “China” OR “Hong Kong*” OR “Singapore*” OR “Macao” OR “Malaysia) AND (“Alzheimer*” OR “AD” OR “dement*” OR “VaD” OR “FTD” OR “Mild cognitive impairment” OR “MCI” OR “memory loss”) AND (all fields) (“assessment” OR “evaluation” OR “scale” OR “test” OR “tool” OR “Instrument” OR “battery” OR “measure*” OR “screen*” OR “diagnos*” OR “inventory*” OR “validat*”).

4.2.4 Study selection

I included studies that fulfilled the following criteria:

- Validation studies of instruments to assess cognition as part of diagnosis (not for the purpose of screening) for patients with any suspected dementia or MCI in a memory clinic or similar setting where people presented with suspected dementia
- Chinese-speaking populations
- Time taken ≤20 minutes (either stated in paper or other studies had stated instrument time ≤ 20 minutes)
- Patient assessed face-to-face
- Published in peer reviewed journal
- Published in English or Chinese

I excluded studies if:

- The instruments were used to test everyday function or behaviour
- They were used in the setting where the dementia prevalence is lower than secondary care (e.g., communities, day care centres) as a screening tool rather than for diagnosis in people presenting with suspected dementia.
- They used telephonic or computerized self-administration tests
- They were informant questionnaires
- They were designed to detect dementia praecox or dementia secondary to head injury

- They were for people with diseases other than dementia and MCI
- Their purpose was to stage severity of rather than testing for suspected dementia or MCI
- They were intended for the learning disability population
- The studies reported qualitative results only
- They did not report cut-off scores, sensitivities, and specificities

4.2.5 Data screening and selection

I exported all searches to Endnote and removed duplicates. Then, I screened the titles and abstract of retrieved articles, and I excluded papers if they were irrelevant. Subsequently, the second rater, a clinical psychologist working in Taiwan - Jen-Chieh Lai (J-CL) joined the process of full-text screening. We both independently assessed the remaining articles for inclusion, based on the eligibility criteria (see *4.2.4 Study selection*). We resolved the disagreements about selection of studies by discussing with my supervisors until we reached a consensus.

4.2.6 Data extraction and definition

J-CL and I then independently extracted data from each paper that was included, using a standard format. We recorded data on population, recruitment settings, specification of illness, study design, purpose of the test, time taken to administer test, total items, total scores, cut-off score, sensitivity, specificity, validity, reference standard, and blinding. We compared our ratings and reached a consensus in case of any disagreements, and we were then able to agree on all the remaining papers. I went through the same screening, selection, and extraction procedures alone when I conducted the second- and third-time update searches.

I contacted the study authors if the potential studies omitted a piece of essential information to be included in the systematic review (e.g., recruitment processes)

or the sub-group analysis of meta-analysis (e.g., numbers of participants, cut-off scores, sensitivity, specificity in subgroups).

4.2.7 Quality assessment

The Centre for Evidence-based Medicine (CEBM) criteria (Heneghan, 2009) for assessing the quality of studies of diagnostic tests ask for answers to five questions, which are scored as 0, 1 or 2, giving a total possible score of 10; higher scores indicate higher quality of study. However, we scored studies as 0 or 2 since we found it difficult to distinguish partially fulfilled from unfulfilled criteria. (See *Appendix 2: Operational criteria for the scoring of the Centre for Evidence-based Medicine (CEBM) diagnostics criteria* for the details scores given to included studies). The questions are:

1. Was the diagnostic test evaluated in a representative spectrum of patients?
2. Was the reference standard applied regardless of the index test result?
3. Was there an independent, blind comparison between the index test and an appropriate reference (‘gold’) standard of diagnosis?
4. Are test characteristics presented or possible to calculate (sensitivity, specificity, positive and negative predictive value)?
5. Were the methods for performing the test described in sufficient detail to permit replication?

The second rater (J-CL), the third rater (Esther Hui, E-H), and I independently assessed the quality of each study using the CEBM diagnostics criteria. I was responsible for all the included studies, and J-CL and E-H were responsible for the studies with clinical and community-based controls, respectively.

4.2.8 Meta-analysis

Meta-analyses of diagnostic tests often report several measures that help decide whether a test is “good”. *Table 3* shows the most commonly reported measures in the meta-analyses of diagnostic test accuracy.

Random-effect univariate and bivariate analysis are the two main approaches to calculate the summary likelihood ratio (LR) measures, in which the former independently summarises the sensitivity and specificity values, the latter explicitly accounts for the (negative) correlation between sensitivity and specificity, resulting in less heterogeneity than univariate analysis.

Table 3: Definition of common measures of test accuracy

Test accuracy indicators	formula	Definition
Sensitivity	TP / (TP + FN)	Proportion of patients with dementia correctly identified
Specificity	TN / (FP + TN)	Proportion of those without dementia correctly identified
Positive predictive value	TP / (TP + FP)	Proportion of those with dementia among the test positives
Negative predictive value	TN / (FN + TN)	Proportion of those without dementia among the test negatives
Positive likelihood ratio	(TP / TP+FN) / (FP/ FP+TN)	Ratio of the proportion of test positives among those with dementia to the proportion of test positives among those without
Negative likelihood ratio	FN / (TP+FN) / TN / (FP+TN)	Ratio of the proportion of test negatives among those with dementia to the proportion of test negatives among those without
Diagnostic odds ratio	(TP × TN) / (FP × FN)	Ratio of the odds of positivity in those with true dementia to the odds of positivity in those without

Note. TP = true positive; TN = true negative; FP = false positive; FN = false negative

4.2.8.1 Main analysis

In the meta-analysis, I included tests that were used in more than one study (multiple cut-off scores, sensitivities, and specificities). As certain articles had more than one test, the number of studies can be higher than the number of articles, and the same articles may be included in different subgroup analyses.

I conducted the meta-analysis on two separate groups according to disease classification (dementia vs MCI) and I separately reported the best cut-off scores according to where the non-dementia group were recruited; the latter could be clinical controls vs community-based controls: the patients' spouse, relatives, friends or caregivers, as well as those recruited from the communities. I did this because the populations were likely to be different, and cut-off scores required for distinguishing the two groups would vary.

I used R software version 4.1 for all statistical analyses. I analysed sensitivity and specificity of each short cognitive test separately using a univariate random-effects model to form the summary statistics. Through applying the package "meta" and function "metaprop", the logit transformation and Clopper–Pearson method were used. First, I calculated the total effect size using the number of events and the number of samples from proportion-type data. Second, I used the forest plots to graphically present the combined sensitivity and specificity. Third, I assessed the statistical heterogeneity among the trials with I^2, which describes the percentage of total variation across studies due to the heterogeneity. I considered an I^2 value (Higgins et al., 2021):

- 0% to 40% might not be important.
- 30% to 60% may represent moderate heterogeneity.
- 50% to 90% may represent substantial heterogeneity.
- Over 75% indicates considerable heterogeneity and I took this into account when considering generalisability of the results; >75% suggests that the results are not generalisable.

There are a couple of factors that can influence the I^2 value (Higgins et al., 2022):

(1) Magnitude and direction of effects.
(2) Strength of evidence for heterogeneity, such as the P value from the Chi^2 test, a confidence interval for I^2, when the number of studies is small, there is a considerable level of uncertainty in the I^2 value.

In general, 75% of sensitivity or specificity is not good enough to rule out a disease, the sum of sensitivity and specificity should be at least 150% (Power et al., 2013). I therefore defined either sensitivity or specificity of individual papers

of meta-analysed tests below 75% as ‘unacceptable’, 75-90% as ‘satisfactory’, and higher than 90% as ‘excellent’.

4.2.8.2 Other analysis

4.2.8.2.1 Summary of diagnostic accuracy

I generated hierarchical summary receiver operating characteristic (HSROC) and summary receiver operating characteristic (SROC) plots with the command of “metandi” and “metandiplot” on Stata, and “reitsma” on R software (when number of studies <3) to graphically display how the summary sensitivity changes with the summary specificity and vice versa.

4.2.8.2.2 Sensitivity analysis on univariate and bivariate models

As there is debate in the literature as to the best method for meta-analysing data on diagnostic test accuracy, I conducted a sensitivity analysis to compare univariate and bivariate methods by conducting a bivariate random-effects model (Rutter & Gatsonis, 2001) on Stata/MP 17.0 with the command of “metadta”.

4.2.8.2.3 Meta-regression

In addition, I conducted a random effects meta-regression on Stata/MP 17.0 to consider the causes for heterogeneity (population, subtype, reference standard, scoring system), with the command “gsort” on Stata (Nyaga & Arbyn, 2022).

4.2.8.2.4 Publication bias

To investigate publication bias, I used R software version 4.1 to apply Deeks’ funnel plot asymmetry test (Deeks et al., 2005) for diagnostic test accuracy meta-analysis (Macaskill et al., 2010). This conducts a regression of

diagnostic log odds ratio against 1/sqrt (effective sample size), with $p < 0.05$ for the slope coefficient indicating significant asymmetry.

4.3 Results

4.3.1 Study selection

I retrieved 19,722 references, see *Figure 1* for the PRISMA diagram (Page et al., 2021) and *Appendix 3* for PRISMA-DTA checklist (McInnes et al., 2018), out of which 37 studies met the inclusion criteria (see *Table 4* for the characteristics of included studies). All the studies were found through databases searches, none through hand searches. Among all studies that met the inclusion criteria, 21 studies were analysed in the meta-analysis in this current chapter (*Chapter 4)*, 18 studies that used other brief cognitive assessments in a single Chinese-speaking population were reviewed in *Chapter 5* as a narrative review. All 37 studies underwent translation and cultural adaptation procedures, which were described in *Chapter 6*.

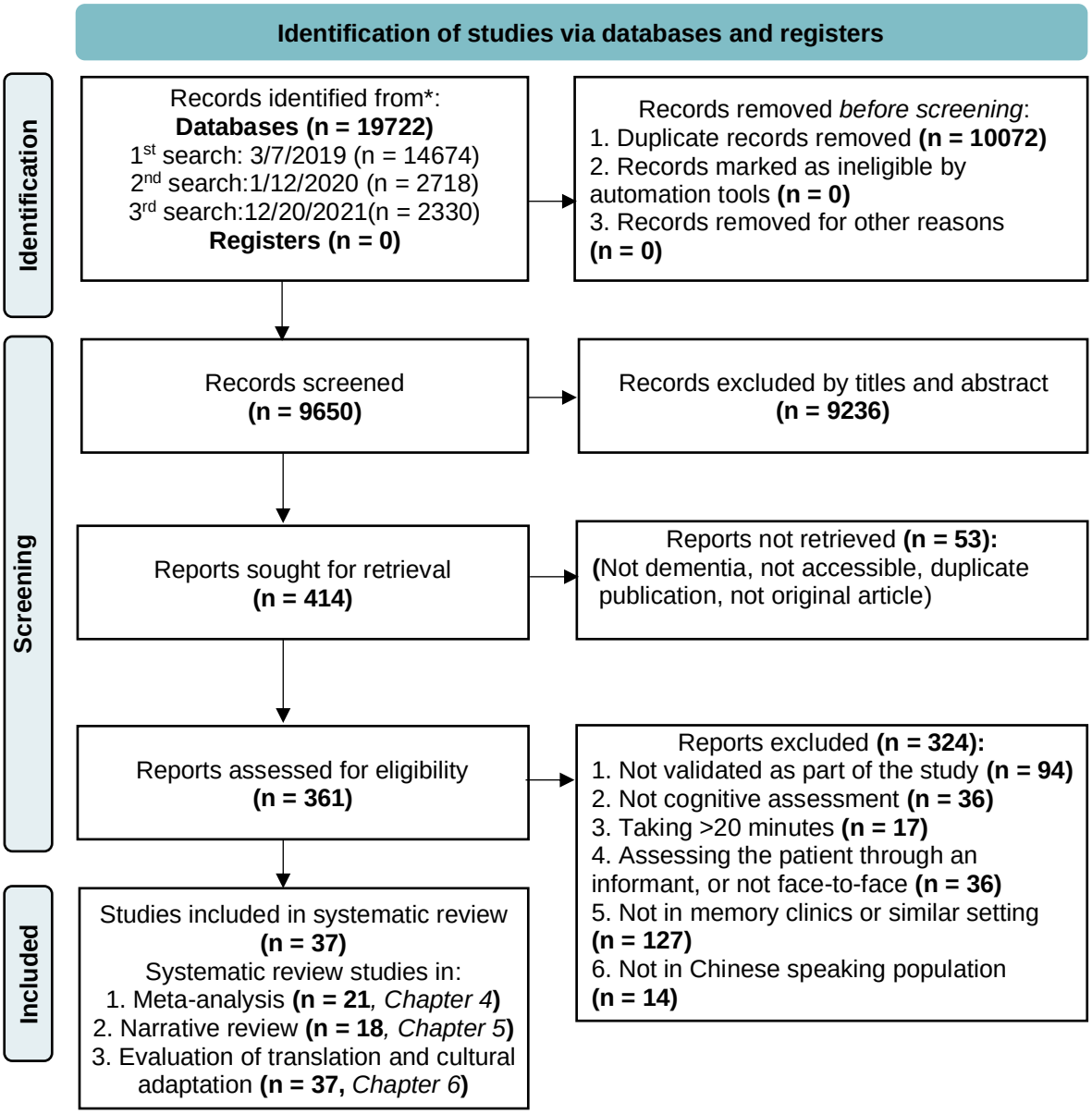


Figure 1: PRISMA flow diagram of references identified and included in the review

4.3.2 Study characteristics and quality analysis

Table 4 displays the included articles, instrument characteristics, the quality of each of the studies, which are grouped according to clinical settings and further grouped by country and cognitive test.

There were 15 studies that recruited all participants from clinical settings comprising eight, four, two, and one studies in China, Taiwan, Hong Kong, and Singapore, respectively. There were 22 studies that recruited controls from community settings, including 12, five, three, and two in China, Taiwan, Singapore, and Hong Kong, respectively.

None met the highest score of CEBM:

1. All studies (expect Chang et al., (2010)'s study) did not validate the diagnostic test in a representative spectrum of patients (like those in whom it would be used in practice), they excluded participants with specific conditions, such as eyesight problems, stroke, or psychiatric issues.
2. Two studies did not apply reference standard regardless of the index test results, they recruited participants from the dementia and non-dementia groups using different standard measurements (Chong et al., 2010; Liew et al., 2015).
3. One study was not blinded to other cognitive tests scores (Yeung et al., 2014).
4. Two studies provided no information about the gold standard (Li et al., 2018; Malhotra et al., 2013).
5. Three studies did not specify whether they were blinded or not (Chong et al., 2010; Huang et al., 2018; Zhang et al., 2019).
6. 14 studies provided insufficient details on the assessments (Chen et al., 2015; Chong et al., 2010; Fang et al., 2014; Lee et al., 2018; X. Y. Li et al., 2018; Li et al., 2021; X. D. Li et al., 2018; Low et al., 2020; Malhotra et al., 2013; Qian et al., 2021; Shi et al., 2012; Wen et al., 2008; Wong et al., 2013; Yeung et al., 2014).
7. The studies listed in 4, 5, 6 provided insufficient detail for replication.

Table 4: Characteristics of included studies and quality score (Arranged by country, setting, test name)

SETTINGS, TEST CODE & AUTHOR			TEST NAME	ILLNESS	SETTING	TIME (MIN)	TOTAL SUB-TESTS OR ITEMS	TOTAL SCORE	IRR	TRR	CE BM	ANALYSIS	
	C1	Li et al., 2021	FAB-Phonemic	AD, aMCI, naMCI	Memory clinics	5	6 items	18	0.997	0.819	6		△
	C2	Shi et al., 2012	HVLT	dementia, AD, aMCI	Memory clinics	NA	3 free recall trails 1 recognition trial	36	NA	NA	6		△
	C3	Yang et al., 2019	[M-ACE], MMSE	mild dementia, MCI	Neurology department	<5	4 domains	30	NA	NA	8	✓	△
	C4	Li et al., 2018	Mini-Cog	MCI	Neurology department	3	5 domains	9	NA	NA	4		△
	C5	Wen et al., 2008	MoCA*	MCI	Neurology department	<10	11 subtests	30	NA	NA	4	✓	
	C6	Li et al., 2018	MoCA	amMCI, asMCI, naMCI,	Memory clinics	NA	NA	30	NA	NA	4		
	C7	Xu et al., 2003	MMSE	dementia	Neurology department	15	12	NA	NA	0.75	8	✓	
A	C8	Wei et al., 2018	TMT (A & B)	VaD, MCI, VaMCI	Memory clinics	150/ 300 secs	2 tasks	NA	NA	NA	6		△
	C9	Wang et al., 2017	ACE-III	dementia	Memory clinics NC: Physical examination centres	NA	5 domains	100	NA	NA	8	✓	
	C10	Fang et al., 2014	ACE-R	mild AD, aMCI	Memory clinics	NA	5 domains	100	0.994	0.967	4	✓	
	C11	Guo et al., 2002	AFT	AD, aMCI	Memory clinics NC: Community	1	1 domain	unlimited	NA	NA	4		△
B	C12	Guo et al., 1991	BNT	AD (mild, moderate), aMCI	Memory clinics NC: Community	NA	30 items	30	NA	NA	4		△

SETTINGS, TEST CODE & AUTHOR			TEST NAME	ILLNESS	SETTING	TIME (MIN)	TOTAL SUB-TESTS OR ITEMS	TOTAL SCORE	IRR	TRR	CE BM	ANALYSIS	
	C13	Qian et al., 2021	DRS-CV	AD, aMCI	Memory clinics NC: Urban centres	NA	5 domains	144	NA	NA	6		
	C14	Guo et al., 2004	MDRS	Mild AD	Memory clinics NC: Community	15	5 domains	144, 124 (illiterate)	NA	NA	4	✓	
	C15	Chen et al., 2016	MoCA-BC	AD, MCI	Memory clinics NC: Community	NA	9 domains	30	0.96	NA	8	✓	
	C16	Huang et al., 2018	MoCA-BC	AD (mild, moderate), MCI	Memory clinics NC: Spouses and friends	NA	9 domains	30	NA	NA	6	✓	
	C17	Zhang et al., 2019	5 items from MoCA-BC	AD, MCI	Memory clinics NC: Caregivers	NA	5 items	NA	NA	NA	4		
	C18	Guo et al., 2010	[QCST], MoCA, MMSE	mild AD, MCI	Memory clinics NC: Community	8-15	9 items	90	NA	0.93	6	✓	△
	C19	Huang et al., 2019	[silhouettes-A], MoCA, STT-B, & other tests	AD, MCI	Memory clinics NC: Relatives and friends	NA	8 animals & 7 inanimate objects	15	0.99	0.91–0.98	8	✓	△
	C20	Zhao et al., 2013	[STT-A], [STT-B]	AD, aMCI	Memory clinics NC: Community	Around 5	2 subtests	NA	NA	NA	8/8	✓	
	H1	Chan et al., 2005	[CDT], [T&C]	dementia	Memory clinics	90.9, 65.6 secs	1 item2 items	10, 1/0	NA	NA	8/8	✓	△
A	H2	Yeung et al., 2014	[MoCA*], MMSE	dementia, MCI, CI	Memory clinics NC: Geriatrics department	10-15	7 domains	30	0.987	0.92	4	✓	
	H3	Wong et al., 2013	[ACE-R], MMSE	dementia, MCI, CI	Memory clinics NC: Spouses	Around 15	5 domains	100	1	1	6	✓	
B	H4	Chu et al., 2015	[MoCA*], MMSE	AD, aMCI	Memory clinics NC: community social centres	10	8 domains	30	0.96	0.95	4	✓	
A	S1	Malhotra et al., 2013	SPMSQ	dementia, MCI	Outpatient cognitive assessment clinics	NA	10 items	10 (errors)	NA	NA	4		△

SETTINGS, TEST CODE & AUTHOR			TEST NAME	ILLNESS	SETTING	TIME (MIN)	TOTAL SUB-TESTS OR ITEMS	TOTAL SCORE	IRR	TRR	CE BM	ANALYSIS
	S2	Chong et al., 2010	[FAB]	CI (dementia, MCI)	Memory clinics NC: Community (SLMS programme)	5	6 domains	NA	NA		2	△
	S3	Liew et al., 2015	MoCA	Major and Mild NCD	Memory clinics NC: Community (SLMS programme)	NA	NA	NA	NA	NA	2	✓
B	S4	Low et al., 2020	VCAT	AD, CI (AD, MCI)	Memory clinics NC: Community	15.7	11 items	30	NA	NA	4	△
	T1	Lin et al., 2003	CDT	AD, MCI	Memory clinics	NA	3 items	3	0.99	0.88-0.9	6	✓
	T2	Chiu et al., 2008	CDT (Rouleau scoring system)	mild AD, QD	Memory clinics	NA	2 subtests	11/10	0.83/0.87	NA	8	△
	T3	Tsai et al., 2016	MoCA*, MMSE	dementia, MCI	Geriatric psychiatry, neurology, and geriatric departments	15	12 items	30	NA	0.92	6	✓
A	T4	Chen et al., 2015	RUDAS	AD	Neurology, geriatric, and geriatric psychiatry departments	10	6 categories	30	0.88	0.9	6	△
	T5	Tsai et al., 2018	BHT-Cog part	dementia, MCI	Neurology department NC: Volunteers, community (The Yilan Study)	4	5 items	16	NA	NA	4	△
	T6	Chang et al., 2010	CVVLT	MCI-VaD, MCI-AD	Memory clinics NC: Volunteers, spouses	NA	4 trials 3 recall trials	NA	NA	NA	8	△
	T7	Chang et al., 2012	[MoCA*], MMSE	Very mild AD	Neurology department NC: Normative database during routine examinations	NA	7 domains	30	NA	NA	6	✓
	T8	Tsai et al., 2012	MoCA	AD, MCI	Memory clinics NC: Relatives or volunteers	10	7 domains	30	0.88	0.88	6	✓
B	T9	Lee et al., 2018	[Qmci], MoCA*, MMSE	dementia, MCI	Neurology department NC: convenience sampling	<5	6 subtests	100	0.87	1.00	6	✓ △

Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively; NA = Not available (the information is not provided in the study); The setting of A/ B refers to clinical or community-based controls, respectively; The test in bracket is the main validated test; * aside MoCA

refers to the adjusted scoring system; ✓ and △ refer to the studies included in the meta-analysis and the narrative review, respectively; MCI = Mild Cognitive Impairment; CEBM = the Centre for Evidence-based Medicine diagnostics criteria. The tests in the brackets are the target tests to be validated. ACE-R = Addenbrooke's Cognitive Examination Revised; AFT = Animal Fluency Test; BHT-cog = Brain Health Test-Cog part; BNT = Boston Naming Test; CDT = Clock Drawing Test; CFT-C = Rey-Osterreith Complex Figure Test-Copy; CVVLT = Chinese version of the Verbal Learning Test; DRS/MDRS = Mattis dementia rating scale; FAB = Frontal Assessment Battery; HVLT = Hopkins Verbal Learning Test; JLO = Judgment of Line Orientation; M-ACE = Mini-Addenbrooke's Cognitive Examination; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; MoCA-BC = Montreal Cognitive Assessment Basic; QMCI = Quick Mild Cognitive Impairment Screen; RUDAS=Rowland Universal Dementia Assessment Scale; SPMSQ = Short Portable Mental Status Questionnaire; STT = Shape Trail Test; T&C = Time and Change Test; TMT = Trail-Making Test; VCAT = Visual Cognitive Assessment Test.

4.3.3 Instruments in meta-analyses

There were 21 articles included in the meta-analysis: 15 studies in the population of dementia and 14 studies in the population of MCI. As some articles validated tests for both MCI and dementia, the number of studies is higher than the number of articles. The tests that could be meta-analysed were:

- For dementia (7 tests): ***Addenbrooke's Cognitive Examination III & Revised (ACE-III &-R)***, ***Clock Drawing Test (CDT)***, ***Mattis Dementia Rating Scale (DRS)***, ***Mini Mental State Examination (MMSE)***, **and** ***Montreal cognitive assessment (MoCA), Shape Trail Test-A, and B (STT-A&B)***
- For MCI (4 tests): ***ACE-III &-R, MMSE, MoCA, Montreal cognitive assessment-Basic (MoCA-BC)***

4.3.3.1 Outline of tests included in the meta-analysis

I have set out below a brief description of each test, covering cognitive domains sampled and time taken to administer:

- ***ACE-R & ACE-III*** were both developed from an earlier version of the Addenbrooke's Cognitive Examination (ACE); the former includes all MMSE items, while the latter does not. Both assess attention and orientation, memory, fluency, language, and visuospatial abilities. Both take < 20 minutes (Hodges & Larner, 2016).
- ***Clock Drawing Test (CDT)*** covers aspects of planning and visuospatial ability; there is no time limit, but it usually takes a short time to administer (Lin et al., 2003).
- ***Dementia Rating Scale (DRS)*** includes 37 tasks divided into five domains of attention, initiation/perseveration, construction, conceptualization, and memory (Qian et al., 2021), and takes around 15 minutes.
- ***Mini Mental State Examination (MMSE)*** consists of tasks assessing orientation, attention, calculation, memory, language, and visual construction (Tsai et al., 2016), taking 15-20 minutes.

- ***Montreal Cognitive Assessment (MoCA)*** covers the same cognitive domains as MMSE but adds two more aspects - executive functions and conceptual thinking domains (Tsai et al., 2016), taking 15-20 minutes.
- ***Montreal Cognitive Assessment-Basic (MoCA-BC)*** taps executive function, language, orientation, calculation, conceptual thinking, memory, visuo-perception, and attention (Chen et al., 2016), taking approximately 15 minutes.
- ***Shape Trail Tests (STT-A & B)*** are a variant of the original Trail Making Test. SST-A taps language and attention, while the STT-B deals more with executive function and memory (Zhao et al., 2013), taking around 5 minutes for both tests.

Appendices 4 and *5* provide details of the characteristics of studies that examined patients who had a diagnosis of dementia or MCI – e.g., setting, subtype, reference standard, numbers of participants, cut-off scores, sensitivity and specificity, and positive and negative likelihood ratio (+LR, -LR). *Appendix 6* provides a 2 x 2 table on the numbers of true positive, false negative, true negative, and false positive. *Appendix 8* provides information on the overall performance of the eight tests (HSROC & SROC plots). In the following sections, I present the diagnostic performance of each test according to where the non-dementia group was recruited, since different populations may be associated with different cut-off scores.

4.3.3.2 Diagnostic test accuracy and optimal cut-off scores for all cause dementia

As shown in *Figure 2*, the combined sensitivity (I^2) and specificity (I^2) of MoCA were 91.9% (67%) and 94.3% (22%). Using an adjusted MoCA scoring system, the best cut-off scores for identifying dementia from clinical controls were 18/19 in Hong Kong (Yeung et al., 2014) and 20 in Taiwan (Tsai et al., 2016). When controls were recruited from the community, it had the best cut-off score of 19/20 for AD in Hong Kong (Chu et al., 2015). Using non-adjusted scoring system, the recommended cut-off scores were 21/22 for AD in Taiwan (Tsai et al., 2012), as

well as 20/21 (years of education<6) and 22/23 (years of education > 6) for major neurocognitive disorder (NCD) in Singapore (Liew et al., 2015).

For MMSE, the combined sensitivity (I^2) and specificity (I^2) were 92.4% (50%) and 89.8% (50%), respectively. When the controls were recruited from clinics, the suitable cut-off score was 24/25 for dementia in Hong Kong (Yeung et al., 2014); it was 22 and 20 for dementia in the population with or without education in China, respectively (Xu et al., 2003), and 23/24 and 25/26 for AD (Fang et al., 2014) and mild dementia (Yang et al., 2019) in China. Compared to the community-based controls, the optimal cut-off score was 26/27 for mild AD in Taiwan (Chang et al., 2012), and 24/25 for AD (Chu et al., 2015) and 25/26 for all-cause dementia (Wong et al., 2013) in Hong Kong.

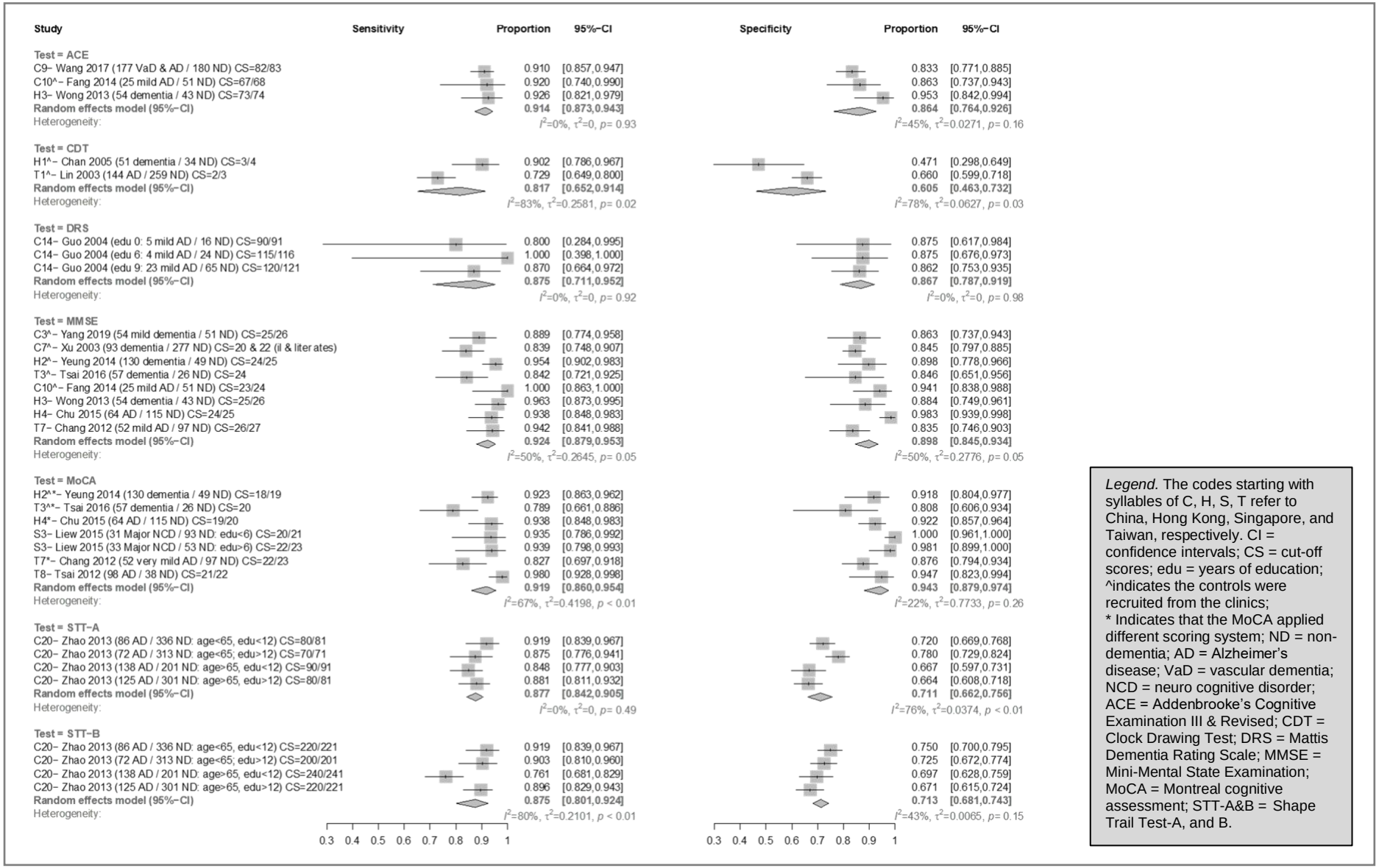
For ACE (ACE-III & ACE-R), the combined sensitivity (I^2) and specificity (I^2) were 91.4% (0%) and 86.4% (45%). In China, when controls were recruited from clinics, the appropriate ACE-R cut-off score for mild AD was 67/68 (Fang et al., 2014). Comparing with the community-based controls, ACE-III's best cut-off score was 83 for VaD and AD in China (Wang et al., 2017); ACE-R's optimal cut-off score was 73/74 for dementia in Hong Kong (Wong et al., 2013).

For DRS, the combined sensitivity (I^2) and specificity (I^2) were 87.5% (0%) and 86.7% (0%) in identifying dementia from community-based controls, with cut-off scores of 90/91, 115/116 and 120/121 for AD with zero, six, and nine year(s) of education in China (Guo et al., 2004), respectively.

For STT-A and STT-B, the combined sensitivity (I^2) and specificity (I^2) were 87.7% (0%) and 71.1% (76%), 87.5% (80%) and 71.3% (43%) in identifying dementia from community-based controls, respectively. For AD, STT-A and -B best cut-off scores were 80 and 220 for those aged < 65 with <12 years of education, 70 and 200 for those aged < 65 with >12 years of education, 90 and 240 for those aged > 65 with <12 years of education, 80 and 220 for those aged > 65 with >12 years of education in China (Zhao et al., 2013).

For CDT, the combined sensitivity (I^2) and specificity (I^2) were 81.7% (83%) and 60.5% (78%) in identifying dementia from clinical controls, with optimal cut-off

scores of 3/4 for dementia in Hong Kong (Chan et al., 2005) and 2/3 for Alzheimer's disease (AD) in Taiwan (Lin et al., 2003).



Legend. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. CI = confidence intervals; CS = cut-off scores; edu = years of education; ^indicates the controls were recruited from the clinics;
* Indicates that the MoCA applied different scoring system; ND = non-dementia; AD = Alzheimer's disease; VaD = vascular dementia; NCD = neuro cognitive disorder; ACE = Addenbrooke's Cognitive Examination III & Revised; CDT = Clock Drawing Test; DRS = Mattis Dementia Rating Scale; MMSE = Mini-Mental State Examination; MoCA = Montreal cognitive assessment; STT-A&B = Shape Trail Test-A, and B.

Figure 2: Meta-analyses for diagnostic test accuracy on diagnosing Dementia

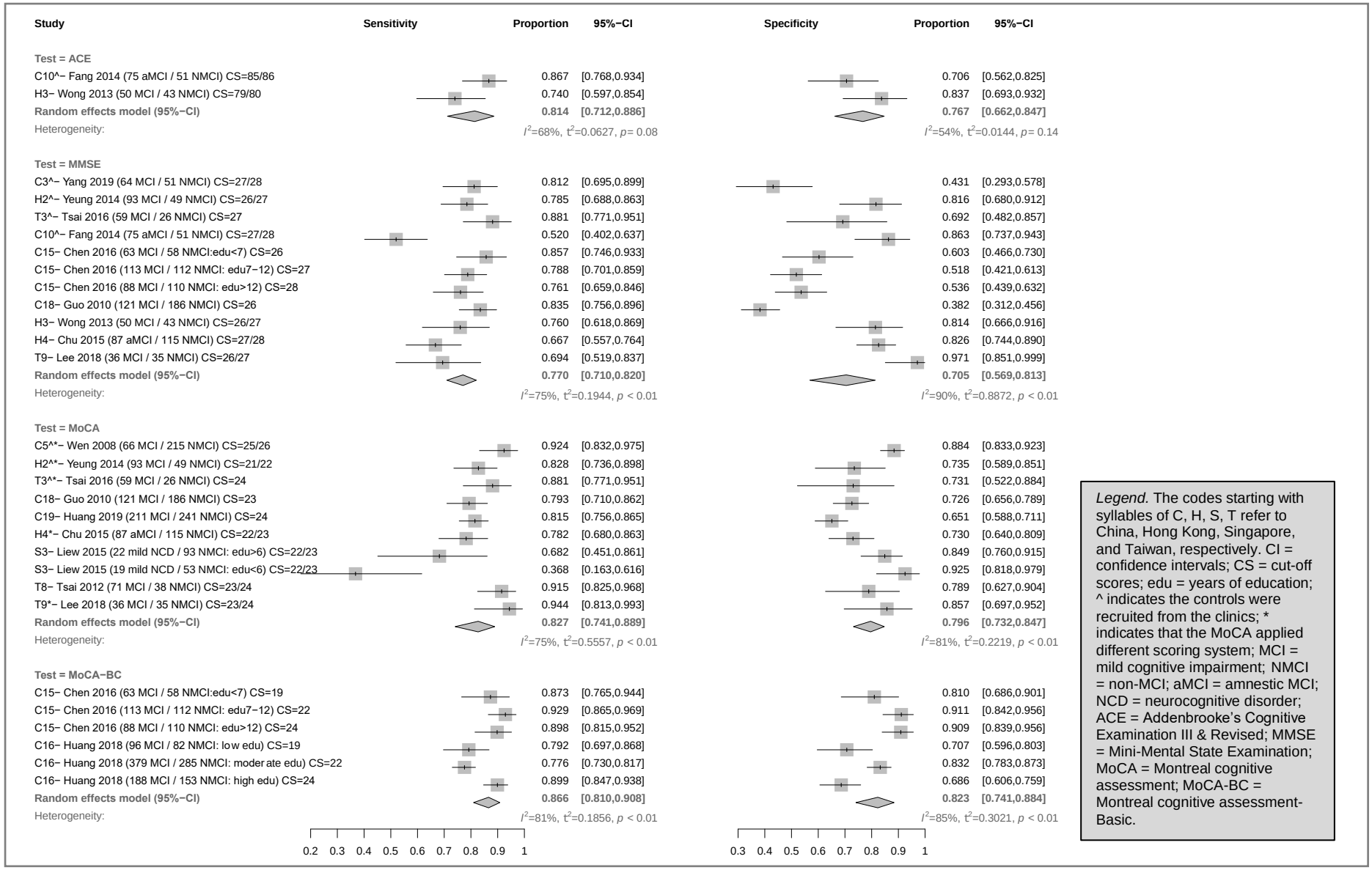
4.3.3.3 Diagnostic test accuracy and optimal cut-off scores for MCI

As indicated in *Figure 3*, for the ACE (ACE-R), the combined sensitivity (I^2) and specificity (I^2) were 81.4% (68%) and 76.7% (54%), with cut-off scores of 85/86 for differentiating aMCI from clinical controls in China (Fang et al., 2014), and 79/80 for recognising MCI from community-based controls in Hong Kong (Wong et al., 2013).

For MoCA-BC, the combined sensitivity (I^2) and specificity (I^2) were 86.6% (81%) and 82.3% (85%), with optimal cut-off scores of 19, 22, 24 for detecting MCI with low (<7), moderate (7-12), and high (>12) education levels from community-based controls in China, respectively (Chen et al., 2016).

For MoCA, the combined sensitivity (I^2) and specificity (I^2) were 82.7% (75%) and 79.6% (81%). With MoCA adjusted scoring system, the suitable cut-off scores were 25/26, 21/22, 24 for differentiating MCI from clinical controls in China (Wen et al., 2008), Hong Kong (Yeung et al., 2014), and Taiwan (Tsai et al., 2016), respectively. When it comes to community-based controls, the best cut-off scores applying original system were 23 and 23/24 for MCI in China (Guo et al., 2010) and Taiwan (Tsai et al., 2012), 22/23 for mild NCD with all education levels in Singapore (Liew et al., 2015). When applying MoCA adjusted scoring system, 22/23 for aMCI in Hong Kong (Chu et al., 2015) and 23/24 for MCI in Taiwan (Lee et al., 2018) were recommended as cut-off scores.

For MMSE, the combined sensitivity (I^2) and specificity (I^2) were 77.0% (75%) and 70.5% (90%), with best cut-off scores of 27/28 in China (Yang et al., 2019), 27 in Taiwan (Tsai et al., 2016), and 26/27 in Hong Kong (Yeung et al., 2014) when identifying MCI from clinical controls. While recognising MCI from community-based controls, the optimal cut-off scores were 26 in China (Guo et al., 2010), and 26/27 in Hong Kong (Wong et al., 2013) and Taiwan (Lee et al., 2018). The suggested cut-off scores for detecting aMCI were 27/28 in China (Fang et al., 2014) and Hong Kong (Chu et al., 2015) when compared to clinical and community-based controls, respectively.



Legend. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. CI = confidence intervals; CS = cut-off scores; edu = years of education; ^ indicates the controls were recruited from the clinics; * indicates that the MoCA applied different scoring system; MCI = mild cognitive impairment; NMCI = non-MCI; aMCI = amnestic MCI; NCD = neurocognitive disorder; ACE = Addenbrooke's Cognitive Examination III & Revised; MMSE = Mini-Mental State Examination; MoCA = Montreal cognitive assessment; MoCA-BC = Montreal cognitive assessment-Basic.

Figure 3: Meta-analyses for diagnostic test accuracy on diagnosing MCI

4.3.3.4 Sensitivity analysis of univariate and bivariate analysis

Appendix 7 shows the random-effect bivariate model. The confidence interval at 95% and the combined sensitivity and specificity were similar to the random-effect univariate model. In terms of heterogeneity, the random-effect bivariate model produces the generalised I^2 and Tau^2 values, which summarise the overall heterogeneity by taking into account the correlation between sensitivity and specificity. Some generalised I^2 and Tau^2 values were less than those of univariate analytic models. For instance, the I^2 values of sensitivity and specificity were 50% and 50% in MMSE, 67% and 22% in MoCA for dementia, whereas it was 30% in MMSE and 0% in MoCA for the generalised model. The I^2 values of specificity in STT-A and that of sensitivity in STT-B were 76% and 80% in dementia patients, respectively, but 0% and 33% in the generalised model. However, these differences did not affect the findings on the best-performing tests.

4.3.3.5 Meta-regression

I performed random-effect meta-regression on three tests (ACE, MMSE, and MoCA) to determine if the absolute sensitivity and specificity of each test differed by adding different covariates. Those covariates included clinical contexts (clinical controls vs community-based controls), populations (China vs other areas), subtypes (AD vs non-AD; MCI vs other types), reference standards (DSM vs others; Peterson's criteria vs others), and MoCA's scoring systems (original vs adjusted). *Appendices 9 to 13* provided a summary of likelihood ratio regression for each factor. The results showed that MMSE's specificity was significantly higher in other Chinese-speaking populations than in China (88% vs 70%), MMSE's sensitivity was significantly higher in MCI than other subtypes (80% vs 60%), and MoCA's original scoring system had significantly higher sensitivity (96% vs 88%) and specificity (98% vs 89%) than the modified scoring system. The remaining factors had no impact on any of the tests.

4.3.3.6 Publication bias

For the meta-analysis of diagnostic accuracy tests, the *P*-values for testing the asymmetry of the Deeks' funnel plots were 0.0054 and 0.6245 for patients with dementia and MCI, respectively (see *Appendix 14: Deek's funnel plots*). The findings revealed that publication bias exists in research in dementia populations but not in MCI populations.

4.4 Discussion

4.4.1 Summary of findings

This was the first systematic review and meta-analysis that focused on the use of brief cognitive assessment instruments with Chinese-speaking patients who have dementia or MCI. On the basis of the meta-analysis, the ACE was found to have the best validity, with sensitivity (I^2) and specificity (I^2) values of 91.4% (0%) and 86.4% (45%) in dementia, and 81.4% (68%) and 76.7% (54%) in MCI. The most commonly used test was MMSE, which was evaluated in 13 studies. None of the studies met the highest standard of CEBM.

4.4.2 Brief cognitive assessment instruments with the best psychometric properties

For the meta-analysis, I was able to compare the diagnostic accuracy of eight tests (ACE, CDT, DRS, MMSE, MoCA, MoCA-BC, STT-A, & B) in populations of dementia and MCI patients presenting to specialist diagnostic settings, such as neurology departments or memory clinics. ACE, DRS, MoCA, and MMSE all performed satisfactorily in detecting dementia, while the rest of the tests (CDT, STT-A & B) performed at an unacceptably low level (either sensitivity or specificity was lower than 75%) in detecting dementia.

When it comes to detecting MCI, MoCA-BC had the highest sensitivity and specificity among the three tests (MoCA-BC, MoCA, and ACE). However, when

comparing the heterogeneity among those tests, MMSE, MOCA, and MoCA-BC all had a high heterogeneity (>75%) in detecting MCI. ACE (ACE-R& III) had an acceptable heterogeneity (<75%) in detecting both dementia and MCI, with a slightly lower sensitivity and specificity.

4.4.3 Heterogeneity

Studies brought together in a systematic review can inevitably have a range of variability (or heterogeneity), such as clinical heterogeneity (variability in the participants, interventions, and outcomes), methodological heterogeneity: (variability in study design, outcome measurement tools, and risk of bias), and statistical heterogeneity (variability in the intervention effects being evaluated in the different studies and is a consequence of both heterogeneities above) (Egger et al., 2022). I conducted separate meta-analyses for two clinical populations (dementia, MCI) to reduce the clinical heterogeneity, sensitivity analysis of different statistical methods and meta-regression to explore the heterogeneity. If studies in a meta-analysis are heterogenous then it reduces the external validity.

I examined bivariate and univariate methods, and the results showed that both models produced similar confidence intervals at 95%, as well as the overall sensitivity and specificity. Regarding heterogeneity, only the ACE (ACE-R & ACE-III) had acceptable I^2 values in both dementia and MCI, independent of statistical methods (univariate or bivariate model). For dementia, the I^2 values of STT-A & B were more than 75% in the univariate model; however, they were lower than 75% in the bivariate model. For MCI, the I^2 values of MMSE, MoCA, and MoCA-BC were likewise higher than 75% in the univariate model but below 75% in the bivariate model (69%, 72%, and 74%) although they remained relatively high (see *Appendix 7: Random-effect bivariate model analysis*). Nevertheless, the differences did not affect the conclusion that ACE is the test with satisfactory sensitivity and specificity and acceptable heterogeneity among individuals with dementia and MCI. The results are in line with previous findings, indicating univariate and bivariate meta-analysis approaches yielded similar summary estimates and confidence intervals for sensitivity and specificity (Dahabreh et al.,

2017; Simel & Bossuyt, 2009), and that the difference is insufficient to alter clinical decision-making (Simel & Bossuyt, 2009). It is crucial to emphasise that head-to-head comparisons (identical cut-off points, age, and education groups) of meta-analyses on the topic of diagnostic test accuracy are vital, as differences in study design, sample, method, and clinical setting may have a confounding effect on the analysis results (Deeks et al., 2022; Huo et al., 2021).

It is important to note that a heterogeneity (I^2) higher than 75% obtained in the meta-analysis suggest the results are not generalisable. Also, the diagnostic accuracy of the tests may be unreliable if the reasons for high heterogeneity cannot be well understood. Due to the low number of studies (n = 2) in ACE for MCI, as well as CDT (n = 2) and DRS (n = 3) for dementia, the uncertainty in the I^2 value must be acknowledged, despite the fact that ACE is the best-performing test.

The result of the meta-regression found that the factors of population (China vs other areas), MCI subtype (MCI vs others), and the dementia scoring systems (original vs adjusted) contributed to some level of heterogeneity of MMSE and MoCA tests. Other factors could not be fully examined due to the limited studies included in the meta-analysis, such as the choice of cut-off scores, educational levels of participants. Future research with a larger sample size should evaluate those factors that may influence the diagnostic test accuracy for detecting dementia and MCI in Chinese-speaking-populations.

Publication bias was discovered in studies of individuals with dementia, but little is known about the mechanisms underlying publication bias in research looking at diagnostic test accuracy (Egger et al., 2022).

4.4.3 Strengths and weaknesses of the review

4.4.3.1 Strengths

This is the first comprehensive systematic review and the first meta-analysis to date on brief (<20 minutes) cognitive assessment instruments used with Chinese-

speaking population to evaluate patients with suspected dementia and MCI. I used a thorough search strategy and robust assessment of quality and data extraction by three independent reviewers in line with recommendations. I compared the outcomes between univariate and bivariate random-effect methods to confirm my findings. I examined the possible reasons of high heterogeneity through meta-regression. Also, I included not only English language papers, but also Chinese language papers identified in English databases, in which there were four papers published in simplified Chinese from China. I had no restriction on the date of publication, and updated the literature search twice, which allowed my review to gather results thoroughly and enables future research to readily update my findings.

4.4.3.2 Weaknesses

First, I did not include some representative Chinese databases, such as China National Knowledge Infrastructure (CNKI) for two main reasons. Firstly, the Chinese academic publishing system differs from western systems and faces unique challenges, such as the variable and disputed quality of Chinese journals (e.g., non-transparent editorial peer review and assessment processes), a lack of specialisation and focus in newly developing and specialised fields, and the problems of low-quality papers (e.g., broker selling papers and researcher purchasing authorship) (Wang et al., 2021). Secondly, the systematic review and meta-analysis of Chinese dementia screening tools revealed that the published language was one of the factors contributing to the high heterogeneity of MoCA tests, with Chinese-language publications demonstrating much higher sensitivity and specificity (Huo et al., 2021). With the assistance of my Chinese colleague, I was able to gain access to a few Chinese databases. However, I noticed that some studies lacked clarity in describing recruiting processes, such as “participants aged over 80 who attended our out-patient department or admitted to our hospital”, “all patients with bvFTD were registered in the CARDS study and diagnosed of bvFTD between January 2007 and November 2014”. Therefore, to maintain a consistent level of quality throughout the electronic resources, I

decided to include only English databases without language restrictions, allowing for the inclusion of high-quality Chinese publications.

Second, I did not involve any information specialist or medical librarian in the process of finding potential publications and developing a search strategy but had attended several database researching courses run by UCL library.

4.4.4 Future research

I only considered dementia as a general diagnosis and did not consider specific instruments used with specific forms of dementia; future research can consider examining the cognitive tests that are used to differentiate or stage dementia patients. Lastly, I only included studies which recruited patients with dementia or MCI from memory clinics or neurology departments; therefore, those studies which recruited dementia participants partly from communities or health centres were excluded. Future research should evaluate the tests administered in these various community-based settings independently in order to apply the results to a specific context.

4.5 Conclusions

For Chinese-speaking populations, the ACE (ACE-R & ACE-III) had the best psychometric properties among all the brief cognitive assessments used to detect dementia and MCI in clinical settings.

I used the findings from my systematic review and meta-analysis to inform my empirical study in Taiwan. In the next chapter, I report the features of brief cognitive assessments that were not included in the meta-analysis, and I score the procedures of translation and cultural adaptation reported in each study.

Chapter 5. Narrative review of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia

In this chapter, I will describe how I conducted a narrative review of the validation studies that were excluded from the meta-analysis as they used particular tests that were not used in other Chinese-speaking populations. Then I will report the features of brief cognitive assessments used in detecting suspected dementia in different Chinese-speaking populations.

5.1 Research aims and questions

One of the key aims of my PhD was to conduct a thorough investigation of the brief cognitive assessment tools used in detecting suspected dementia in Chinese-speaking populations. Thus, my main research question of the narrative review was:

1. In the case of brief cognitive assessment tools used in detecting suspected dementia (dementia, MCI) solely in one Chinese-speaking population, how long do the tests take to administer, what are the key domains covered, and what are the tests' performance?

5.2 Methods

5.2.1 Study selection

Out of the 37 studies that met the inclusion criteria, 21 studies were analysed in the meta-analysis in *Chapter 4*, 18 studies that used other brief cognitive assessments in a single Chinese-speaking population were reviewed in this current chapter (Chapter 5) as a narrative review. All 37 studies underwent translation and cultural adaptation procedures, which were described in Chapter 6 (see *Figure 1: PRISMA flow diagram of references identified and included in the review*)

5.2.2 Narrative synthesis

My narrative synthesis of the brief cognitive assessments that were not included in the meta-analysis (*Chapter 4*) started with extracting all the data from the included studies including the type of illness, diagnostic criteria, and cognitive domains related to study features in a specific Chinese-speaking population. Finally, I tabulated the administration time and cognitive domains tapped for each assessment. Then, I described those studies for each Chinese-speaking population.

I describe other variables, including the administration, the sub-tests, and performance of each test in detecting of dementia and MCI. As was done in the meta-analysis, I define either sensitivity or specificity of individual papers of meta-analysed tests below 75% as unacceptable, 75-90% as satisfactory, higher than 90% as excellent (Power et al., 2013).

5.3 Results

The narrative review of individual tests that only appeared in one Chinese-speaking population, including the administration, sub-tests, and performance data, is discussed below. *Table 5* provides information on the administration time and cognitive domains of all tests, *Appendices 4* and *5* detail the setting, subtype, reference standard, numbers of participants, cut-off scores, sensitivity, and specificity, and positive and negative likelihood ratio (+LR & - LR) for diagnosing dementia and MCI.

5.3.1 Studies in China

When all participants were recruited in the clinical settings, there were five brief cognitive assessments used in China. The tests are ranked here according to test performance (sensitivity, specificity; best test first):

- ***Hopkins Verbal Learning Test (HVLT)*** consists of three free recall trials and one recognition memory test. The participant is asked to read aloud and freely recall the 12 words immediately after the examiner reads them. After 20-25 minutes, there is a delayed recall trial, followed by a recognition memory trial. The total HVLT learning score was found to be excellent in detecting patients with AD (sensitivity: 94.7%, specificity: 92.5%) or dementia (94.7%, 93.4%) in different age groups, while it proved unacceptable (69.1.%, 70.7%) in detecting patients with amnestic MCI (aMCI) (Shi et al., 2012).
- ***Mini-Addenbrooke's Cognitive Examination (M-ACE)*** was derived from the Addenbrookes Cognitive Examination-III (ACE-III). The sensitivity and specificity were excellent and satisfactory, respectively, in detecting individuals with mild dementia (96.0%, 87.0%), but were satisfactory and unacceptable for MCI (88.0%, 72.0%) (Yang et al., 2019).
- ***Frontal Assessment Battery-Phonemic (FAB-P)*** is a new version of FAB, and in this version the original verbal fluency subtest was replaced by a Chinese phonemic fluency test. It proved to be satisfactory in detecting AD from non-dementia (93.4%, 82.9%) and nonamnestic MCI (naMCI) (86.1%, 82.7%) populations, while unacceptable in detecting aMCI from normal (77.0%, 64.2%); AD from aMCI (77.5%, 70.7%); naMCI from normal (62.3%, 58.3%); and aMCI from naMCI (56.6%, 64.2%) (Li et al., 2021).
- ***Trail-Making Test (TMT)*** includes TMT-A and TMT-B and mainly tests executive function. TMT-A consists of 25 consecutive numbers. TMT-B is 25 numbers enclosed in 13 circles and 12 squares, which has been culturally adapted for the Chinese population. Both tasks showed satisfactory performance in detecting patients with VaD (TMT-A: 85.7%, 81.6%; TMT-B: 81.6%, 83.9%), satisfactory sensitivity and excellent specificities in detecting patients with AD (TMT-A: 77.8%, 92.0%; TMT-B: 83.3%, 91.8%). The sensitivities were unacceptable, and the specificities were satisfactory in detecting patients diagnosed with MCI (TMT-A: 48.4%, 78.4%; TMT-B: 51.8%, 80.2%) or vascular MCI (VaMCI) (TMT-B: 62.9%, 75.9%) (except TMT-A for VaMCI: 70.5%, 67.7%) (Wei et al., 2018).

- ***Mini cognitive scale (Mini-Cog)*** includes two tasks – a three-word recall test and a clock drawing test. It was shown to be satisfactory in detecting patients with MCI (85.7%, 79.4%) in China (Li, et al., 2018).

Six brief cognitive assessments were tested in China with community-based controls:

- ***Boston Naming Test (BNT)*** - asks the participant to name each picture correctly within 20 seconds. Semantic cues will be given if they provide any wrong answer. If still wrong, they will be asked to select an answer from a choice of three – the correct "name", "name of an item from the same category", and "name of an item of similar quality/condition". Its sensitivities were satisfactory, excellent, and unacceptable in recognising patients with mild (79.0%, 81.0%), moderate AD (95.0%, 81.0%), and aMCI (61.0%, 81.0%), respectively, while the specificities were all satisfactory (Guo et al., 2006).
- ***Quick Cognitive Screening Test (QCST)*** comprises several subtests, including word list recall, naming test, animal fluency test, similarity test, colour trail test-1 min, clock drawing test, finger construction test and digit span. It was satisfactory when detecting patients with MCI (87.6%, 84.3%) (Guo et al., 2010).
- ***Animal Fluency Test (AFT)*** requires the participant to generate names of animals within 60 seconds. It was found to be satisfactory in detection of patients with AD (85.0%, 81.0%) using non-zodiac animals (Guo et al., 2007).
- ***Silhouettes-A*** includes silhouette drawings of 15 animals and 15 inanimate objects that need be named. Its sensitivities were satisfactory, but the specificities were unacceptable in detecting patients with AD (78.4%, 46.4%) and MCI (79.6%, 65.1%) (Huang et al., 2019).
- ***Rey-Osterreith Complex Figure Test (CFT)*** has been widely used to assess the visuo-constructional ability and visual memory with applying copying (CFT-C) and recall subtests. It is less affected by language and culture (Zhang et al., 2021). The CFT-C's sensitivity was unacceptable while specificity was satisfactory when detecting patients with AD (52.5%, 83.9%) and MCI (46.9%, 76.8%) (Huang et al., 2019).

- ***Judgment of Line Orientation (JLO)*** *test* is a 30-item test designed to simply measure visuospatial perception without involving constructional-motor demands, such as copying and assembling blocks (Spencer et al., 2013). It was found to be unacceptable in detecting patients with MCI (59.7%, 53.2%) and AD (64.0%, 65.9%) (Huang et al., 2019).

5.3.2 Studies in Hong Kong, Singapore, and Taiwan

When all participants were recruited in clinical settings, there were two short cognitive instruments validated and used in Taiwan, and only one in Hong Kong, and Singapore each, respectively.

- ***Rowland Universal Dementia Assessment Scale (RUDAS)*** is a 6-item questionnaire designed to minimise the influence of cultural learning and language diversity. Items include registration, body orientation, praxis, drawing, recall, and language. It was found to be satisfactory in Taiwanese participants for detecting dementia (76.0%, 81.0%) and MCI (79.0%, 91.0%) patients (Chen et al., 2015).
- ***Short Portable Mental Status Questionnaire (SPMSQ)*** is a 10-item instrument developed to measure orientation to time and place, memory, current event information, and calculation. In Singapore, its ability to identify patients with dementia in all education groups was satisfactory (78.0%, 75.0%), and only the ability (72.0%, 43.0%) to detect those who received fewer than six years of education was unacceptable (Malhotra et al., 2013).
- ***Time and Change (T&C)*** *test* is a simple, rapid (45 seconds), and performance-based task related to real-world function, namely telling time, and creating a sum of money from change that is laid out in front of the person. Its specificity was satisfactory while sensitivity was unacceptable in a study detecting a small number of patients with dementia (74.5%, 88.2%) in Hong Kong (Chan et al., 2005).
- ***Clock Drawing Test (CDT) - Rouleau Scoring System*** consists of two components (command and copy), which use a 10-point scale to evaluate

the quality of the clock face's integrity (0 - 2), the presence and order of the numbers (0 - 4), as well as the placement and presence of the hands (0 - 4). In Taiwan, its ability to identify mild AD (60.0%, 72.0%) and questionable dementia (74.2%, 56.4%) is unacceptable (Chiu et al., 2008).

When the controls were from the community, there were two short cognitive instruments validated and used in Singapore, and three in Taiwan.

- ***Chinese version of the Verbal Learning Test (CVVLT)*** consists of 9 two-character nouns presented over 4 learning trials, followed by recall tests after 30-second, 10-minute, and then a delayed word recognition test. In Taiwan, the value of using this (total score) was excellent in identifying AD patients (92.0%, 91.0%), while the sensitivity dropped to only satisfactory for those who aged over 75 (86.0%, 92.0%) (Chang et al., 2010).
- ***Brain Health Test-Cog part (BHT-Cog)*** was developed indigenously in Taiwan and was the second half of BHT. It included orientation to time, immediate and delayed recall, category verbal fluency test, and a clock drawing test. It had excellent sensitivities in detecting patients with dementia (91.5%, 87.3%) and MCI (91.5%, 64.9%), while its specificities were satisfactory and unacceptable in detecting dementia and MCI, respectively (Tsai et al., 2018).
- ***Frontal Assessment Battery (FAB)*** was developed for testing executive function at the bedside and has been adapted for Asian contexts, with the verbal fluency subtest being substituted by category fluency because of differences in linguistics. It had an excellent sensitivity and satisfactory specificity (92.0%, 78.7%) in detecting cognitive impairment (MCI and dementia) in Singapore, only the specificity (70.3%) in detecting those who received more than five years of education was unacceptable (Chong et al., 2010).
- ***Quick mild cognitive impairment screen (Qmci-TW)*** based on the AB Cognitive Screen then added story recall and reweighted all the subtests. In Taiwan, its sensitivity was excellent, and specificity was satisfactory in detecting AD (94.0%, 78.0%), while its sensitivity was unacceptable and specificity was excellent in recognising MCI (69.0%, 97.0%) (Lee et al., 2018).

- ***Visual Cognitive Assessment Test (VCAT)*** is a 11-item visual-based cognitive test for use in culturally diverse population without a need for further adaptation and is weighted more toward the episodic memory domain. It had a satisfactory sensitivity and an unacceptable specificity (75.4%, 71.1%) in detecting cognitive impairment (MCI, AD) in Singapore (Low et al., 2020).

Table 5: The administration time and cognitive domains of all tests (including the studies included in the meta-analysis)

Test name & settings (A/B)	Study code	Time (min)	Number of domains	DSM-5's five domains				
				A	E	M	L	P
ACE-III, ACE-R (A, B)	C9, C10, H3	16-24		●	●	●	●	●
MoCA (A, B)	C5, C6, C18, C19, H2, H4, S3, T3, T7, T8, T9	10-15		●	●	●	●	●
MoCA-BC (B)	C15, C16, C17	15*	5	●	●	●	●	●
QCST (B)	C18	8-12		●	●	●	●	●
Qmci (B)	T9	<5		●	●	●	●	●
RUDAS (A)	T4	10		●	●	●	●	●
VCAT (B)	S4	15.7		●	●	●	●	●
Mini-Cog (A)	C4	3			●	●	●	●
MMSE (A, B)	C3, C7, H2, H3, H4, S4, T7, T9	10-15	4	●		●	●	●
BHT-cog (B)	T5	4		●		●	●	
CDT (A)	H1, T1, T2	1.5		●	●			●
CFT-C (B)	C19	10*	3	●	●			●
M-ACE (A)	C3	<5		●			●	●
Silhouettes test (B)	C19	3-5				●	●	●
BNT (B)	C12	≈ 15*					●	●
DRS/MDRS (B)	C13, C14	15		●		●		
SPMSQ (A)	S1	<5*		●		●		
STT-A (B)	C20	≈ 5	2	●			●	
STT-B (B)	C19, C20	≈ 5			●	●		
T&C (A)	H1	≈ 1		●				●
AFT (B)	C11	1					●	
CVVLT (B)	T6	NA				●		
FAB-Phonemic (A)	C1	5			●			
FAB (B)	S2	5	1		●			
HVLT (A)	C2	10*				●		
JLO (B)	C19	15*						●
TMT (A)	C8	2.5-5			●			

Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. The letter of A, E, M, L, P of DSM-5 refer to the cognitive domains of complex attention, executive function, learning and memory, language, perceptual-motor function, respectively. The setting of A / B refers to clinical or community-based controls, respectively. ACE-R&III = Addenbrooke's cognitive Examination Revised & III; AFT = Animal Fluency Test; BHT-cog = Brain Health Test-Cog part; BNT = Boston Naming Test; CDT = Clock Drawing Test; CVVLT = Chinese version of the Verbal Learning Test; DRS/MDRS = Mattis dementia rating scale; FAB = Frontal Assessment Battery; HVLT = Hopkins Verbal Learning Test; M-ACE = Mini-Addenbrooke's Cognitive Examination; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; MoCA-BC = Montreal Cognitive Assessment Basic; QCST = Quick Cognitive Screening Test; Qmci = Quick Mild Cognitive Impairment Screen; RUDAS = Rowland Universal Dementia Assessment Scale; SPMSQ = Short Portable Mental Status Questionnaire; STT = Shape Trail Test; T&C = Time and Change Test; TMT = Trail-Making Test; VCAT = Visual Cognitive Assessment Test; *indicates that the information was obtained from other articles as the included articles did not specify.

5.4 Discussion

5.4.1 Summary of findings

5.4.1.1 The best tests in different Chinese-speaking populations

Some cognitive assessment tools for dementia or MCI were validated in only one Chinese-speaking population. The findings were therefore not replicated. In this review, I have identified the top-performing tests for each Chinese-speaking population in detecting all-cause, all-stage dementia, and MCI. In China, the tests with the best sensitivity and specificity in detecting dementia and MCI from non-dementia and non-MCI controls were the Hopkins Verbal Learning Test (HVLT: 94.7%, 93.4%) (Shi et al., 2012) and the Quick Cognitive Screening Test (QCST: 87.6%, 84.3%) (Guo et al., 2010). In Taiwan, the best-performing tests in distinguishing dementia and MCI from non-dementia and non-MCI controls were the Brain Health Test-Cog part (BHT-Cog: 91.5%, 87.3%) (Tsai et al., 2018) and the Rowland Universal Dementia Assessment Scale (RUDAS: 79.0%, 91.0%) (Chen et al., 2015). In Singapore, the Frontal Assessment Battery (FAB: 92.0%, 78.7%) performed the best in differentiating cognitive impairment (CI, including dementia and MCI) from non-CI (Chong et al., 2010). There was no test used in Hong Kong that had both sensitivity and specificity higher than 75%. Those brief cognitive tests which had good performance in respect of dementia and MCI diagnosis can be further validated in other Chinese-speaking populations in the future.

5.4.2 Strengths and limitations

The primary strength of the methodology used in this study is its ability to summarise and convey results narratively when statistical or other formal methods of data estimation are not feasible or acceptable. It permits analysis of similarities and differences across studies, exploration of relationships within the data, and evaluation of the evidence quality. It provides a summary of information

relevant to a particular review question that may be used to advise practitioners (Lisy & Porritt, 2016).

The general guidelines for narrative review, which were published in 2006, included four steps – (1) creating a theoretical framework that explains how, why, and for whom interventions (in my study, cognitive assessments) are effective; (2) creating a preliminary review (developing an initial description of the findings of included studies); (3) examining relationships in the data: look for emerging patterns in the data to find any reasons for variations in the direction or size of effect across included studies; and (4) evaluating the robustness of the product of the review (e.g., quality assessment of the included studies) (Popay et al., 2006).

My study followed these four suggested steps. Some of them are described in *Chapter 4* (e.g., the framework for conducting the study and the assessment of the quality of each study/cognitive test). However, there are still certain limitations. I was only able to describe the diagnostic performance of each test used in Chinese-speaking populations and was unable to identify a trend across the included research.

5.4.3 Findings of systematic review, meta-analysis, and narrative synthesis related to the validation study

My meta-analysis study that included 21 studies demonstrated that the ACE (ACE-R & ACE-III) had the best sensitivity and specificity with acceptable heterogeneity among all the brief cognitive assessment instruments used to detect dementia and MCI in clinical settings. In my narrative synthesis, the best-performing tests varied among Chinese-speaking populations, and none of them performed well in both dementia and MCI.

The most recent version of ACE, the ACE-III, has not been formally made accessible on the host website for use by Taiwanese clinicians. Therefore, a key aim of my PhD was to culturally adapt and validate the Taiwanese version of the

ACE-III (T-ACE-III), along with considering the acceptability of a supplementary test of executive function. This will be covered in the following chapters.

Chapter 6. The translation and cultural adaptation procedures employed for brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia

In this chapter, I will describe the reported procedures used for translation and cultural adaptation of brief cognitive assessment tools, as they were described in every publication which I included, both in the meta-analysis and in the narrative review.

6.1 Research aims and questions

The evaluation of the translation and cultural adaptions of brief assessment tools was performed to look into each step of the process, ensuring that every study which validated brief cognitive assessment accurately described how they made modifications, depending on their own cultural contexts. Thus, my main research question was:

1. What were the reported procedures for translation and cultural adaptation of brief cognitive assessment tools used in every publication which I included, both in the narrative review and meta-analysis, in respect of different Chinese-speaking populations?

6.2 Methods

6.2.1 Study selection

The selection of studies included 21 studies that were used in the meta-analysis (see *Chapter 4. Systematic review and meta-analysis of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia*) and 18 studies that were included in the narrative review (see *Chapter*

5. Narrative review of brief cognitive instruments used with Chinese-speaking populations to evaluate suspected dementia).

6.2.2 Translation and cultural adaptation procedures

The Manchester Translation Reporting Questionnaire (MTRQ) and Manchester Cultural Adaptation Reporting Questionnaire (MCAR) scales were developed at the Centre for Primary Care at the University of Manchester specifically for evaluating the quality of "translation" and "cultural adaptation" procedures, respectively (Mirza et al., 2017; Waheed et al., 2020). The process incorporates a coding scheme of seven outcomes about the overall quality of reported procedures (see *Appendix 15* for the details of scores for each scale used to rate the studies). Five scores can be given:

- **Score 0:** When the translation/ cultural adaptation procedures were not mentioned
- **Score 1**: If the translation/ cultural adaptation procedures were mentioned, but with no details
- **Score 2**: If the translation/ cultural adaptation procedures were mentioned, but with insufficient details for replication (2a) or there is reference to another publication (mainly original authors of the test) which provides insufficient details for replication (2b)
- **Score 3**: If any pre-existing guidelines about the translation/ cultural adaptation procedures were provided (such as possible scenarios of cross-cultural adaptation from Beaton et al., 2000), but with insufficient details for replication
- **Score 4**: If the translation/ cultural adaptation procedures were mentioned with sufficient details directly in the study so as to enable replication (4a) or there is reference to another publication which provides sufficient details for replication (4b)

The seven steps of the validation and adaptation procedures are as followed (Mirza et al., 2017):

- **1. Translation:** Direct translation from English to Chinese languages, without any cultural modification, typically with the help of a native or proficient Chinese speaker or an official translator.
- **2. Back translation**: Creating a retroversion of the first translation from Chinese languages back to English, often with the aid of a native or proficient Chinese speaker or an official translator.
- **3. Users in coproduction**: Potential users of the assessment, such as participants from a Chinese-speaking population, who provide feedback that informs the translation of the assessment.
- **4. Expert recommendations**: Experts in translation, Chinese languages, or assessment-related topics who provide comments that inform the translation of the assessment.
- **5. Revisions based on step-by-step feedback**: Every time a modification or recommendation is suggested and approved; the translated assessment is updated continuously.
- **6. Involvement of the original authors:** Authors of the original assessment who provide comments that affect how the translated assessment is developed.
- **7. Pilot Study**: Administering translated test to gauge its feasibility and acceptability among prospective users.

The professionals involved in the validation and adaptation procedures included: psychiatrist, psychologists, physicians, neurologists, geriatricians, bilingual experts, researchers, and accredited translators.

I evaluated all studies, while Jen-Chieh Lai evaluated the studies with clinical controls, and Esther Hui was responsible for the studies that used community controls. Subsequently, I used the Kappa statistic to calculate interrater reliability for MTRQ and MCAR. Finally, my supervisors, GL or NM reviewed with me any papers that had inconsistent scores, and we determined the final score for each item in all the scales, after discussing and reaching a consensus.

6.3 Results

6.3.1 Inter-rater reliability of the MTRQ and MCAR

The inter-rater reliability coefficients between J-CL and myself for MTRQ and MCAR were 0.835 and 0.907, respectively.

6.3.2 The quality of each study and the reported procedures for translation and cultural adaptation

6.3.2.1 Limited reporting on translation and cultural adaptation

Table 6 displays the quality of translation and cultural adaptation for the cognitive assessment tools, as well as the reported steps involved in the process; it also has information on the professionals involved in the process of translation and adaptation. Four studies (Chang et al., 2012; Chu et al., 2015; Lin et al., 2003; Low et al., 2020) met the highest score of 4 on the MTRQ and MCAR, twenty-seven studies scored 0 to 4, six studies scored 0.

In the papers included for this review, there was a paucity of information reported in the translation and cultural adaptation processes. The scores of the included papers for MTRQ and MCAR were 1.0 and 2.1 in China, 1.2 and 1.6 in Hong Kong, 2.0 and 2.0 in Singapore, 2.0 and 2.9 in Taiwan.

6.3.2.2 Actual steps reported and professionals involved

Regarding the steps reported, 48.6% (n = 18) of the papers reported translation, 29.7% (n = 11) reported backtranslation, 13.5% (n = 5) reported involvement of the original authors, 10.8% (n = 4) reported revisions based on step-by-step feedback, 5.4% (n = 2) reported expert recommendation, and 2.7% (n = 1) reported carrying out a pilot study as well as users in co-production. About the individuals involved in the process, 13.5% (n = 5) of the papers involved a physician or neurologist, 8.1% (n = 3) of the papers involved a psychiatrist or

psychologist, 8.1% (n = 3) involved a bilingual expert/researcher, 2.7% (n = 1) involved a geriatrician, and 2.7% (n = 1) involved an accredited translator. Only one study met the highest standards for both MTRQ and MCAR in Hong Kong, reporting all seven steps and involving the greatest number of experts (Chu et al., 2015).

It is essential to highlight that both MTRQ and MCAR are only used to evaluate a study if the procedures of translation of cultural adaptation have been appropriately stated on the papers, not the actual steps taken by the researchers. In my thesis, the papers that referred to other publications (e.g., 2b, 4b) without describing the processes that were followed, and individuals who were involved on their own, were not ticked off on *Table 6*. Also, those studies that emphasised that their cognitive screening tool is culture-fair (STT; Zhao et al., 2013) or language-neutral (VCAT; Low et al., 2020) or had minimal language demands (CDT; Lin et al., 2003), were given a score of 4 on MCAR, since they would have recognised the importance of reporting information while conducting a validation study.

Table 6: Translation steps undertaken and the individual professionals involved in the papers included in the review

TEST CODE & AUTHOR		MTRQ	MCAR	Translation steps undertaken							Individuals involved				
				Translation	Back Translation	Users in co-production	Expert recommendation	Revisions based on step-by-step feedback	Involvement of original authors	Pilot study	Psychiatrist/ Psychologists	Physicians/ Neurologists	Geriatricians	Bilingual experts/ researchers	Accredited translators
C1	Li et al., 2021	0	2a												
C2	Shi et al., 2012	0	0												
C3	Yang et al., 2019	4a	2a	✓	✓			✓				✓		✓	
C4	Li et al., 2018	0	0												
C5	Wen et al., 2008	0	4a												
C6	Li, Jia, et al., 2018	0	0												
C7	Xu et al., 2003	2b	4a	✓											
C8	Wei et al., 2018	2a	0	✓	✓										
C9	Wang et al., 2017	1	4a	✓								✓			
C10	Fang et al., 2014	1	4a	✓											
C11	Guo et al., 2002	0	2a												
C12	Guo et al., 1991	0	2a												
C13	Qian et al., 2021	1	1	✓											
C14	Guo et al., 2004	2b	2b	✓	✓										
C15	Chen et al., 2016	1	1	✓	✓				✓						
C16	Huang et al., 2018	2b	1	✓	✓				✓						
C17	Zhang et al., 2019	2b	2b	✓	✓				✓						
C18	Guo et al., 2010	0	4a												
C19	Huang et al., 2019	2a	2a	✓				✓							
C20	Zhao et al., 2013	0	4a												
	Average	***1.0***	***2.1***												

H1	Yung et al., 2005	0/0	0/0												
H2	Yeung et al., 2014	0	0												
H3	Wong et al., 2013	2a	4a	✓			✓								
H4	Chu et al., 2015	4a	4a	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
	Average	***1.2***	***1.6***												
S1	Malhotra et al., 2013	1	0		✓										
S2	Chong et al., 2010	2b	4b												
S3	Liew et al., 2015	1	0	✓	✓				✓		✓	✓		✓	
S4	Low et al., 2020	4a	4a												
	Average	***2.0***	***2.0***												
T1	Lin et al., 2003	4a	4a												
T2	Chiu et al., 2008	0	0												
T3	Tsai et al., 2016	2b/0	4b/0	✓											
T4	Chen et al., 2015	3	3	✓	✓			✓						✓	
T5	Tsai et al., 2018	N/A													
T6	Chang et al., 2010	4a	4a												
T7	Chang et al., 2012	0	4a												
T8	Tsai et al., 2012	2a	4a	✓	✓						✓	✓			
T9	Lee et al., 2018	3	3	✓											
	Average	***2.0***	***2.9***												

Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. NA = Not available (the information is not provided in the study); MCI = Mild Cognitive Impairment CEBM = the Centre for Evidence-based Medicine diagnostics criteria; MTRQ = Manchester Translation Reporting Questionnaire; MCAR = Manchester Cultural Adaptation Reporting Questionnaire.

6.4 Summary of findings

Only four studies stated in detail how they translated and culturally adapted the tests into their own settings, and only one of them took all the steps and involved the most professionals. Reported processes of translation and cultural adaptation, which are crucial when validating a test for a new culture/population, were still insufficient. The information provided should contain not only the alterations they made for cultural adaptation, but also the processes they used as well as information about the professionals who participated in the adaptation.

Chapter 7. Validation of the Taiwanese version of ACE-III for detecting suspected dementia in Taiwan

In the previous two chapters, I reported a systematic review and meta-analysis followed by a narrative review. These reviews found that the ACE-III performed well in the detection of dementia in Chinese-speaking populations where it had been tested. In this chapter, I outline the initial versions of the Addenbrooke's Cognitive Examination (ACE & ACE-R) as published by the original authors (Mathuranath et al., 2000; Mioshi et al., 2006), its clinical usage and the current translated versions of the Addenbrooke's Cognitive Examination III (ACE-III). Then, I detail the methods I used to carry out the empirical study in Taiwan, which included gaining ethical approvals, adapting the ACE-III for use in Taiwan (T-ACE-III), interviewing patients, testing it with patients attending a memory clinic in Taiwan, and carrying out an analysis of its psychometric properties, based on my findings.

This study was published in 2021 in *Archives of Clinical Neuropsychology* (Yu et al., 2021). See *Appendix 16* for a copy of the published paper.

7.1 Introduction

7.1.1 The predecessors of ACE-III (ACE, ACE-R)

The features of the MMSE and different versions of the Addenbrooke's Cognitive Examination (ACE) are shown in *Table 7*. The original ACE was introduced as a cognitive screening tool in 1990 (Mathuranath et al., 2000). It measured the cognitive domains of orientation (10 points), attention (8 points), memory (35 points), verbal fluency (14), language (28), and visuospatial ability (5) (see *Table 8*); this yielded a maximum score of 100. It took 15 to 20 minutes to administer and incorporated all of the components of the MMSE. It was found to be a useful tool in differentiating FTD and AD amongst patients in the early stages of dementia (Mathuranath et al., 2000).

The Addenbrooke's Cognitive Examination Revised (ACE-R) was developed in 2006 not only to allow for cross-cultural usage and translation but also to tackle some existing issues to improve the sensitivity and specificity of ACE, such as dealing with ceiling effects for the naming test and the small number of items for visuospatial processing. The ACE-R demonstrated a significantly higher sensitivity and specificity for detection of dementia than the ACE (Mioshi et al., 2006).

The Addenbrooke's Cognitive Examination III (ACE-III) was developed in 2012 to address the copyrights issue for MMSE items contained in the ACE-R, since the MMSE was no longer freely accessible after 2001 (Seshadri & Mazi-Kotwal, 2012). Other changes made were (Hsieh et al., 2013): Firstly, in the case of attention testing, spelling of the word 'WORLD' backwards was replaced with counting back from 100 in 7s to deal with issues relating to difficulty level (Ganguli et al., 2016). Secondly, in the language domain, due to hearing or attention problems, healthy adults often failed on the verbal repetition item ('no ifs, ands or buts') (Valcour et al., 2002), and it also caused some translation difficulties between languages (Werner et al., 1999) and so it was replaced with other phases; ACE-R's items evaluating language-comprehension (e.g., 3-stage command and 'close your eyes') was considered lacked sensitivity to detecting cognitive impairment (Hsieh et al., 2013; Mathuranath et al., 2000) and was omitted; the item assessing language-writing was changed from writing a single sentence to writing two or more sentences; naming items of a pencil and a clock were replaced with a book and a spoon and there were 10 other objects to be named. Thirdly, in the visuospatial domain, intersecting pentagons were substituted by two intersecting 'infinity' symbols. The ACE-III one of the tests in a toolkit of tests recommended by the Alzheimer's Society in the UK (Ballard et al., 2015).

In 2015, a shorter version of the ACE-III - the Mini-Addenbrooke's Cognitive Examination (M-ACE) was derived from the ACE-III, using a data-driven scaling method. It was demonstrated to be more sensitive than the MMSE and less prone to ceiling effects (Hsieh et al., 2015).

Table 7: Domains and items of ACE versions & MMSE

		ACE	ACE-R	ACE-III	M-ACE	MMSE
Orientation	Time	5	5	5	4	5
	Place	5	5	5		5
Registration		3	3	3		3
Attention / Concentration: serial 7 s, WORLD backwards		5 (Best performed task)	5	5 (Serial 7s only)		5
	Sub-total	**18**	**18**	**18**	**4**	**18**
Memory	Recall	3	3	3		3
	Anterograde memory	28	19	19	14	
	Retrograde memory	4	4	4		
	Sub-total	**35**	**26**	**26**	**14**	**3**
Verbal fluency: Letters and animals in 1 min		**14**	**14**	**14**	**7** (Letters or animals)	
Language	Naming	12	12	12		2
	Comprehension	8	8	7		4
	Repetition	5	4	4		1
	Reading	2	1	1		
	Writing	1	1	2		1
	Sub-total	**28**	**26**	**26**		**8**
Visuospatial abilities	Intersecting pentagons	1	1	1 (Intersecting infinity figures)		1
	Wire (Necker) cube	1	2	2		
	Clock drawing	3	5	5	5	
Perceptual abilities	Dot counting	-	4	4		
	Fragmented letters	-	4	4		
	Sub-total	**5**	**16**	**16**	**5**	
Total score		**100**	**100**	**100**	**30**	**30**

Note. The content of this table is based on the Hodges and Larner (2016) paper; ACE = Addenbrooke's Cognitive Examination; ACE-R = Addenbrooke's Cognitive Examination Revised; ACE-III = Addenbrooke's Cognitive Examination III; M-ACE = Mini-Addenbrooke's Cognitive Examination; MMSE = Mini-Mental State Examination.

7.1.2 Performance of ACE-III in clinical settings

7.1.2.1 MCI

Mild cognitive impairment (MCI) is a term used to describe people who have cognitive changes between normal ageing and early dementia (Petersen, 2004) and whose cognitive impairment does not lead to significant changes in everyday functioning. It may represent the prodromal stage of a dementia. In order to detect MCI, a doctor can not only look to the clinical history but also consider using sensitive and brief cognitive screening tools. Some studies found the ACE-III to be sensitive at detecting amnestic MCI (aMCI) since it incorporates sensitive memory test items that are useful in detecting early AD (Matias-Guiu, Valles-Salgado, et al., 2017). For all types of MCI, it has been shown to have high sensitivity but low specificity (Potts et al., 2021).

7.1.2.1 Dementia

Compared to other screening tools, such as the MMSE and the MoCA, the ACE-III has been found to be one of the most sensitive cognitive screening instruments for the detection of dementia (Hodges & Larner, 2016). Also, it has been found to readily differentiate between patients with AD and FTD, with distinct profiles of performance on items from the domains of language and memory (Elamin et al., 2016; Hsieh et al., 2013; Matias-Guiu, Cortés-Martínez, et al., 2017). The Verbal-Language/Orientation-Memory (VLOM) ratio derived from the ACE-III contrasts memory and language test scores has been found to be particularly useful in differentiating AD from FTD (Hodges & Larner, 2016; Mathuranath et al., 2000). In addition, the ACE-III has also been found to be helpful in distinguishing between individuals with subjective cognitive impairment and those with early-onset dementia (Elamin et al., 2016).

7.1.3 Current translated versions of ACE-III

According to the official website of the FRONTIER Research Group (www.ftdrg.org), there are translated versions of ACE-III for use in diverse populations, such as Spanish, Italian, Korean, and Chinese. When I first started my PhD project in 2019, the website included more than 30 versions of the ACE-III available for use, but it did not have a Taiwanese version. *Table 8* displays the available versions of ACEs (ACE, ACE-R, ACE-III, & M-ACE) used in East Asia. Only Japan and Korea translated and adapted the original version of the ACE when it was first developed. After the first revision when it became the ACE-R, Chinese and Cantonese versions were created for China and Hong Kong. Up until that point, neither the ACE nor ACE-R was available in Taiwan. As outlined in *Chapters 1* and *2*, the culture (see *1.2.2 The multi-ethnicities living in* Taiwan.) and language (*see 2.2 Chinese-speaking culture*) in Taiwan is sufficiently different from China and other Chinese-speaking populations to require distinct adaptation and validation of cognitive tools. Although a Taiwanese study, which adapted the ACE-III, was published in 2019 (Lin et al., 2019), it did not follow published guidance on translation and adaptation, and a language expert did not appear to have been consulted. More importantly, this study did not use any gold standard to reach the diagnosis of MCI but used the MMSE and MoCA scores. Also, no reliability measures were gathered. The different versions of the ACE that are available in East Asia are shown in *Table 8*.

Table 8: Reported translations of ACE, ACE-R, ACE-III, & M-ACE in East Asia

Language and version		ACE	ACE-R	ACE-III	M-ACE
Chinese-speaking population	**Cantonese** (Hong Kong)		Wong et al., (2013)		
	Chinese (China)		Fang et al., (2014)	Wang et al., (2017)	Pan et al., (2022)
	Taiwanese (Taiwan)			Lin et al., (2019); Yu et al., (2021)	
Japanese (Japan)		Yoshida et al., (2011)	Hitomi et al., (2012); Yoshida et al., (2012)	Takenoshita et al., (2019)	Senda et al., (2020)
Korean (Korea)		Heo et al., (2012)	Kwak et al., (2010)	Lee (2016) [1]	

Note. The information contained in this table was extracted on the 27th of September 2022. ACE = Addenbrooke's Cognitive Examination; ACE-R = Addenbrooke's Cognitive Examination-Revised; ACE-III = Addenbrooke's Cognitive Examination-III; M-ACE = Mini-Addenbrooke's Cognitive Examination. [1] Korean version of ACE-III was translated and adapted by Lee (2016), but no study was published.

7.2 Aims

The main aim of my research study was to culturally adapt and validate the ACE-III for use in Taiwan. My research questions were –

1. What are the sensitivity, specificity, and best cut-off scores for the T-ACE-III test to distinguish people presenting with suspected dementia from those without dementia in Taiwan?
2. What is the diagnostic accuracy for the T-ACE-III to distinguish people presenting with AD and non-AD from those without dementia in Taiwan?
3. Are there any demographic factors that have an impact on the participants' performance on the T-ACE-III?

7.3 Methods

7.3.1 Data collection plan

After contacting several institutions in Taiwan, the neurology department at Shuang-Ho Hospital agreed to collaborate with me.

The data collection was conducted at Shuang-Ho Hospital in Taiwan from June to October 2020, with the support of Drs Chaur-Jong Hu, Chien-Tai Hong, Li-Kai Huang, and Lung Chan. One of the primary functions of the dementia centre in Shuang-Ho Hospital is to carry out training for dementia diagnosis and care, using up-to-date, evidence-based, screening tools, as well as to providing dementia-friendly clinical services. I obtained ethical approval from the UCL Research Ethics Committee (UCL-REC: 16845/001) and Taipei Medical University-Joint Institutional Review Board (TMU-JIRB: N202004089), including Data Protection Registration (NO: Z6364106/2020/01/48) in the UK (see *Appendix 17-18*).

I also obtained permission to revise, adapt and use the cognitive tests for Taiwan from the original authors and relevant institutions (see *Appendix 19-21*). The adapted and validated Taiwanese version of T-ACE-III now is available at www.ftdrg.org (The University of Sydney, 2023) (also see *Appendix 28*).

7.3.2 Cultural Adaptation

7.3.2.1 Existing guidelines

When I use the word ‘China’, I am referring to the People's Republic of China, excluding the Special Administrative Regions of Hong Kong and Macau. Mandarin Chinese is officially spoken in Taiwan and China, while the written form of Simplified Chinese is used in China and Singapore, and Traditional Chinese is used in Taiwan, Hong Kong, and Macau. The most significant difference between the two written forms of Chinese is the form of the individual characters (see *2.2.2 Written forms of Chinese*, for more details).

I used the Cross-Cultural Adaptation Process (Beaton et al., 2000) to guide my cultural adaptation of the ACE-III for the Taiwanese population. This adaptation process is used when assessments in one language are used in other populations with the same language, but where there may be differing cultures or written language characters. So, for example, the term used for the leader of state differs ('Chairman' in China and 'President' in Taiwan). The form of written characters (simplified Chinese in China, traditional Chinese in Taiwan), the culture and the history of the two populations also differ. The role of different Chinese dialects is considered important when culturally adapting neuropsychological assessments due to a significant difference in performance on abstract thinking tasks. For example, in response to the question *'What are the similarities between laughing and crying?'* in the Cognitive Abilities Screening Instrument (CASI), Chinese speakers can say 'they are all emotional reactions'. However, Hokkien speakers cannot explicitly state this. They can only say 'they cry when sad and laugh when happy', since there is no equivalent phrase or usage for "emotion" in their language (Chen, 1996). Further details are outlined in the following section and in *Appendix 22: Discussion on T-ACE-III adaptation.*

7.3.2.2 Steps taken in adapting T-ACE-III

When constructing the Taiwanese (Traditional Chinese) version of ACE-III (T-ACE-III), I largely relied on the simplified Chinese version (C-ACE-III), as well as the English (UK), Korean, and Japanese versions, all of which are accessible through the University of Sydney (www.ftdrg.org). *Figure 4* represents the adaption process which I followed. I first translated the original C-ACE-III, including instructions and answer sheets, into Traditional Chinese. This translated version was subsequently discussed with Cheng-Chang Yang (C-CY), a postgraduate researcher at Shuang-Ho Hospital with a background in language psychology and linguistics, and it was adapted to Taiwanese culture (*Appendix 22*). Items were modified in the domains of verbal fluency, memory, language, and visuospatial ability (see *7.3.3 Adaptation of items for the T-ACE-III*). Two clinical psychologists, Jen-Chieh Lai and Pei-Jung Wu (J-CL & P-JW), two neurologists, Chien-Tai Hong, and Li-Kai Huang (C-TH, L-KH), and my four

supervisors (as clinicians or consultants in the fields of psychiatry or psychology) acted as an expert committee that assisted me in refining the test further. The final version of the T-ACE-III for testing was created after undergoing preliminary testing with a small sample of Taiwanese colleagues and elderly relatives, and then incorporating their feedback. The changes made are set out in paragraph 7.3.3 below.

Figure 4: The adaptation processes for the T-ACE-III

7.3.3 Adaptation of items for the T-ACE-III

Details regarding the adaptation of specific items are set out in *Table 9*. The Taiwanese version of the ACE, the T-ACE-III, has the same maximum score of 100 as the original version and covers the same five cognitive domains: attention and orientation (18 points), memory (26 points), fluency (14 points), language (26 points), and visuospatial abilities (16 points). A higher score indicates better cognitive function.

Some items could be directly translated without further adaptation, e.g., questions relating to orientation. Some items were revised to enable a better understanding of items when the test was administered to those in the Taiwanese population. Items in four cognitive domains were specifically adapted:

- **Verbal fluency (letter fluency test, LFT)**: As there are no letters used in a graphical, non-alphabetic Chinese language system, the methods used to generate a list of words in an LFT may be fundamentally distinct from those used by users of languages with alphabetic systems (e.g., English,

Spanish, Japanese). The character (字) is the basic meaningful unit in both spoken and written Chinese (Eng et al., 2018), and 'Bopomofo' (字 spelled as ‘ㄗ’) in Taiwan or ‘Hanyu Pinyin’ (字 spelled as ‘zì’) in China are the phonetic systems that composed the sounds, in which those phonemes provide no meaning, functioning similarly to letters of the alphabet in English. However, those phonemes cannot be seen in the Chinese characters (so words cannot be “sounded out” as they can be seen with letters in English words). Chinese phonetic systems aid in pronunciation when learning Chinese and can generate characters when typing on a computer. In short, how we pronounce the characters and how we write the characters are barely related, so children see a new character and do not know how to pronounce it, but in English, people can pronounce a new word when they first see it. Therefore, using the first letter to develop LFT is impossible in Mandarin (Chinese) (Hung et al., 2016). In the English version of the ACE-III, the LFT asks participants to generate as many words as possible beginning with the letter ‘P’ in a minute. The simplified Chinese version, used in China, asks participants to use a single-syllable Chinese word (車[che], meaning car) with a meaning – in English, this is like asking people to generate from the word ‘car’, the words and phrases ‘car keys’ and ‘park a car’ and the sound ‘car’ in the words ‘carpet’ and ‘cartoon’. However, as the network activated by semantic and phonemic processing in the brain may be altered by an individual's linguistic background, it is unclear if LFT scoring rules created for Western languages are applicable to Chinese, which has a logographic writing system (Eng et al., 2018). In order to reduce complexity and difficulty, the items used to assess **letter fluency** in the traditional Chinese version for Taiwan were replaced with one assessing **category fluency** (asking people to name “vegetables and fruits”), which is less affected by education levels and is more appropriate, given the differences in language.

- **Memory**: In the anterograde memory section, the question in the simplified Chinese version relating to a person's name and address in China was

changed into the Taiwanese equivalents. In the retrograde memory section, the question about the chairman was replaced with a question about the president in Taiwan's history.

- **Language**: In the object naming section, the pictures of ‘spoon’ and ‘kangaroo’ in the UK English version were replaced with ‘pencil’ and ‘panda’ in the simplified Chinese version, while the two pictures remained the same as UK English version (spoon & kangaroo); The pictures of ‘harp’, and ‘barrel’ were substituted with pictures of ‘drum’, and ‘light bulb’, since those objects are more common in Taiwan. In the comprehension part of the test, the questions were now: Point to the animal which is marsupial (Q1); point to the animal which is a reptile (Q2); point to the item which is found in the Antarctic (Q3); point to the item which has a nautical connection (Q4). In the reading section, the UK English version uses five terms ‘sew’, ‘pint’, ‘soot’, ‘dough’, ‘height’, which are irregular pronounced words, while the simplified Chinese version used in China uses eight single characters that are not all irregularly pronounced. In the traditional Chinese version, I replaced these with five terms that are all irregularly pronounced, as in the UK English version.
- **Visuospatial ability**: The fragmented letters K, M, A, T were replaced with Arabic numbers 8, 2, 4, 9, which is less affected by education level.

Table 9: The modified items in the Simplified and Traditional Chinese versions of ACE-III compared to the English (UK) version

Revised domain	English version (UK)	Simplified Chinese version (China)	Traditional Chinese Version (Taiwan)
Verbal fluency- Letter	I will give you a letter of the alphabet and I'd like you to generate as many words as you can beginning with that letter, but not names of people or places. For example, if I give you the letter "C", you could give me words like "cat, cry, clock" and so on. But you can't give me words like Catherine or Canada. Do you understand? Are you ready? You have one minute. The letter I want you to use is the letter "P.	I will give you a single-syllable Chinese word, please use this word to generate as many words or phrases as possible, but it cannot be a name or place. For example, If I speak the word "hai (sea)", you can generate "hai yang (ocean)", "hai siao (tsunami)"," da hai" (great ocean). But it cannot be "Shanghai". Are you ready? You have one minute, please, to generate phases with "car".	Now, can you name as many vegetables and fruits as possible. You have 1 minute.
Memory- Anterograde memory- Name and address	Harry Barnes 73 Orchard Close Kingsbridge Devon	Wang Chun-Ming No. 18 Harbin Road, Haidian District, Beijing (With other normal Chinese addresses and names in Part 4)	Wang Zhi-Wei No. 16, Chuangxin Road, Nanzi District, Kaohsiung City (With other normal Taiwanese addresses and names in Part 4)
Memory- Retrograde Memory- Famous people	1. Name of the current Prime Minister 2. Name of the first female Prime Minister 3. Name of the USA president 4. Name of the USA president who was assassinated in the 1960s	1.The name of the current chairman of China? 2. The name of the first President of the People's Republic of China 3. The name of the only female emperor throughout Chinese history 4. The name of current American president	1.The name of the current president of Taiwan. 2. The name of the first elected president in Taiwan 3. The name of the Taiwan president who first experienced a gun shooting event and was injured 4. The name of current American president
Language- Word repetition	1. Caterpillar 2. Eccentricity 3. Unintelligible 4. Statistician .	1. zhi zutian dikuan (When the hearts know content, it feels safe and assured that there is a place for them in this world) 2. xin an cai gen xiang (Conduct yourself with integrity and there's nothing to fear)	1. xiang sheng bei shi (Care about the fame more than facts) 2. bai yun cang gou (Thing can change quickly and unpredictably) 3. qu wei cun zhen (discard the false and retain the true)

		3.ren dao wu qiu pin zi gao (There is nothing more noble than those who act without greed) 4. fa guang de wei bi du shi jin zi (All that glitter is not gold)	4.nang ying zhao shu (pursue knowledge under difficulties) Added four phases for Hokkien users 1. ah-á thiaⁿ lûi (to be completely beyond one's comprehension) 2. tsiȧh pá uānn iau (prove futile) 3. thâi-ke-kà-kâu (to punish a wrongdoer as a warning for others) 4. nńg-thôo-tshim-kṳ̀t (everybody leaves a load on a willing horse.)
Language- Proverb repetition	1. All that glitters is not gold. 2. A stitch in time saves nine.	1. chun jiang shui nuan ya xian zhi (The duck knows first when the river becomes warm in spring) 2. ye jing yu qinhuangyuxi (Achievement is founded on diligence and wasted upon recklessness)	*The simpflied Chinese proverbs remained, and* *added* two proverbs for Hokkien users: 1. Tsit iūnn bí tshī pah-iūnn lâng (One kind of rice provides for one hundred kinds of people) 2. Tshit tsē peh pê káu huat-gê (Everything happens in due time)
Language- Object naming	spoon; book; kangaroo or wallaby; penguin; anchor; camel or dromedary; harp; rhinoceros or rhino; barrel, keg, or tub; crown; crocodile or alligator; piano accordion, accordion, or squeeze box.	Change pictures: 1. from spoon to pencil 2. from kangaroo to panda	Remained the same as the UK English version: 1. Spoon 2. Kangaroo Change pictures: 3. from harp to drum (revised based in Taiwanese culture) 4. from barrel to bulb (referred to Japanese version)
Language- Comprehension	1. Point to the one which is associated with the monarchy 2. Point to the one which is a marsupial 3. Point to the one which is found in the Antarctic 4. Point to the one which has a nautical connection	1. Point to the one wore by a king 2. Point to the animal living in Wolong, Sichuan 3. Point to the one which is found in the Antarctic 4. Point to the one which has a nautical connection	1. Point to the animal which is a marsupial 2. Point to the animal which is a reptile 3. Point to the one which is found in the Antarctic 4. Point to the one which has a nautical connection

Language- Reading	Ask the participant to read the following words: 1. sew 2. pint 3. soot 4. dough 5. height	Ask the participant to read the following words: 1. jade sea 2. soft shift 3. warrior Hope 4. suburban surname	Ask the participant to read the following words: 1. attentively 2. so far 3. publish 4. dragon boat 5. quality
Perceptual abilities- Identifying letters	Ask the participant to identify the letters: K, M, A, T	Ask the participant to identify the letters: K, M, A, T	Ask the participant to identify the Arabic numbers: 8, 2, 4, 9

7.3.4 Participants

Participants were eligible if they were between the ages of 45 and 100 and had been referred to a neurological clinic at Taipei Medical University Shuang-Ho Hospital in Taiwan within the preceding six months with memory issues or suspicions of dementia. Excluded were those previously diagnosed with severe mental disorders or with visual, hearing, or motor impairments that severely hindered comprehension or performance on cognitive tasks.

7.3.5 Dementia diagnosis and recruitment

Table 10 displays the schedule of recruitment. Four neurologists based in the clinic diagnosed participants as having or not having dementia according to diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) prior to their recruitment for the study (American Psychiatric Association, 2013). They used medical history, daily functioning, reported cognitive symptoms, laboratory results, and neuroimaging findings to make the diagnosis. Clinical observation of significant cognitive decline in one or more cognitive domains and interference with independence in everyday activities were judged by the clinicians at consultation. Clinicians also classified dementia subtypes, based on the International Classification of Diseases, Tenth Revision (ICD-10). Neuropsychological assessments were conducted by three psychologists. I only had access to participants' CDR, MMSE, and CASI scores, but not to other test data.

Participants were informed about the study by neurologists or me in the clinic or through clinic posters and participant information sheets. Referring to Table 10, initially, I screened patients on the appointment list to find those who could qualify (CDR 0 – 2, dementia group: with a dementia diagnosis; non-dementia group: without a diagnosis of dementia nor MCI; both groups had no additional severe mental illness or sensory or motor impairments). I then informed their neurologists, who reviewed their eligibility and directed those who met the criteria and expressed interest to me, or I confirmed their willingness myself if the neurologists

were too busy (see *Appendix 23* for the poster, *Appendices 24-25* for the information sheet). I first gave the participants the participant information sheet, then explained the objectives and procedures of the study and confirmed their willingness to participate. I asked those who wished to take part in the study to complete the informed consent form (see *Appendices 26-27* for the consent form). I then requested personal data specified on the case report form such as the participant's date of birth, proficiency in Chinese reading and writing, second language, occupation and position, existence of sensory impairments, history of major, chronic, or psychiatric illness, and current and previous medications. The cognitive interview then began with the T-ACE-III, followed by the supplemental tests (discussed in *Chapter 8*).

Table 10: Schedule of the recruitment

Evaluation / Procedure	Screening	Registration	Validation
DSM-5 major NCD diagnostic criteria, ICD-10	✓		
Assess Eligibility (CDR, MMSE, CASI)	✓		
Informed Consent		✓	
Case Report Form		✓	
T-ACE-III			✓
Brixton Spatial Anticipation Test			✓
Two WAIS- IV- Processing Speed Subtests			✓
The Coin-in-the-Hand test			✓

Note. DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; NCD = neurocognitive disorder; ICD-10 = International Classification of Diseases 10th Revision; CDR = Clinical Dementia Rating; MMSE = Mini Mental Status Exam; CASI = Cognitive Abilities Screening Instrument; T-ACE-III = Taiwanese version of the Addenbrooke's Cognitive Examination III; WAIS-IV = Wechsler Adult Intelligence Scale-Fourth Edition

7.3.6 Power calculation

I determined the required number of participants for each group using "easyROC" (http://www.biosoft.hacettepe.edu.tr/easyROC/). With a power value of 90%, type I error of 0.05, area under the ROC curve of 0.70, and allocation ratio of 0.5, the required number of participants for the dementia and non-dementia groups were 48 and 24, respectively.

7.3.6 Statistics

I analysed data using SPSS (Version 26). I first described the demographic features of dementia and non-dementia groups, including sex, age, education, language, and raw cognitive test scores. I also provided cognitive test results, including the mean and standard deviation (SD) for parametric data and the median and interquartile range (IQR) for nonparametric data. From the 47th participant onwards, as my skills to administer cognitive tests improved and stayed consistent, I began recording the time required for assessments.

I calculated the diagnostic accuracy of the T-ACE-III to distinguish between dementia and non-dementia groups based on the area under the curve (AUC) with receiver operating characteristic (ROC) analysis. I utilised the Youden Index to determine the optimal cut-off point, which integrates sensitivity and specificity information into statistics and ranges from 0 to 1. A value of 1 indicates there are no false positives or false negatives (indicating a perfect test). I assessed for reliability using the domain-total internal consistency (Cronbach's -coefficient) of the T-ACE-III. I also tested for convergent validity by calculating the correlation between the T-ACE-III scores and the tests previously administered in the neurology clinic (CDR, MMSE, CASI), using a two-tailed Spearman's test.

To explore the impact of demographic factors on T-ACE-III scores, I used multiple linear regression analysis, with the T-ACE-III scores as the dependent variable and age, years of education, sex, socioeconomic status (managerial/nonmanagerial), and language as the independent variables. I considered possible independent predictors of dementia diagnosis by putting demographic variables (age, years of education, socioeconomic status (SES), and sex) and T-ACE-III scores into logistic regression analysis with dementia diagnosis as the outcome. In a post-hoc analysis, I examined the effects of education and age on language (speaking solely Taiwanese versus speaking both Mandarin and Taiwanese) using logistic regression. A ‘p’ value of <.05 was regarded as statistically significant.

7.4 Results

7.4.1 Participants' characteristics

I invited 110 people to participate in this study; 91 (82.7%) provided informed consent. Among the reasons for refusal were other doctor's appointments, poor sleep quality, and unwillingness to participate in this study.

Among those who consented, 90 out of 91 people completed the T-ACE-III (98.9%). In the dementia group, there were 57 participants with complete data, including 21 patients with Alzheimer's disease (AD), 15 with vascular dementia (VaD), 6 with Parkinson's disease dementia (PDD), 3 with frontotemporal dementia (FTD), 1 with dementia with Lewy bodies (DLB), 2 patients with unspecified dementia, 9 with mixed-type dementia (3 VaD+PDD, 4AD+VaD, 1AD+FTD, and 1AD+Trauma), and 33 in the non-dementia group. All participants completed the T-ACE-III.

Table 11 displays the demographic data of the participant groups, as well as their scores on neuropsychological assessment and on the T-ACE-III. Those diagnosed with dementia were older (74.2 vs. 66.4) and had lower levels of education than those without dementia (7.7 and 11.2 years of education). Fourteen (24.6%) of the dementia group's participants spoke only Taiwanese. All participants of the non-dementia group, however, spoke either Mandarin Chinese (15.2%) or both Taiwanese and Mandarin Chinese (84.8%).

Table 11: Demographic and neuropsychological characteristics of participant groups

Variable (number)	Dementia (N = 57)	Non-dementia (N = 33)
Male	**24 (42.1%)**	**12 (36.4%)**
Female	**33 (57.9%)**	**21 (63.6%)**
Education, years*	**7.70 ± 4.4**	**11.24 ± 4.0**
Illiterate	6 (10.5%)	1 (3.0%)
Elementary school	25 (43.9%)	5 (15.2%)
Junior high school	9 (15.8%)	6 (18.2%)
Senior high school	12 (21.1%)	11 (33.3%)
College	5 (8.8%)	9 (27.3%)
Higher education	0 (0.0%)	1 (3.0%)
Age, years*	**74.2 ± 8.4**	**66.4 ± 7.5**
< 60	5 (8.8%)	7 (21.2%)
61 - 70	13 (22.8%)	17 (51.5%)
71 - 80	23 (40.4%)	7 (21.2%)
81 - 90	15 (26.3%)	2 (6.1%)
> 90	1 (1.8%)	0 (0.0%)
Language used		
Taiwanese only	14 (24.6%)	0 (0.0%)
Chinese only or Chinese & Taiwanese	43 (75.4%)	33 (100.0%)
CDR score		
0	0 (0.0%)	28 (84.9%)
0.5	36 (63.2%)	5 (15.2%)
1	15 (26.3%)	0 (0.0%)
2	6 (10.5%)	0 (0.0%)
MMSE score*	**18.4 ± 5.2**	**27.7 ± 1.9**
CASI score*	**60.8 ± 16.3**	**91.2 ± 4.9**
Total T-ACE-III score*	**55.0 ± 15.8**	**90.8 ± 6.6**
Orientation/attention*	11.3 ± 4.0	17.4 ± 1.1
Memory*	11.7 ± 5.3	23.5 ± 2.3
Verbal fluency*	4.6 ± 3.2	10.5 ± 2.2
Language*	16.6 ± 5.1	24.2 ± 2.3
Visuospatial*	10.8 ± 4.0	15.3 ± 1.3

Note. CASI = Cognitive Abilities Screening Instrument; CDR = Clinical Dementia Rating Scale; MMSE = mini-mental state examination; T-ACE-III = Taiwanese version of Addenbrooke's Cognitive Examination III; * represents the results presented as Mean ± SD.

7.4.2 Administration time

I measured the time taken for testing in 43 participants. The T-ACE-III took a mean (SD) of 21.3 (6.7) minutes for all participants, 24.0 (6.1) minutes for those with dementia and 16.3 (4.6) minutes for those without dementia.

7.4.3 The T-ACE-III Scores for Discriminating Dementia from Non-Dementia

Figure 5 shows the ROC analysis. *Table 12* shows the sensitivity and specificity at different cut-off points and the Youden index. The T-ACE-III had an area under the ROC curve (AUC) of 0.990 (p = .001, 95% confidence interval (CI) = 0.977–1.000). As shown in *Table 12*, when the cut-off score was 73/74, the T-ACE-III had a specificity of 1 and a sensitivity of 0.895, with a Youden index value of 0.895, indicating the optimal cut-off point; when the cut-off score was 86/87, the T-ACE-III had a specificity of 0.788 and a sensitivity of 1, with a Youden index value of 0.788.

Six people with dementia and one without dementia were illiterate. I calculated and supplied the corrected T-ACE-III cut-off values (adding 3 points back for two items in language domains) for the population who had not received any education in *Table 13: T-ACE-III study- Cut-off scores for the T-ACE-III (adjusted) in differentiating dementia from non-dementia.*

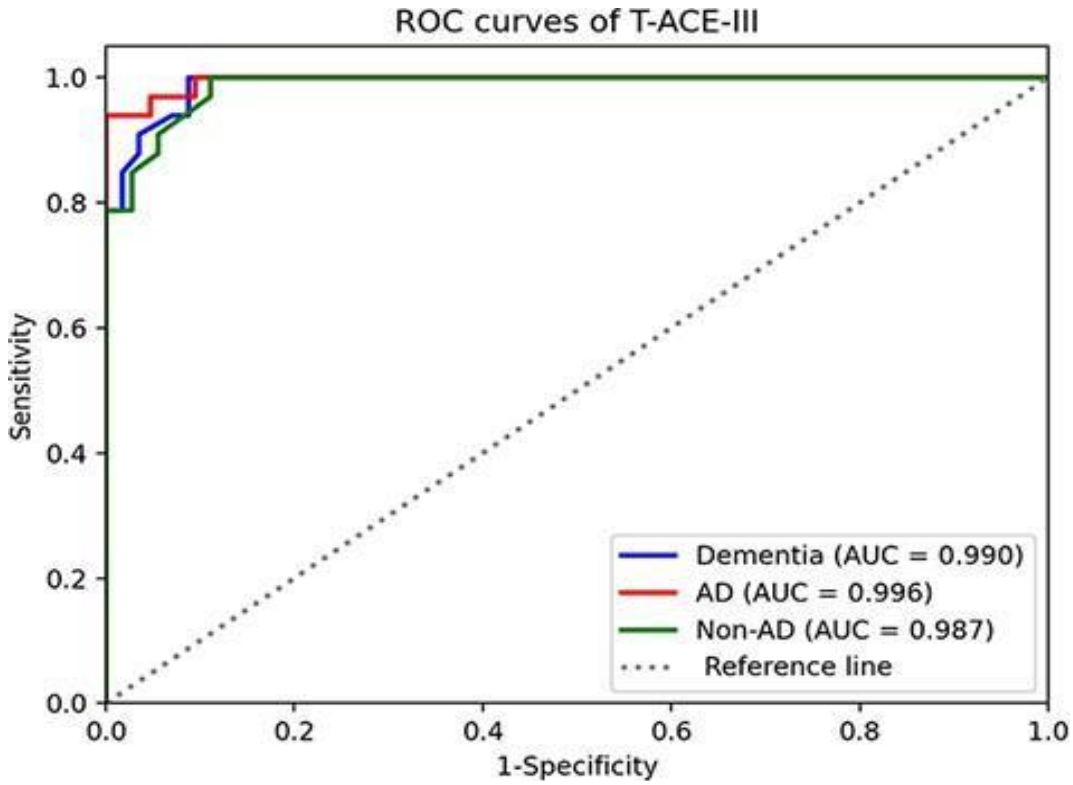


Figure 5: ROC curves of T-ACE-III.

Table 12: Cut-off scores for the T-ACE-III in differentiating dementia from non-dementia, sensitivity values, specificity values, and Youden index

Cut-off scores	Sensitivity	Specificity	Youden Index
73/74	0.895	1.000	0.895
74/75	0.895	0.970	0.865
75/76	0.912	0.970	0.882
76/77	0.930	0.939	0.869
77/78	0.947	0.939	0.886
80	0.965	0.909	0.874
82/83	0.965	0.879	0.844
83/84	0.982	0.848	0.830
84/85	0.982	0.818	0.800
85/86	0.982	0.788	0.770
86/87	1.000	0.788	0.788

Table 13: Cut-off scores for the T-ACE-III (adjusted) in differentiating dementia from non-dementia

Cut-off scores	Sensitivity	Specificity	Youden Index
73/74	0.895	1.000	0.895
74/75	0.895	0.970	0.864
75/76	0.912	0.970	0.882
76/77	0.930	0.970	0.900
77/78	0.947	0.970	0.917
78/79	0.965	0.939	0.904
80	0.965	0.909	0.874
82/83	0.965	0.879	0.844
83/84	0.982	0.848	0.831
84/85	0.982	0.818	0.801
85/86	0.982	0.788	0.770
86/87	1.000	0.788	0.788

7.4.4 The T-ACE-III score for discriminating AD and non-AD from non-dementia

As also seen in *Figure 5*, the AUCs of T-ACE-III for identifying AD and other dementia subtypes versus no dementia were 0.996 (p = .001, 95%CI = 0.986–1.000) and 0.987 (p = .001, 95%CI = 0.969–1.000), respectively, indicating that high diagnostic accuracy was not affected by the type of dementia.

7.4.5 Reliability and convergent validity of the T-ACE-III

The T-ACE-III had an internal consistency of 0.895. As expected, there was a negative correlation between the CDR score and the total T-ACE-III score (r = -0.9, p 0.01) and a positive correlation between the MMSE and CASI scores from the clinician assessment and the total T-ACE-III score (both r = 0.9, p = 0.01).

7.4.6 Association of demographic factors with the T-ACE-III score in linear regression

As shown in *Table 14*, using multiple linear regression with the T-ACE-III scores as the dependent variable, I found Taiwanese speakers tended to score lower compared to others (only Taiwanese compared to both Mandarin and Taiwanese, p =.005, B = -0.300, 95% CIS for B = −30.868 to −5.856) and people with more years of education also tended to have higher T-ACE-III scores (for every additional year of education, p = .009, B = 0.329, 95% CIS for B = 0.404 to 2.760). However, T-ACE-III scores were not correlated with age (for every year older, p = .213, B = -0.127, 95% CIS for B = −0.810 to −0.183), sex (male compared to female, p = .292, B = -0.105, 95% CIS for B = −13.406 to −4.085), and employment divided into managerial/ nonmanagerial jobs (p = .614, B = 0.052, 95% CIS for B = −6.864 to 11.550). The variance inflation factors (VIFs) were from 1.157 to 1.798 (lower than 10), indicating that no multi-collinearity existed among the independent variables (education years, age, sex, language spoken, and SES). The model accounted for 31.9% of the variance in the T-ACE-III scores overall model (F = 7.6, p = < .001).

7.4.7 Association of demographic factors with the language used

As also shown in *Table 14*, further analysis of the language used (Taiwanese only) through logistic regression, found that people with fewer years of education tended to use Taiwanese only (for every additional year of education, p = < .001, 95% CIS for Exp(B) = 0.586–0.847). The model accounted for 35.7% of the variance in the overall model (p< .001). However, age was not associated with the language used (Taiwanese only) (p =.935, 95% CIS for Exp(B) = 0.990–1.141).

7.4.8 Association of demographic factors with diagnosis of dementia on logistic regression

In the logistic regression model (see also *Table 14*), lower T-ACE-III ($p = .021$, 95% CIS for Exp(B) = 0.260–0.898) scores were associated with diagnosis of dementia. With T-ACE-III scores in the model, years of education ($p = .302$, 95% CIS for Exp(B) = 0.808–1.985), age ($p = .155$, 95% CIS for Exp(B) = 0.679–1.063), sex ($p = .667$, 95% CIS for Exp(B) = 0.005–28.690) and managerial/nonmanagerial jobs ($p = .221$, 95% CIS for Exp(B) = 0.140–4833.824) were not significantly independently associated with dementia diagnosis. The model accounted for 92.7% of the variance ($p < .001$). I was unable to add language to the model as everyone who spoke only Taiwanese was diagnosed with dementia.

Table 14: Association of demographic factors with T-ACE-III score and diagnosis of dementia

		Independent variables					
Dependent variables	Coefficients	T-ACE-III (higher)	Education (increasing years)	Age (higher)	Sex (male)	SES (managerial/ non-managerial)	Language (Taiwanese)
Linear regression (T-ACE-III score)	Standardized Coefficients β	-	0.329	-0.127	-0.105	0.052	-0.300
	95% CIS for B	-	0.404-2.760	-0.810-0.183	-13.406-4.085	-6.864-11.550	-30.868 – -5.856
	T	-	2.672	-1.256	-1.060	0.506	-2.921
	p	-	0.009	0.213	0.292	0.614	0.005
	R^2	0.319					
Logistic regression (Outcome diagnosis of dementia)	B	-0.727	0.236	-0.163	-0.945	3.259	-
	Exp(B)	0.483	1.267	0.850	0.389	26.031	-
	95% CIS for Exp(B)	0.260-0.898	0.808-1.985	0.679-1.063	0.005-28.690	0.140-4833.824	-
	p	0.021	0.302	0.155	0.667	0.221	-
	NagelkerkeR^2	0.927					
Logistic regression (Language used-Taiwanese only)	B	-	-0.35	0.061	-	-	
	Exp(B)	-	0.704	1.063	-	-	
	95% CIS for Exp(B)	-	0.586-0.847	0.990-1.141	-	-	
	p	-	<0.001	0.094	-	-	
	NagelkerkeR^2	-	0.357	0.058	-	-	

Note. B = unstandardized regression weight; Exp(B) = adjusted odds ratio; 95% CIS for Exp(B) = 95% confidence interval for adjusted odds ratio; T-ACE-III, Taiwanese version of Addenbrooke's Cognitive Examination-III. SES = socio-economic status. I could not include language in logistic regression as there were no Taiwanese speakers in the non-dementia groups.

7.5 Discussion

7.5.1 Summary of findings

I culturally modified and validated the T-ACE-III for Taiwanese populations, which is the first study that used this in detecting dementia in Taiwan. I determined the optimal cut-off scores for distinguishing dementia patients from non-dementia participants. In this study, my main findings were:

1. The T-ACE-III showed an excellent ability for distinguishing people with dementia from non-dementia, with an AUC of 0.990, a sensitivity of 0.895, and a specificity of 1, at the cut-off point of 73/74.
2. The T-ACE-III's high diagnostic accuracy was not impacted by the type of dementia, with AUCs of 0.996 and 0.987 for identifying AD and other dementia subtypes versus no dementia, respectively.
3. The T-ACE-III scores were significantly associated with the language spoken (only Taiwanese, or both Mandarin and Taiwanese) and years of education. Indicators of poorer T-ACE-III scores were Taiwanese users and fewer years of education.
4. Inclusion of other demographic variables (age, sex, years of education, and managerial/ nonmanagerial jobs) into the model did not improve the detection of dementia over and above T-ACE-III scores.

The T-ACE-III distinguished dementia from non-dementia in a memory clinic population with superior accuracy and a higher ROC AUC (0.990) than existing tests, such as CASI (0.985) or MMSE (0.970) in my study (see *Appendix 29 ROC curves of T-ACE-III, CASI, and MMSE*). It also had higher detection accuracy compared to the CASI (0.941), MMSE (0.85 - 0.972), and MoCA (0.80 - 0.964), as found in other studies in Taiwan (Chang et al., 2012; Tsai et al., 2016; Wang et al., 2013). The T-ACE-III scores had a high correlation with scores of currently used measures (MMSE, CASI, and CDR) in Taiwan to differentiate dementia from non-dementia, demonstrating good convergent validity (> 75%, Abma et al., 2016). It had a good internal consistency (0.895 , Cortina, 1993) and a very high acceptance (90 out of 91 people completed it). It was administered relatively quickly compared to the current assessments. According to the neurologists and

psychologists working in the Taiwanese clinics, the CASI, CDR, and Neuropsychiatric Inventory, when used for diagnosing dementia, take approximately one hour in total for each patient (personal communication, C-CY).

Age is the most significant risk factor for dementia, with a cut-off value of 65 years old, below which dementia is uncommon (Mukadam et al., 2022; van der Flier & Scheltens, 2005). Reflecting this, the non-dementia group in my study were younger (mean age 66.4 vs. 74.2) and had a higher level of education (mean years of education 11.2 vs. 7.2) than those in the dementia group as would be expected in people presenting to a memory clinic with suspected dementia (see *7.4.1 Participants' characteristics*), which is also consistent with numerous previous studies (Chang et al., 2010; Chang et al., 2012; Chong et al., 2010; Chu et al., 2015; Huang et al., 2019; Li et al., 2021; Liew et al., 2015; Low et al., 2020; Shi et al., 2012; Tsai et al., 2018; Wei et al., 2018; Xu et al., 2003; Yeung et al., 2014).

Personality traits, medical or psychological disorders, and cultural backgrounds may all have an influence on the onset of subjective cognitive decline (SCD) (Jessen et al., 2014). One study has pointed to a correlation between a high prevalence of SCD and the highest (tertiary) levels of education (Röhr et al., 2020). Rohr et al. indicated that this may be due to a heightened awareness of minor cognitive changes. A previous study (Kuhn et al., 2019) found that people with SCD from the clinic (age 68.3, 12.9 years of education) were younger and more educated than those from the community (age 71.7, 12.7 years of education) and healthy controls (aged 72.3, 11.5 years of education). More importantly, Individuals with SCD in a memory clinic context had an increased risk of dementia than those without SCD in contrast to those with SCD in a community-based setting who did not(Slot et al., 2019). Consequently, these individuals should return to the clinic for evaluation if they found their cognitive function or memory deteriorating.

On a policy level, the Taiwanese government launched a nine-year compulsory education system in 1968; therefore, most participants I interviewed in 2020, who were over the age of 65, had only completed six years of education. This would help explain why those in non-dementia group who were younger had higher

levels of education. The Lancet Commission has identified childhood education as the main risk factor in early life for dementia, and it should be a priority for all countries and individuals worldwide (Livingston et al., 2020). Taiwan has audited and set to start the 12-year Basic Education Project in 2014, and the new era of education began in 2019, which is encouragingly in line with the global trend.

7.5.2 Factors associated with level of cognitive functioning

To gain a more comprehensive understanding of the influence of different Chinese languages (local dialect, such as Hokkien) on T-ACE-III scores, I recruited nine more participants to each group. Multiple linear regression analysis revealed that only years of education and spoken language being Taiwanese were significant predictors of T-ACE-III scores (31.9%), while other demographic factors did not account for any variance. Language spoken (those who only spoke Taiwanese compared to those who spoke both Mandarin and Taiwanese) and years of education were significant predictors of ACE scores, and this may be because those who only spoke Taiwanese had less education, which supports previous research indicating that education levels are negatively correlated with the use of Taiwanese language (Tsao & Lu, 1999). People who spoke exclusively Taiwanese had less education, which may have been a risk factor for dementia and contributed to lower T-ACE-III scores. Nevertheless, the clinical diagnosis was made prior to testing with T-ACE-III, so was independent of the scores.

7.5.3 Factors associated with status of dementia

In terms of the potential predictors of the diagnosis of dementia, logistic regression analysis found that T-ACE-III scores explained the vast majority of the variance and there was no room for other demographic factors to further contribute. In summary, in the cultural context of Taiwan, although demographic factors of language and education years have some influence on the T-ACE-III

scores, it does not add to knowing how effective the T-ACE-III scores are at classifying people as having dementia or not. A previous study (Lin et al., 2019) attempted to validate the traditional Chinese version of ACE-III in four day-care community centres rather than in a clinical setting, with the aim of detecting MCI in Taiwan. The translation did not follow cultural adaptation guidelines or use any gold standards in respect of MCI diagnosis. My own study followed specific guidelines for translation and cultural adaptation (Beaton et al., 2000), and incorporated input from a bilingual postdoctoral researcher (C-CY), two neurologists (C-TH, L-KH), two clinical psychologists (J-CL, P-JW) from Taiwan, two psychiatrists (GL, NM), and two clinical psychologists (NK, JS) from the UK. It is now freely available on the website of the FRONTIER Research Group (www.ftdrg.org) for anyone to use.

7.5.4 Limitation of the validation study

There are a few limitations to this study. Firstly, to minimize selection bias, I recruited a consecutive sample of individuals and discovered that all those who solely spoke Taiwanese in the current study had dementia, which may be linked to early-life education deprivation as a result of Taiwan's education policies. From 1945 to 1987, the KMT (Nationalist) government adopted a strict Mandarin Language Policy in Taiwan's schools, prohibiting students from speaking local dialects such as Hokkien (Sandel, 2003). Consequently, most relatively older, or non-schooled people were unable to learn Chinese. In addition, six participants in the dementia group and one person in the non-dementia group were illiterate but I have provided the cut-off values of the adjusted T-ACE-III for people who were illiterate (*Table 13: T-ACE-III study- Cut-off scores for the T-ACE-III (adjusted) in differentiating dementia from non-dementia*). I recommend that future research use a larger sample size to establish different T-ACE-III cut-off values for participants with different educational levels and levels of literacy, so as to provide more tailored cut-off scores in relation to educational and literacy needs, thereby enhancing the diagnostic accuracy. Secondly, I calculated convergent validity but did not include other measures of reliability, such as test–

retest reliability in the study. Thirdly, I made a mistake on one question in the attention domain (orientation) due to unclear instructions in the simplified Chinese version that ask, ‘where do you live?’ instead of ‘where are you (we) at the moment?’. This question has been corrected on the final version of T-ACE-III and each participant’s score on this item has been replaced with their scores on an identical MMSE item. Fourthly, the study could not be double-blind since patients received a diagnosis as part of their routine clinical evaluation. This could lead to potential rater bias when administering psychometric test as both the patient and the researcher were aware of the diagnosis, although the T-ACE-III cut-off score was not known. Lastly, the administration time was slightly longer than 20 minutes, which may have been due to a number of factors, including the mask policy during the Covid-19 pandemic that influenced communication, the fatigue induced by some evening interviews, and the noise from the clinic next door during the interviews. However, these external factors were not able to be avoided or predicted.

7.6 Conclusions

The T-ACE-III is a valid, reliable, acceptable, and effective tool to assist in the diagnosis of dementia with a very high AUC in ROC analysis and a short administration time compared to other test batteries. It has higher validity in distinguishing dementia from non-dementia than currently-used dementia routine diagnostic tools, and its ease of administration and accuracy may make it the cognitive assessment tool of choice in clinical settings in Taiwan.

Chapter 8. Acceptability of tests of executive function, speed of processing and performance validity

In the previous chapter, I detailed how I carried out the validation study of the T-ACE-III in Taiwan, including background work (ethics approval, authors' permissions, and collaboration invitations) prior to the formal study, and the actual steps of data collection and analysis that I took to conduct the study. In this chapter, I will begin by outlining the reasons why I chose the supplementary cognitive tests in my study. These were the Brixton Spatial Anticipation Test, speed of processing subtests from the Wechsler Adult Intelligence Scale IV, and the Coin-in-the-Hand (CIH) test. I will introduce each test, present the results of my study, and discuss the implications of my findings.

8.1 Background

8.1.1 Executive function

Diamond (2013) noted that "*Executive functions (EFs; also called executive control or cognitive control) refer to a family of top-down mental processes needed when you have to concentrate and pay attention, when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible* (Burgess & Simons, 2005; Espy, 2004; Miller & Cohen, 2001)*." (p.135)*. However, as it is a complicated function, there is no full agreement as to what executive function is and what it performs (Baggetta & Alexander, 2016). Executive function is affected by dementia. According to DSM-5 (Please see *2.1.1 Diagnostic criteria for dementia*) the cognitive domains relevant to executive function include **planning, decision-making, working memory, responding to feedback, inhibition, and mental flexibility**. Certain components of executive functions, such as **inhibitory control and mental set-shifting**, can be impaired even in the earliest stages of Mild Cognitive Impairment (MCI) (Junquera et al., 2020; Kim et al., 2016). In addition, people with the dysexecutive phenotype of MCI have been shown to have a higher likelihood of developing dementia within

a year than the amnestic variant of MCI (Junquera et al., 2020). It is therefore important to include executive functions (especially switching and category generation) in the evaluation of dementia, as they may be affected early in the stages of the disease, cause varying deficits, and impact on everyday functioning (Lamar et al., 2010). In my study, I aimed to consider more sub-domains of executive functions covered by the T-ACE-III, and explore both test performance and test acceptability in patients with dementia in Taiwan.

A systematic review (de Faria et al., 2015) identified the most commonly used tests (usage > 20% of 25 selected articles) for measuring executive function in individuals aged over 50. These tests included the Trail Making Test-Form B (TMT-B), Verbal Fluency Test (letters F, A, and S), Verbal Fluency Test (animals), Clock Drawing Test, Digit span subtest from the Wechsler Adult Intelligence Scale, Stroop test, and the Wisconsin Card Sorting Test. Most tests identified in this review tapped a specific aspect of executive functions, of which four were described by DSM-5: **mental flexibility, planning, working memory**, and **inhibitory control**. The remaining two domains were **verbal fluency** and **processing speed**. Only one of the 25 selected publications examined five sub-domains (mental flexibility, verbal fluency, planning, working memory, processing speed) of executive functions (Parra et al., 2012), whereas the majority of studies combined three to four sub-domains into a holistic evaluation. The most focus was placed on **mental flexibility**, whereas **processing speed** received the least attention.

It is worth noting that the classification of **processing speed** as a sub-domain of executive functions is not universally accepted. Nonetheless, for the purpose of this study, **processing speed** has been included as one of the aspects of executive functions. Conversely, the sub-domain of **inhibitory control** has not been considered due to its features have not been thoroughly characterized and the there are few tests to choose from (O'Callaghan et al., 2013). My thesis focuses on five specific sub-domains of executive functions, namely **mental flexibility, verbal fluency, planning, processing speed,** and **working memory**.

8.1.2 Further exploration of executive functions

In the next section, I address the three sub-domains of executive function (**mental flexibility, processing speed,** and **working memory**) that were not tested by the T-ACE-III.

8.1.2.1 Mental flexibility

The Brixton Spatial Anticipation Test (Burgess & Shallice, 1996) requires a simple motor or verbal response, making it ideal for participants who are not fluent in speaking. It assesses mental flexibility, such as the ability to detect and follow a new rule (see *Figure 6*). Mental flexibility is regarded as a crucial component of executive function and can be broadly defined as *"the ability to adjust behaviours in response to environmental changes* (Webster et al., 2016*, p.235)*" or *"the ability to shift attention between tasks, attributes of a stimulus, responses, perspectives, or strategies"* (Perone et al., 2018*, p.299)***.** The Brixton test has been widely used in Europe but appears to have never been used with any Chinese-speaking population. It was designed in part to tackle the issue of interpreting errors when using the Wisconsin Card Sorting Test, such as "guessing behaviours" being considered as using a prior correct rule (Burgess & Shallice, 1996; Grant & Berg, 1948). In addition, the test does not need a complex verbal or sophisticated motor response and has no time constraint. Therefore, it is attractive in meeting the goal of recruiting people with different levels of education or those who have a motor disability.

8.1.2.2 Processing speed

In the case of processing speed, this is the ability to do simple, repeated cognitive tasks rapidly and precisely, often involving several motor and cognitive skills. It has attracted less attention compared to other components of executive function (de Faria et al., 2015), but it has now been highlighted as one of the diagnostic factors for neurocognitive disorders in the cognitive domain of "complex attention"

in DSM-5 (Lu et al., 2017). There is a positive correlation between processing speed and successful ageing, which is characterised as (1) absence of major disease (2) high function, and (2) active engagement with life (Bosnes et al., 2022). The Processing Speed Index (PSI) from the WAIS-IV has a moderate accuracy (AUC area = .78) in differentiating patients with AD from those without dementia (Adishesa & Halim, 2015). The PSI in WAIS-IV has been adapted into a Taiwanese version (Wechsler, 2015a, 2015b), but little research has been conducted on the processing speed performance of patients with dementia in Taiwan.

8.1.2.3 Working memory

In the case of working memory, this generally includes retaining and actively processing no longer perceptually present information (Baddeley & Hitch, 1994; Diamond, 2013; Smith & Jonides, 1999). Working memory is divided into verbal and nonverbal (visual-spatial) working memory. Making meaning of anything that develops over time requires working memory, such as writing a paragraph, performing mathematics, and reasoning (finding links between things that at first glance appear unconnected). Working memory differs from short-term memory in that the former not only retains but also manipulates the information, whilst the latter merely retains it (Diamond, 2013).

8.1.2.3 Selected tests for executive functions in my study

The present study aimed to explore the acceptability of these supplementary tests for patients with dementia in Taiwan. As demonstrated in *Table 15*, the T-ACE-III only covers the **verbal fluency** and **planning** aspects of executive function but does not cover other aspects to any detailed extent. I therefore chose the Brixton Spatial Anticipation Test to test **mental flexibility** and **working memory** and two tests from the WAIS-IV to test **processing speed** and **working memory** as

supplementary tests to evaluate three additional areas of cognitive functioning that the T-ACE-III does not cover.

Table 15: The executive domains covered by the tests used in this thesis

Sub-domains of executive functions	Test
Mental flexibility	Brixton Spatial Anticipation Test
Verbal fluency	T-ACE-III – verbal fluency subtest
Planning	T-ACE-III – Clock drawing test
Processing speed	Digit Symbol Coding & Symbol Search subtests of WAIS-IV
Working memory	Brixton Spatial Anticipation Test, Symbol Search subtests of WAIS-IV

Note. WAIS-IV = Wechsler Adult Intelligence Scale-Fourth Edition.

8.1.3 Assessment of Malingering in Taiwan

In Taiwan, it is estimated that around one-fifth to half of applicants for labour insurance disability payment feign their memory impairment in order to qualify for payment (Lin, 2011). It is therefore, important to develop a practical tool for Taiwan's clinicians to assess for malingering (Yeh et al., 2019).

The Coin-in-the-hand (CIH) test (Kapur, 1994) (content of test described below: *8.2.2.3 the Coin-in-the-hand test*) was designed to assist clinicians in differentiating patients with neurocognitive impairment and those who are exaggerating or faking memory complaints. It only takes 5 minutes to administer and has been found to be helpful in a range of clinical populations, including those with dementia (Schroeder et al., 2012), head injury (Kelly et al., 2005), amnesia (Hanley et al., 1999) and memory impairment (Cochrane et al., 1998). In 2019, there was a study of a digital version of the CIH (Daugherty et al., 2019), and one pilot study of the test in Taiwan (Yeh et al., 2019). This test is recommended to be used in combination with other regularly used tests (Schroeder et al., 2012).

8.1.3 Aim

I aimed to explore the acceptability of using these supplementary tests (Brixton Spatial Anticipation test, WAIS-IV speed-of-processing subtests, and the CIH test) in individuals with suspected dementia presenting in a clinic setting in Taiwan. I also aimed to investigate how Taiwanese patients with dementia perform on those tests. My main research question was:

1. What is the acceptability of the supplementary tests in Taiwan?

8.2 Methods

8.2.1 Participants

This study was given ethics approval as part of the ethics application for the T-ACE-III study described in *Chapter 6* (please find *6.3 Methods*). All participants gave informed consent prior to testing.

8.2.2 Measures

8.2.2.1 The Brixton Spatial Anticipation Test

The Brixton Spatial Anticipation Test, which measures the ability to detect and stick with a rule, and also to change rules when there is a rule change, is a part of the Hayling and Brixton test battery (Burgess & Shallice, 1996). Participants are presented with a 56-page test booklet, with each page contains two rows of 5 circles numbered 1 to 10. Only one circle is coloured blue on each page, and its position varies on the basis of simple rules which change without warning. One page at a time is shown to participants, and they are asked to point to where they think the blue circle will appear on the next page.

Before the formal testing begins, the following question is posed “if I turn the page and the blue circle is at position 8, and I turn the page again, it goes to position 9, where do you think it goes to if I turn the page now?” If their response is 10, it

indicates they have understood how this test operates and what they have to do. If they forget the instructions during the test, they can also ask any questions for clarification.

As shown in *Figure 6*, the participant is shown page 1 where the blue circle is at location 5. They are then required to respond (or point to) the place where the blue circle is on the following page prior to turning the page, which the researcher does without offering any hints about the rule.

The maximum number of errors in 55 trials is 55 - higher scores reflect worse performance. In previous research, patients with MCI or early dementia made 20.0 errors compared to 18.7 errors of non-dementia controls (van den Berg et al., 2009). Cut-off scores are 29 for differentiating dementia from non-dementia in the original study (Burgess & Shallice, 1997) and 33 for in a population aged 66–80 (Bielak et al., 2006).

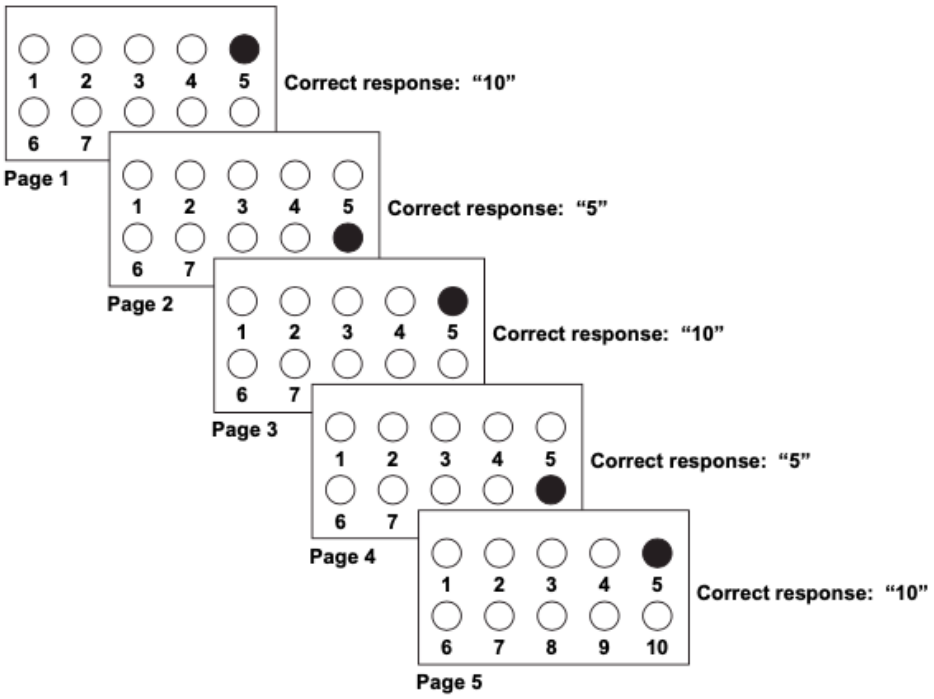


Figure 6: A set of responses from the Brixton test (from van den Berg et al., 2009, p. 697)

8.2.2.2 Speed of processing subtests from the Wechsler Adult Intelligence Scale-IV

The WAIS-IV-Processing Speed Index (PSI) one of four indices that make up the full-scale intelligence quotient derived from the WAIS-IV (Wechsler, 2008). The WAIS-IV-PSI includes three tasks, the Digit Symbol Coding (CD), Symbol Search (SS) subtests, and the supplemental Cancellation (CA) subtest, measuring the ability to process simple coloured visual stimuli quickly and efficiently. The PSI requires participants to process and respond to visual stimuli as quickly as possible, taking around 6-8 minutes for all three tests.

As specified in the test instructions, the participants were given some practise trials before the scored trials; for instance, I demonstrated how to find the two targets from the array of search figures in one trial, and the patient then practised in two trials for the Symbol Search subtest. For the Digit Symbol Coding subtest, I demonstrated how to replace the numbers with symbols for three items, and they practised for six items. Before the scored trials began, I explained about the time limit, and that they should perform the task as quickly as possible. When I say "time's up!", they should stop completing the form.

I only included the Digit Symbol Coding (CD) and Symbol Search (SS) subtests of the WAIS-IV-Processing Speed Index in my study.

8.2.2.3 The Coin-in-the-Hand test

In the Coin-in-the-hand (CIH) test (Kapur, 1994), a new Taiwan ten-dollar coin (diameter = 2.5 cm, cupronickel, 7.5g) was used in each of 10 trials. Firstly, the participants were asked to remember in which of the tester’s hands a coin is held. Secondly, they were asked to close their eyes and count backward from 10 to 1. Thirdly, after finishing counting, they were asked to open their eyes and point to the hand that held the coin. Both hands are then opened to show which one holds the coin. The total possible score is 10 (1 for each trial answered correctly). The

sequence of hands was suggested by Hanley et al. (1999) as follows: right, left, left, right, right, left, right, left, right, left. The test can be completed in 5 minutes.

Kapur (1994) administered this test to five amnestic patients and two suspected malingerers. People with amnesia were correct in all answers while the suspected malingers performed no better than chance level. Cochrane, Baker, and Meudell (1998) found no memory-impaired groups nor healthy controls made >1 error, while the simulated malingerers who were asked to feign memory problems made > 3 errors (Mean score = 6.1). Similarly, Hanley, Baker, and Ledson (1999) found that amnesics and controls group made >1 error, while the simulated malingerers made > 5 errors (Mean score = 4.1). These studies suggested that the CIH test might be a useful screening instrument for differentiating between individuals who feigns memory problems and patients with true memory impairments.

8.2.3 Translation and cultural adaptation

Since there was no cultural adaptation of the Brixton Spatial Anticipation test or the CIH test that was needed, I directly translated the administration instructions of both tests from English into traditional Chinese after consulting with a bilingual neurology postgraduate researcher, Cheng-Chang Yang. I also gained permission to use both tests in the Taiwanese population from their creators (Brixton test: Professor Burgess, a professor from the Institute of Cognitive Neuroscience at UCL; CIH: Professor Kapur, a member of my supervision panel, see *Appendices 20 - 21*), incorporating the original author's thoughts on administration processes for the CIH exam. Before formal recruiting of participants, I interviewed three nurses, one doctor, and one psychologist at the neurology department to check whether the instructions were clear and to adjust the content based on their feedback. No changes were made for test questionnaires, but only some terms or phases we used in interviewing the participants (see *Appendices 30 – 31* for the Brixton test and the CIH test).

For the WAIS-IV speed-of-processing tests, the traditional Chinese versions were available in Taiwan; I utilised them without further translation or cultural adaption.

8.2.4 Completion of tests

I administered the cognitive assessments to each participant. If a participant had trouble understanding the instructions at the beginning despite receiving further guidance, or if they were unable to complete the test midway due to factors such as fatigue or comprehension issues, their performance was classified as "incomplete."

8.2.5 Statistical Methods

I analysed data using SPSS statistics (version 21). I operationalised measuring acceptability by examining ‘completion rates’. I then further considered whether people with all severities of dementia completed them. I firstly documented how many participants agreed to take part and completed each test as well as how many did not, followed by a descriptive analysis that provides demographic data of those completing the tests. I considered if people with all grades of severity – as measured by the CDR – completed them, including those with and without dementia. Then, I compared the performance of individuals from the dementia and non-dementia groups on the Brixton Spatial Anticipation test (number of errors and standardised score), the WAIS-IV processing speed test, and Coin-in-the-hand (CIH) test scores, using independent T-tests, with the level of statistical significance set at 0.05.

8.3 Results

8.3.1 Completion rate

As shown in *Table 16*, the overall completion rates for the Brixton Spatial Anticipation Test, the Symbol Search (SS) and Digit Symbol Coding (CD) subtests of the WAIS-IV, and the Coin-in-the-Hand (CIH) test, among the 90 participants that completed the T-ACE-III, was high for the participants from the non-dementia group (97 - 100%). However, people with dementia had a high

completion rate for the CIH test (98.2%), but low completion rates for the remaining tests (57.9 - 66.7%).

Comparing the completion rates of individuals with different CDR scores in the dementia group, those with very mild dementia (scored 0.5, n = 36) performed satisfactorily on all the supplementary tests (77.8 - 100.0% completion rate, n = 28-36), while those with mild and moderate dementia (scored 1 and 2, n = 15 and 6), had low completion rates (26.7 - 46.7% and 16.7%, n = 4 - 7 and 1), with the exception of the CIH test (completion rate of 93.3% and 100.0%, n = 14 and 6) (see *Table 16*).

Table 16: The completion rate of dementia and non-dementia groups

Test	Group (N)	CDR = 0 (0, 28)	CDR = 0.5 (36, 5)	CDR = 1 (15, 0)	CDR = 2 (6, 0)	All (57, 33)
Brixton Spatial Anticipation Test	Dementia (57)	-	30 (83.3%)	7 (46.7%)	1 (16.7%)	**38 (66.7%)**
	Non-dementia (33)	28 (100.0%)	5 (100.0%)	-	-	**33 (100.0%)**
WAIS-IV - Symbol Search	Dementia (57)	-	30 (83.3%)	4 (26.7%)	1 (16.7%)	**35 (61.4%)**
	Non-dementia (33)	27 (96.4%)	5 (100.0%)	-	-	**32 (97.0%)**
WAIS - IV - Digit Symbol Coding	Dementia (57)	-	28 (77.8%)	4 (26.7%)	1 (16.7%)	**33 (57.9%)**
	Non-dementia (33)	27 (96.4%)	5 (100.0%)	-	-	**32 (97.0%)**
Coin-in-the -Hand test	Dementia (57)	-	36 (100.0%)	14 (93.3%)	6 (100.0%)	**56 (98.2%)**
	Non-dementia (33)	28 (100.0%)	5 (100.0%)	-	-	**33 (100.0%)**

Note. CDR = Clinical Dementia Rating; WAIS-IV = Wechsler Adult Intelligence Scale-Fourth Edition.

8.3.2 Demographic and clinical profile

Please see *Chapter 7* for the participant's' demographic and clinical profile (see *7.4.1 Participants' characteristics* and *Table 11: Demographic and neuropsychological characteristics of participant groups*).

8.3.3 Administration time

I measured completion time for 17 people with dementia and 15 people without dementia on the Brixton Spatial Anticipation Test, for 15 and 14 people with and without dementia on two processing speed tests from WAIS-IV, and for 27 and 15 people with and without dementia on the CIH test. The mean time spent for the dementia and non-dementia groups was 9.6 (SD = 2.0) and 7.9 (SD = 1.2) minutes for the Brixton Spatial Anticipation Test, 8.2 (SD = 2.0) and 6.6 (SD = 1.5) minutes for the speed-of-processing sub tests of the WAIS-IV, and 2.5 (SD = 0.8) and 2.0 (SD = 0.4) minutes for the CIH test, respectively.

8.3.4 Performance on supplementary tests

As shown in *Table 17*, the participants from the dementia group had statistically poorer performance than those from non-dementia group the WAIS-IV symbol search (mean score 6.6 vs 10.7, p <.001) and Digit Symbol Coding (mean score 6.9 vs 9.5, p <.001) subtests. They were also much poorer than controls on the Brixton Spatial Anticipation Test – for number of errors (32.2 vs 21.2, p <.001) and for profile score (2.2 vs 4.3, p <.001). The dementia group were slightly lower than controls on the CIH test (mean score 9.7 vs 9.9, p = .02).

Table 17: The performance on supplementary tests by dementia and non-dementia groups in Taiwan (with numbers completing each test)

Test		Dementia (N = 57)	Non-dementia (N = 33)	T value	P value
WAIS-IV Symbol Search (35/32)[a]		**6.6 ± 2.4**	**10.7 ± 3.4**	**5.672**	**<.001**
	CDR = 0 (0/27)	-	10.9 ± 2.9		
	CDR = 0.5 (30/5)	6.4 ± 2.5	9.2 ± 5.6		
	CDR = 1 (4/0)	7.8 ± 1.3	-		
	CDR = 2 (1/0)	6.0 ± 0.0	-		
WAIS-IV Digit Symbol Coding (33/32)[a]		**6.9 ± 3.0**	**9.5 ± 2.6**	**3.638**	**<.001**
	CDR = 0 (0/27)		9.7 ± 2.4		
	CDR = 0.5 (28/5)	6.9 ± 3.2	8.0 ± 3.7		
	CDR = 1 (4/0)	7.3 ± 1.7	-		
	CDR = 2 (1/0)	7.0 ± 0.0	-		
Brixton test (38/33)					
	Number of errors[a]	**32.2 ± 10.2**	**21.2±7.6**	**-5.110**	**<.001**
	CDR = 0 (0/28)	-	20.6 ± 7.4		
	CDR = 0.5 (30/5)	31.3 ± 10.2	24.2 ± 9.0		
	CDR = 1 (7/0)	34.3 ± 9.6	-		
	CDR = 2(1/0)	47.0 ± 0.0	-		
	Standardized score[a]	**2.2 ± 1.6**	**4.3 ± 2.2**	**4.542**	**<.001**
	CDR = 0 (0/28)	-	4.5 ± 2.2		
	CDR = 0.5 (30/5)	2.3 ± 1.6	3.2 ± 2.5		
	CDR = 1 (7/0)	1.7 ± 1.3	-		
	CDR = 2(1/0)	1.0 ± 0.0	-		
The CIH test (56/33)[a]		**9.7 ± 0.7**	**9.9 ± 0.4**	**2.363**	**0.02**
7 points[b.]	CDR = 0 (0/28)	-	0		
	CDR = 0.5 (36/5)	0	0		
	CDR = 1 (14/0)	1	-		
	CDR = 2 (6/0)	0	-		
8 points[b]	CDR = 0 (0/28)	-	1		
	CDR = 0.5 (36/5)	3	0		
	CDR = 1 (14/0)	0	-		
	CDR = 2 (6/0)	1	-		
9 points[b]	CDR = 0 (0/28)	-	0		
	CDR = 0.5 (36/5)	3	0		
	CDR = 1 (14/0)	3	-		
	CDR = 2 (6/0)	1	-		
10 points[b]	CDR = 0 (0/28)	-	27		
	CDR = 0.5 (36/5)	30	5		
	CDR = 1 (14/0)	10	-		
	CDR = 2 (6/0)	4	-		

Note. T-ACE-III = Taiwanese version of Addenbrooke's Cognitive Examination III; WAIS-IV = Wechsler Adult Intelligence Scale-IV. CIH = Coin-in-the-Hand test [a] Represents the results presented as Mean ± SD; [b] Represents the results presented as frequency of specific scores.

8.4 Discussion

In this study, the majority of participants without dementia completed the supplementary tests. However, some participants with dementia did not, which may be due to a number of reasons. People who had previously finished the T-ACE-III may have felt exhausted and may not have wanted to continue with the supplementary tests due to fatigue. It was challenging for people with dementia, particularly those with CDR scores of 1 or 2, to comprehend the rules of the Brixton test and WAIS subtests, which is in line with the previous findings that indicated patients with more severe cognitive deficits may have difficulties in performing complex executive tasks such as the Brixton Test (van den Berg et al., 2009). Those who did not complete the Brixton test either did not comprehend the rules or were repeatedly failing during the test.

In the case of the Brixton Test, the completion rates for individuals without dementia were 100%, the completion rates were satisfactory (83.3%) for patients with very mild dementia (CDR = 0.5), and were low (46.7%, 16.7%) for those with mild (CDR = 1) and moderate dementia (CDR = 2), respectively. The mean error rate in dementia patients (32.2) was higher than in a previous study of patients with MCI or early dementia (20.0) (van den Berg et al., 2009).

In the case of the Symbol Search and Digit Symbol Coding subtests of the WAIS-IV, completion rates were high (96.4-100.0%) for all participants in the non-dementia group, satisfactory for patients with very mild dementia (77.8%, 83.3%), and low for those with mild (26.7% in both tests) and moderate (16.7% in both tests) dementia.

In the case of the CIH test, both the dementia and non-dementia groups had good completion rates, regardless of their CDR scores.

8.5 Conclusions

The Brixton Spatial Anticipation Test, as well as the Coding and Symbol Search subtests of the WAIS-IV-PSI, exhibited a low acceptability (operationalised as completion rate) in people presenting to a memory clinic who had mild and moderate dementia. Therefore, they are recommended to only administer to people with very mild dementia. In addition, sufficient rest should be provided before applying the supplementary tests. While executive tasks might be particularly useful in discriminating AD from FTD, due to the limited number of participants with FTD in this study, comparisons in performance across different dementias, particularly those with frontal lobe dysfunction, could not be carried out.

Chapter 9. General Discussion

In this chapter, I firstly summarise the findings of my systematic review and meta-analysis of studies looking at brief cognitive assessment tools used in Chinese-speaking populations and the narrative review of studies. I address the key findings of the T-ACE-III study and the supplementary tests study from my research in Taiwan. I then compare my findings with those from previous studies to discuss similarities and differences, and the methodological challenges in conducting a meta-analysis of diagnostic test procedures. Finally, I outline the merits and limitations of my systematic review and my empirical study, the clinical and policy implications of my findings, and suggestions for future research on the topic of validation and cultural adaptation of cognitive assessments.

9.1 Summary of findings

9.1.1 Systematic review, meta-analysis, and narrative synthesis

In my systematic review and meta-analysis study *(Chapter 4)*, I used a systematic search strategy with English databases to find all relevant articles in Chinese and English (see *4.2.2 Data sources, 4.2.3 Search strategy, 4.2.4 Study selection*). I performed a robust assessment of study quality (see *4.2.7 Quality assessment*), so as to identify the optimal brief cognitive assessments used to detect suspected dementia in Chinese-speaking populations. In addition, I conducted a sensitivity analysis to compare the different findings from univariate and bivariate approaches in the meta-analysis (see *4.3.3.4 Sensitivity analysis of univariate and bivariate analysis*).

I identified the ACE-R and the ACE-III, as having the best diagnostic performance and acceptable heterogeneity on the basis of 21 studies included in the meta-analysis. I observed a low sensitivity or specificity in CDT, STT-A, and STT-B in detection of dementia (see *4.3.3.2 Diagnostic test accuracy and optimal cut-off scores for all cause dementia*), as well as a high level of between-study heterogeneity in MMSE, MoCA, and MoCA-BC for MCI (see *4.3.3.3 Diagnostic*

test accuracy and optimal cut-off scores for MCI). Among the 37 studies included in the systematic review, MMSE was the most commonly tested brief cognitive test among Chinese-speaking populations, with 13 studies using it.

I found the population (China vs others) and MCI subtype (MCI vs others) significantly impacted MMSE scores, and the dementia scoring system (original vs adjusted) had an effect on MoCA in the meta-regression (see *4.3.3.5 Meta-regression*).

In my analysis of reported translation and cultural adaptation of brief cognitive instruments used in Chinese-speaking populations (*Chapter 6*), I found that only four studies fully reported on the procedures they used for translation and cultural adaptation (see *6.3.2.1 Limited reporting on translation and cultural adaptation*). At that time, the most recent version of the ACE, the ACE-III for Taiwan, was not published or on the website (www.ftdrg.org, or www.sydney.edu.au/brain-mind/resources-for-clinicians/dementia-test.html). I therefore decided to validate it in Taiwan.

9.1.2 The T-ACE validation study

In *Chapter 6*, I presented the adaptation of the ACE-III for a Taiwanese population (T-ACE-III). The ACE-III covers a broad range of cognitive domains and exhibited promising sensitivity and specificity with low heterogeneity between studies in my systematic review but was not yet available for Taiwan. I validated the T-ACE-III at a hospital in the northern part of Taiwan, recruiting 90 participants. There were 57 participants in the dementia group (21 AD, 15 VaD, 6 PDD, 3 FTD, 1 DLB, 2 unspecified dementia, and 9 mixed-type dementia), and 33 participants in the non-dementia group. I followed published guidance for translation and cultural adaptation as well as recommendations from the experts in the background of psychiatry, clinical psychology, neurology, and linguistics.

The T-ACE-III demonstrated an excellent ability to differentiate dementia from non-dementia, with an AUC of 0.990, a sensitivity of 0.895, and a specificity of 1.

The diagnostic power was unaffected by the type of dementia, with AUCs of 0.996 and 0.987 for identifying AD versus other dementia subtypes (non-AD), and versus no dementia, respectively.

In a linear regression model, I discovered that two demographic variables, language used and years of education, were significantly associated with T-ACE-III scores. However, when predicting the status of dementia, the T-ACE-III score was the only significant variable, and demographic variables (years of education, age, sex, and socioeconomic status) were no longer important.

9.1.3 Summary of the findings for the supplementary tests

Among the four supplementary tests I used (discussed in *Chapter 7*), fewer than half of individuals with mild (CDR = 1) and moderate (CDR = 2) dementia who presented to a memory clinic completed the Brixton Spatial Anticipation test (46.7%, 16.7%). Even fewer completed the Digit Symbol Coding (CDR = 1, 26.7%) and Symbol Search (CDR = 2, 16.7%) subtests of the WAIS-IV, despite most participants having very mild dementia (CDR = 0.5, 77.8% - 83.3%) and those from non-dementia group (96.4%-100.0%) being able to complete the tests. The Coin-In-the-Hand (CIH) test was found to be acceptable to participants from both the dementia and non-dementia groups. Only one participant, who had dementia with a score of 1 on the CDR, did not complete this test.

Possible reasons for a low completion rate were that people who had previously completed the T-ACE-III may have felt tired and may have been unwilling to continue with the supplemental tests, due to exhaustion. It was particularly difficult for dementia patients with CDR scores of 1 or 2, to comprehend the instructions (understanding what to perform in the Brixton test) or to complete the practice trials (Symbol Search and Digit Symbol Coding tests of WAIS-Processing Speed Index) of these tasks.

9.2 Comparison with other research

In this section, I reflect on, compare, and discuss my findings with those of prior research in an effort to shed light on some critical issues related to meta-analyses of diagnostic test accuracy, cross-cultural validation studies, and the use of diagnostic tests.

9.2.1 Systematic review and meta-analysis of dementia tests used with Chinese-speaking populations

There has been only one other systematic review and meta-analysis of cognitive assessment tools (those taking < 35 minutes) used with Chinese-speaking populations (see *Table 18*, on page 141). This study had wider inclusion criteria than mine and identified 81 dementia tools used in all settings (e.g., clinics, community, primary care, specialist care, nursing homes), and sampled 167 studies (Huo et al., 2021). My systematic review and meta-analysis study specifically examined brief (< 20 minutes) cognitive assessment tools with a focus on clinical settings where people had presented for assessment of suspected memory impairment or dementia. These were memory or neurology clinics that considered both dementia and MCI in Chinese-speaking populations. Therefore, only 37 papers were included in the systematic review and 21 papers in the meta-analysis, and 27 brief cognitive instruments were identified for the purposes of the systematic review.

In my systematic review, the test with the best diagnostic performance (sensitivity and specificity > 75%) and acceptable heterogeneity (I^2 < 75%) was the ACE (ACE-R & ACE-III), while other tests (e.g., MMSE, MoCA, & MoCA-BC) had high heterogeneity in the detection of MCI, or had a lower sensitivity or specificity in the detection of dementia (e.g., CDT, STT-A & B) and MCI (e.g., MMSE). Huo et al., (2021) also reported a great deal of variation in both sensitivity and specificity among all screening tools except for ACE-R. This was similar to the findings for MCI in my study. The ACE (ACE-R & ACE-III) was the only test with acceptable

heterogeneity ($I^2 < 75\%$). They also reported that the diagnostic performance of ACE-R, MoCA, HDS, CASI, and GPCOG were superior to that of MMSE. Similarly, my research also indicated that MoCA, MoCA-BC, and ACE (ACE-R & ACE-III) outperformed the MMSE in detecting MCI, while MoCA outperformed the MMSE in detecting dementia. Both papers found that ACE-R performance was less heterogeneous than other tests.

Huo et al. reported that optimal cut-off scores for the MoCA and MMSE varied between studies. In my study, the cut-off scores for distinguishing dementia and MCI from non-dementia group also differed across Chinese-speaking populations, clinical settings, educational levels, and subtypes. For instance, the MoCA cut-off scores for differentiating dementia from clinical controls was 18/19 in Hong Kong and 20/21 in Taiwan. Similarly, the MMSE cut-off score was 27/28 when used in China, and 26/27 when used in Hong Kong for differentiating MCI from clinical controls.

9.2.2 Comparison approach

Several factors can contribute to heterogeneity in results for dementia diagnostic and screening tools, including the language of the publication, selection of controls, and clinical settings, particularly in a Chinese context (Huo et al., 2021). Therefore, I separately compared the diagnostic accuracy of each test based on outcome of interest (dementia or MCI recruited in the clinical settings) and this provides strong evidence for their future use. Furthermore, in the meta-regression, I found that the population (China vs others) and MCI subtype (MCI vs others) influenced MMSE results, the dementia scoring system (original vs adjusted) influenced MoCA results.

In summary, only two systematic review and meta-analysis studies (including mine) have been conducted for Chinese-speaking populations on the topic of brief cognitive assessment tools for dementia and MCI. The other Chinese study (Huo et al., 2021) focused on screening tools and covered a broad range of settings to provide an overview of the diagnostic performance of each test;

however, there should be caution when generalising the results of their meta-analyses due to their high heterogeneity. My own work examined tests for people presenting to secondary care with suspected memory impairment or dementia which provides a direct comparison. No previous meta-analysis on a similar topic adopted this approach. This will also be more relevant to clinicians wishing to use these tools in a clinical setting.

9.2.3 MMSE

The MMSE is globally the most widely used brief cognitive assessment tool, and it was included by most studies in recent systematic reviews (Custodio et al., 2020; Huo et al., 2021; Rosli et al., 2016; Tsoi et al., 2015; Velayudhan et al., 2014). From the time that the MMSE's intellectual property rights were transferred to Psychological Assessment Resources in 2001, its accessibility and use have diminished (Folstein et al., 1975; Powsner & Powsner, 2005).

A systematic review and meta-analysis (Tsoi et al., 2015) examined cognitive tests that are shorter than 20 minutes and used with individuals in community settings, memory clinics, or hospitals with the aim of replacing the MMSE with freely accessible tests that have equivalent diagnostic power. They found that the Mini-Cog and ACE-R performed better than the Clock Drawing Test (CDT), Memory Impairment Screen (MIS), verbal fluency (VF test), Abbreviated Mental Test (AMT), General Practitioner Assessment of Cognition (GPCOG), MoCA, Informant Questionnaire on Cognitive Decline in Elderly (IQCODE), and modified Mini-Mental State Examination (3MS). Nonetheless, for detecting dementia and MCI, there was a high degree of heterogeneity (I^2 value) found in different papers evaluating the Mini-Cog (89 %, 97 %), ACE-R (53 %, 87 %), MoCA (31 %, 93 %), and MMSE (92 %, 94) (Tsoi et al., 2015).

Despite the fact that several tests, like Mini-Cog (e.g., sensitivity = 91%, specificity = 86%) and ACE-R (e.g., sensitivity = 92%, specificity = 89%), are freely accessible and diagnostic accuracy are better than the MMSE (e.g., sensitivity = 81%, specificity = 89%) (Tsoi et al., 2015), the MMSE is still the most

extensively used proprietary test for dementia (Custodio et al., 2020; Huo et al., 2021; Rosli et al., 2016; Tsoi et al., 2015; Velayudhan et al., 2014); therefore, most clinicians may be familiar with it and still wish to continue using it.

In addition, the ACE-R is more time-consuming compared to other brief cognitive assessments, despite its significant diagnostic power (Custodio et al., 2020; Velayudhan et al., 2014) (see *Table 18*, page 141). In 2014 and 2020, systematic reviews of brief cognitive evaluations were completed for English and Spanish-speaking populations, respectively (Custodio et al., 2020; Velayudhan et al., 2014). Both studies also included patients from mixed settings, such as primary and secondary care, day care centres, and communities, and both concluded that the ACE-R may be useful in differentiating between various types of dementia and should be administered when the diagnosis is doubtful. It can be administered and interpreted by multiple professionals (e.g., nurses, psychologists, and physicians) with appropriate training, thus reducing need for specialist neuropsychological assessments (Velayudhan et al., 2014).

Four tests from these reviews that had promising performance were validated in secondary care with English-speaking populations but not with any Chinese-speaking populations: *Cognitive Assessment Screening Test (CAST), Rotterdam-The Cambridge Cognition Examination (R-Camcog), Memory Orientation Screening Test (MOST), Test for the early detection of dementia (TED4-Cog)*. Four tests were validated in memory clinics with Spanish-speaking populations but not with any Chinese-speaking populations: *INECO Frontal Screening (IFS), Memory, Fluency and Orientation (MEFO), Memory Binding Test (MBT), Phototest.* These tests could in future be considered for adaptation and validation in Chinese contexts and thus may expand the clinical repertoire of available cognitive screening tools.

Table 18: Systematic reviews of dementia diagnostic and screening tools

Author and years:	Velayudhan et al., (2014)	Tsoi et al., (2015)	Rosli et al., (2016)	Custodio et al., (2020)	Franzen et al., (2020)	Huo et al., (2021)	My thesis
Language:	English	Published in English	Languages spoken in the Asia Pacific region	Spanish	Languages spoken in non-western populations	Chinese	Chinese
Searched date:	N/A	The earliest available date to September 1, 2014	Articles published between September 1989 and June 2014	Articles published between 1953 and July 30, 2018	The earliest available date to August 2018	The earliest available date to 1 March 2020	The earliest available date to 12 October 2021
Search database:	Medline, Embase, PsychInfo, Web of Science, HMIC Health Management Information Consortium, and the Cochrane library. Hand-searched the International Journal of Geriatric Psychiatry, Ageing and Mental Health, International Psychogeriatrics, and Age and Aging.	Medline, Embase, PsychInfo	CINAHL, MEDLINE, Embase, Cochrane Library, and Science Direct	Medline, Biomed Central, Embase, Scopus, Scirus, PsycINFO, LILACS, and SciELO.	Medline, Embase, Web of Science, Cochrane, Psycinfo, and Google Schola	PubMed, EMBASE, Science Direct, SCOPUS, Web of Science, Cochrane Library Database, PsychoINFO, CINAHL and three Chinese databases (Chinese Biomedical Literature Database, CNKI, and WANGFANG); ISI Proceedings, BIOSIS database, OpenSIGLE website.	MEDLINE, PsychInfo, PsycTESTS, the Web of Science core collection, and The Cochrane library.

Author and years:	Velayudhan et al., (2014)	Tsoi et al., (2015)	Rosli et al., (2016)	Custodio et al., (2020)	Franzen et al., (2020)	Huo et al., (2021)	My thesis
Country of target population(s):	USA and Europe	No restriction	South Korea, China, Hong Kong, Japan, Taiwan, Malaysia, Philippines, Singapore, India, Sri Lanka	Argentina, Chile, Colombia, Peru, Mexico	Brazil, China, Cuban, Hong Kong, Korea, Singapore, Taiwan, or concerned Hispanics/Latinos residing in the USA	China, Hong Kong, Singapore, Taiwan	China, Hong Kong, Singapore, Taiwan
Types of participants:	Dementia	Dementia (MCI as secondary outcome)	Dementia, MCI	Dementia, MCI	Low educated dementia, and MCI	Dementia (cognitive impairment– MCI and dementia were analysed separately)	Dementia, MCI
Settings:	Memory clinics, primary care, community, and hospital in-patient unit.	Community, clinic, hospitals, primary care, others	Did not specify. Tools intended for use by non-specialist health care workers or intended to aid detailed assessment by a physician or designed for the identification of dementia subtype.	Memory clinics, primary care, adult day care	Outpatient, population-based, other, institutionalised	Include but are not restricted to community, primary care, specialist care, hospital wards, residential homes, and nursing homes	Patients with dementia or MCI: clinical settings non-dementia group: either clinical or community-based settings
Administration time (category):	≤ 20 mins	< 5 mins, 5-10 mis, 10-20 mins	Did not specify, but 45 as the longest.	< 15 mins	Did not specify	≤ 35 mins	≤ 20 mins

Author and years:	Velayudhan et al., (2014)	Tsoi et al., (2015)	Rosli et al., (2016)	Custodio et al., (2020)	Franzen et al., (2020)	Huo et al., (2021)	My thesis
Quality assessment or criteria:	The Centre for Evidence-based Medicine (CEBM) guideline and the level of evidence	Standards for Reporting of Diagnostic Accuracy statement. Potential risk of bias: Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2)	QUADAS-2	Content validity, internal consistency, criterion validity, construct validity, reproducibility, diagnostic accuracy, floor and ceiling effects, interpretability	Self-develop eight criteria: illiterate participants, administration language, cross-cultural adaptation, feasibility, validity, blinding, source, and criteria	QUADAS-2	The CEBM diagnostics criteria
Other quality assessments:	N/A	N/A	N/A	MTRQ and MCAR	N/A	N/A	MTRQ and MCAR
Meta-analysis Methods:	N/A	Bivariate random-effects model (forest plots, HSROC plots, subgroup analysis)	N/A	N/A	N/A	Bivariate random-effects model (forest plots, HSROC plots, subgroup analysis)	Univariate random-effects model (forest plots), Bivariate random-effects model (HSROC plots, SROC plot)

Note. MCAR = Manchester Translation Reporting Questionnaire; MCI = Mild Cognitive Impairment; MCRQ = Manchester Cultural Adaptation Reporting Questionnaire; SROC = Summary Receiver Operating Characteristic; HSROC = Hierarchical Summary Receiver Operating Characteristic.

9.2.4 Reported process of translation and cultural adaptation

Only four out of 37 studies fully reported the procedures used for translation and cultural adaptation (see *6.3.2 The quality of each study and the reported procedures for translation and cultural adaptation*). Among them, six studies scored zero on the MTRQ and MCAR scales. Nearly half of the studies reported completing a translation, one third of them reported back translation. Other procedures were rarely reported such as: users in co-production, expert recommendation, revisions based on step-by-step feedback, involvement of original authors and a pilot study (see *6.2.2 Translation and cultural adaptation procedures*). Even though some studies indicated that they used tests with minimal language requirements (Lin et al., 2003), or language-neutral tests (Low et al., 2020), and were given four points (the highest score) in the translation and cultural adaptation processes in my study, it is still unlikely that any tests are completely culture-free (Fernandez & Abe, 2017). For example, the claim that the Trail Making Test (TMT) is a culture-free test has been questioned by Lee et al., (2000), since the correlation of the TMT with Colour Trail Test (CTT) is lower in a Chinese-English bilingual population than in an English monolingual population, after matching the participants' age and educational level in the two groups. This divergence may be a result of mastering a language rich in the feature of temporal sequencing, such as English, which confers an advantage when performing the TMT.

9.2.4.1 Cultural Bias

In the field of neuropsychology, it is proposed that fundamental cognitive processes and characteristics are shared across cultures, but they can be developed and exhibited differently in different cultures (Fernandez & Abe, 2017; Henrich et al., 2010; Uzzell et al., 2007). Cultural differences are generally rooted in the contexts in which these cognitive processes are employed, rather than by the existence of particular cognitive processes in one culture and their absence in another culture (Cole, 1975). Different cultural settings result in the formation

of distinct cognitive profiles because culture determines what should be learnt, at what age, and by what sex (Berry, 1979). Cognitive abilities are generally measured by neuropsychological tests, and the existence of cultural differences has been confirmed in various cognitive domains, including visual perception, memory, and language (Uzzell et al., 2007).

There are two ways in which we can minimise the influence of cultural difference and increase the validity of neuropsychological tests (Fernandez & Abe, 2017):

- **Cultural decentring:** Modifying existing instruments but keeping similar meanings and constructs in a different culture. For example, when developing an instrument for use in several cultures, exclude words or concepts that are clearly specific to one language or culture (Tanzer, 2005).
- **Cultural centring:** Developing tests that have culture-specific meaning and validity regarding a particular cognitive function. For example, the Brain Health Test was developed for specific use within Taiwanese culture as a brief dementia screening instrument with a focus on the "memory" cognitive domain and it also incorporated the indicators of the three ‘highs’ (high blood pressure, high cholesterol, and high blood sugars) as well as depression, and complaints of memory declines (Tsai et al., 2018).

When scores for a given construct appear to be discrepant from the underlying ability being tested, there are good grounds for suspecting the presence of bias (van de Vijver & Tanzer, 2004). Several sources of bias in cross-cultural testing have been identified (Fernandez & Abe, 2017):

- **Construct bias**: When the measurement of the construct does not apply equally to all cultural groups (van de Vijver & Tanzer, 2004). For example, various cultures may evaluate attention tasks differently, depending on which factors are seen as most essential.
- **Methods bias**: When bias occurs related to methodological issues, including sample bias (e.g., differences in size, age, or selection criteria), instrument bias (e.g., differential familiarity with stimulus materials), and administration bias (e.g., language issues, or communication problems).

- **Item bias**: When the same item has different meanings in different cultures (Holland & Wainer, 1993), e.g., the Famous Face Recognition and Naming test (Rizzo et al., 2002) which was developed in Italy contains items which are more familiar to Italians than to non-Italians.
- **Language bias**: The language in which a test is administered can lead to variability when two different language groups are compared. However, even when two groups are matched on key variables (e.g., education, age), one group can nonetheless outperform the other on a culturally fair test, which truly reflects ability differences (van de Vijver & Tanzer, 2004). For example, Chinese speakers are better than Japanese speakers in a mental arithmetic task (Lau & Hoosain, 1999).

9.2.4.2 Bias in dementia instruments

Some well-established neuropsychological tests such as the MMSE and MoCA, that have been used globally are nonetheless prone to bias not only due to factors such as language and culture, but also due to factors such as education, literacy, and age (Ng et al., 2018). When a nationally representative population sample in Singapore (N = 1,092), comprising Chinese, Malays, and Indians were screened for dementia using the MMSE, differences in performance could be readily attributed to linguistic and cultural factors (Ng et al., 2007). For example, difficulties related to the language of administration were observed, especially for serial 7s, repetition of “no ifs, ands, or buts”, request to close your eyes, and sentence writing (Teresi et al., 1995). Significant differences in MMSE scores existed only among less-educated Asians from these three countries and could not be fully explained by cognitive functioning correlates (socioeconomic status, comorbid illnesses, functional health status, and health-related behaviours). When using the MMSE with less-educated Asians, such ethnic non-equivalence must be considered (Ng et al., 2007).

A systematic review of neuropsychological domain-specific tests used in patients diagnosed with dementia in non-western, low-educated, or illiterate populations

was published in 2020 (Franzen et al., 2020). It comprised research conducted in Brazil, China, Cuba, Hong Kong, Korea, Singapore, and Taiwan, and also with Hispanics and Latinos living in the United States. They found that only a few studies applied a rigorous adaptation process, which highlighted the difficulty of developing a test that measures a cognitive construct in the same way as the original test, allowing for differences in language and culture. A lack of research assessing cognitive domains of executive functions, complex perceptual and motor skills, and language in non-western, less-educated patients limits the value of neuropsychological tests used to assess for dementia in such populations.

9.2.4.3 Existed guidelines for translation cultural adaptation

I applied the MTRQ and MCAR scales (Mirza et al., 2017; Waheed et al., 2020) in my narrative synthesis (*Chapter 6*) to each included publication to see whether the reported information on "translation" and "cultural adaptation" was sufficient to allow for future replication of the study in question, and I found that only four studies met the highest criteria. In a 2017 systematic review (Mirza et al., 2017), the ACE, ACE-R, and ACE-III tests were examined to determine the appropriateness of the reporting of each item individually. The results were similar to my findings, with very few papers meeting the highest standard on both scales.

In my validation study of the T-ACE-III, I followed the guidelines of the Cross-Cultural Adaptation Process (see *7.3.2 Cultural Adaptation)* (Beaton et al., 2000) recommended by the American Academy of Orthopaedic Surgeons (AAOS). These guidelines are intended to optimise semantic, idiomatic, experiential, and conceptual similarity between the original and the adapted/translated measure. These guidelines provide translation and cultural adaptation instructions based on the requirements of various scenarios. For instance, I followed Beaton’s guideline to adapt main instrument - the T-ACE-III for my thesis. The T-ACE-III did not require a wholesale translation because a Chinese version already existed; I only needed to modify a few items or terminology to fit the Taiwanese context based on their framework. In contrast, other guidelines, such as the Sousa and

Rojjanasrirat (2011) user-friendly guidelines, are particularly suited for researchers that specifically aim to include "translation processes".

9.2.5 Methodological challenges in reviewing diagnostic test accuracy

Systematic reviews of reliability and accuracy of cognitive assessment tools are important, as they help inform the selection and interpretation of appropriate tests. I conducted the systematic review of the diagnostic test accuracy, a topic which differs from other research topics such as intervention studies, in terms of its unique considerations in literature search, quality assessment, and statistical methods. (Takwoingi & Deeks, 2022). In the following section, I will explain my methodology by elaborating on these distinctions.

9.2.5.1 Literature search

Choosing search terms on diagnostic test accuracy faces challenges due to a lack of consistent study design terminology and appropriate indexing terms. The suggested key indexing items for searching papers in this topic involve “index test”, “target condition”, and “target population”, in order to minimise the chance of missing relevant papers (Takwoingi & Deeks, 2022). As I attempted to investigate all the brief cognitive assessments used in Chinese-speaking populations, I did not name specific tests. The keyword combinations in my study included terms related to "test," "dementia and MCI," and "Chinese (including commonly used dialects)"; and it yielded a total of 19,722 papers. The majority of articles were eliminated after removing duplicate papers (n = 10072) and in the initial screening due to the title and abstract covering irrelevant fields (n = 9236). Those articles eliminated after full-text screening were those that did not recruit participants in a memory clinic or similar setting. To avoid screening an excessive number of irrelevant articles, methodological filters consisting of text words and database indexing terms were created to improve accuracy. However, one should avoid using such filters as the sole strategy in a formal literature search

because these search filters do not work consistently (Beynon et al., 2013), and this advice is in line with the recommendations of the Cochrane Library (de Vet et al., 2008). In addition, prospective researchers are advised to consult a librarian or information specialist with prior experience in researching diagnostic test accuracy (Takwoingi & Deeks, 2022). Based on my own experiences, I would suggest including only terms for "target condition" and "target population" to obtain a wide variety of studies and quickly eliminate irrelevant papers in the title and abstract screening (e.g., Chinese populations, validation purposes in my case) so as to avoid missing relevant articles. Only a few hundreds of publications will require full-text screening as a result.

9.2.5.2 Assessment of quality

The criteria for the assessment of study quality are important since some details can be easily neglected when assessing diagnostic test studies, resulting in biased findings. Firstly, cognitive tests should be evaluated separately when the study includes more than one test (Takwoingi & Deeks, 2022). Secondly, due to lack of detailed reporting, the rating for the methodology's quality frequently slipped into the "Unclear" category. For example, one study using the Quality Assessment of Diagnostic Accuracy Studies-Two (QUADAS-2) rated as “unclear” 26 out of 49 (53%) studies for the reference (gold) standard of diagnosis (Liu et al., 2015). In my study, I used the CEBM diagnostic criteria to assess the quality of the study and eliminated the outcome option "Unclear" (1 point) for a binary result of "Yes" (2 points) or "No" (0 points) to avoid ambiguity, so that the information not reported was scored 0. Two studies that validated two different tests (CDT & T&C in Chan et al., (2005)’s study, MoCA & MMSE in (Tsai et al., 2016)’s study) were evaluated and given two scores for quality evaluation, separately. Thus, the outcome of the assessment of each study's quality can be stated clearly. No study from my systematic review, met the highest standard of CEBM (see *4.3.2 Study characteristics and quality analysis*): (1) 36 studies excluded participants with specific conditions so were not representative of the population. (2) Two studies did not apply reference standard regardless of the

index test results. (3) Two studies were not blinded to other cognitive test scores and three did not specify. (4) Two studies provided no information for the reference standard of diagnosis. (5) Three studies did not specify whether they were blinded or not. (6) 14 studies provided insufficient details on the assessments. Therefore, the studies listed in (4), (5), (6) provided insufficient detail for replication (16 studies).

9.2.5.3 Statistical methods

Statistical methods reported in current reviews are inconsistent as the methods for evaluating statistical heterogeneity are not standardized (Plana et al., 2021). One examination of published meta-analyses studies of diagnostic test accuracy (Stegeman & Leeflang, 2020) found the use of the data available in the meta-analysis to be poorly replicable due to lack of information about data used, and mistakes in data extraction and reporting; only 1 article out of 51 (2%) was completely reproducible. Notably, the authors considered reproducibility to be limited to the capacity to reproduce the data using the information accessible in the meta-analysis. They did not investigate whether the step from review question to search and selection process, or the search and selection process itself, could be repeated. Therefore, no conclusions can be drawn regarding the overall reproducibility of the review (Stegeman & Leeflang, 2020).

To account for the potential (negative) correlation between sensitivity and specificity across studies, some authors (Reitsma et al., 2005) recommended hierarchical bivariate analysis instead of univariate methods. There is no consensus regarding these methods (Harbord et al., 2008; Simel & Bossuyt, 2009). However, I found generally similar summary estimates and marginal confidence intervals for sensitivity and specificity in univariate and bivariate meta-analysis models (see *4.3.3.4 Sensitivity analysis of univariate and bivariate analysis*). This was consistent with earlier findings, which suggested that the bivariate models do not generate substantially different summary estimates for

sensitivity, specificity, and heterogeneity (Dahabreh et al., 2017; Simel & Bossuyt, 2009).

In addition, one systematic review study that investigated the methods used to quantify heterogeneity in Cochrane diagnostic test accuracy reviews concluded that reporting on heterogeneity is poor and authors' subjective judgements about heterogeneity are often inconsistent with the statistical estimated results (Plana et al., 2021). In the methods section of my study, I used the criteria of I^2 value suggested by the Cochrane Library (Higgins et al., 2021) and reported the heterogeneity based on I^2 values calculated from both univariate and bivariate models, with HSROC and SROC plots demonstrating the overall performance of each test, so as to prevent personal misjudgements.

There are other factors, such as differences in design, conduct, recruitment, reference standard, or chosen threshold (cut-off scores) for disease that may lead to between-study variance and therefore heterogeneity in test accuracy. I performed a meta-regression in an effort to determine the factors (e.g., subtypes, populations, and reference standard) that contributed to the high heterogeneity among certain tests (e.g., MMSE, MoCA). However, the sample size of my meta-analysis did not permit me to fully evaluate other potential factors (e.g., cut-off scores, education levels), which can be investigated further in future research.

9.2.5.4 Limitations of reviewing diagnostic test accuracy

There are certain limits to a meta-analysis carried out to examine diagnostic test accuracy. Firstly, it does not assess the test's impact on patients (Ferrante Di Ruffano et al., 2012). Secondly, undertaking a systematic review or meta-analysis on this topic may produce less useful evidence than investigations of patient outcome effect or using a novel diagnostic test on therapeutic and diagnostic decisions. However, investigations on the topic can benefit and inform national or worldwide diagnostic criteria in some research disciplines (Takwoingi & Deeks, 2022). Due to the lack of standardised guidelines for performing a meta-analysis on this topic, I made extra efforts to explore various platforms and

software in order to find the most appropriate approaches and found that demonstrating the findings via numerous methods is more reliable and less biased compared to using a single method. The customised programmes used to generate the forest plots for the study now is available at: https://github.com/GraceUCL/Systematic-review-project. Future research should investigate the statistical methods used in the meta-analyses of diagnostic test accuracy, given that the use of diagnostic tests is dependent on the findings of systematic reviews and meta-analyses on this topic (Takwoingi & Deeks, 2022).

9.2.6 The Chinese versions of ACE-III

As of October 2022, 37 versions of the ACE-III have been developed for various populations. A narrative review of all Chinese versions of the ACE (ACE, ACE-R, ACE-III, M-ACE) was published in 2022 (Cao et al., 2022) and reported that the initial Chinese translation of ACE-III was published in 2017 (Wang et al., 2017), and that other versions were published in 2019 (Li et al., 2019; Wang et al., 2019). In a Chinese context, the Chinese version of the ACE-III was effective not only as a cognitive instrument for dementia diagnosis, but also at detecting MCI (Li et al., 2019; Pan et al., 2021; Wang et al., 2019).

In Taiwan, when I developed the T-ACE-III, the optimal cut-off score for identifying dementia was 73/74, with a sensitivity of 89.5% and a specificity of 100%. In China, the cut-off score was 82/83, with a sensitivity of 91.1% and a specificity of 83.1%. The lower cut-off score observed in my study, compared to the study conducted in China (Wang et al., 2017), could be attributed to differences in the demographic background of the participants. Specifically, the dementia group in their study had a younger age (71.9 vs 74.2 years old) and a higher level of education (11.3 vs 7.7 years) compared to my study. Moreover, they recruited participants in the non-dementia groups from a physical examination centre while I recruited from a memory clinic. In my study, the non-dementia participants who sought evaluation at the memory clinics due to memory complaints (or subjective cognitive decline, SCD) may have had a higher risk of dementia compared to individuals with SCD identified through community-

based settings (Slot et al., 2019) (please refer to section 7.5.1 Summary of findings). These factors may have influenced the determination of the cut-off score.

In my study with the same cut-off point of 82/83, the sensitivity was 96.5% and the specificity was 87.9% (see *Table 12: Cut-off scores for the T-ACE-III in differentiating dementia from non-dementia, sensitivity, specificity and Youden index*). A similar performance (Youden Index = 74.2 in the China study) can be observed with the cut-off point of 85/86 in my study. In addition, the study that validated a Taiwanese version of ACE-III for detecting MCI found that the optimal cut-off score was 76.5, with a sensitivity of 84.0% and a specificity of 85.0%. In Hong Kong and Singapore, no validation studies on ACE-III have been carried out.

The narrative review study by Cao et al. also summarised the pros and cons of using ACE-III, MMSE, and MoCA to diagnose dementia and MCI. The only disadvantage of ACE-III is that it is time-consuming (15 - 20 minutes), whereas it has multiple benefits, including high sensitivity and specificity, excellent performance in individuals with high, middle, and low educational levels (ACU = 0.95, 0.91, & 0.89), and no copyright restrictions (Cao et al., 2022).

There is no validation study of the ACE-III conducted on some rarer forms of dementia in Chinese-speaking populations, although behavioural variant frontotemporal dementia and primary progressive aphasia having been assessed using the ACE-III in English speakers (Hsieh et al., 2013).

9.2.7 The use of diagnostic tests

My PhD focused on the use of brief cognitive assessments in cases of suspected dementia (*see 4.1.2 Differences between screening tests and diagnostic tests*). Therefore, these tests would be used along with medical history and laboratory data (blood test, neuroimaging, mental and physical examination) to make a diagnosis. This is not in the context used for screening populations who have not

presented with concerns to health professionals. The population they are used in makes a big difference: diagnostic tests are typically administered to people who are exhibiting symptoms (high prevalence, for example, dementia patients in memory clinics or neurology departments), whereas screening tests are frequently given to people who are not exhibiting symptoms (low prevalence e.g., dementia patients in communities or day care centres). When utilised with the populations and settings where the tests were developed and validated, they can maintain a high level of reliability and validity. Furthermore, using diagnostic tests as screening tools might be more harmful than helpful (Wang et al., 2003). The following characteristics are necessary for an effective diagnostic test (Reid et al., 1995; Thomas et al., 2015):

- Under strictly controlled circumstances, the test detects cases with evident pathology with high accuracy and reliability.
- The test distinguishes between healthy controls and patients with a broader spectrum of illness, including cases with mild or less typical symptoms.
- The test distinguishes between patients with a specific, well-defined range of illness and healthy controls with high accuracy and reliability.
- The test distinguishes between more diverse groups of controls and patients with high accuracy and reliability, as comorbidities in one or more groups might be confounded with the diagnosis and affect test confidence.
- The test distinguishes the illness amongst patients who would be seen in a typical clinical practice with high accuracy and reliability.

In my study, the T-ACE-III was employed to distinguish patients with dementia with a variety of causes, different stages, and a range of presentations from the people without dementia presented to a typical memory clinical practice.

9.3 Strengths and limitations of my PhD thesis

My thesis considered the range of cognitive tests that Chinese-speaking clinicians can use to help in diagnosis of dementia. I conducted a systematic

review and meta-analysis to find out the optimal brief cognitive assessments that had been used in Chinese-speaking populations, and the extent to which studies reported on the processes of translation and cultural adaptation. This systematic review also helped to inform the validation project that formed the basis of my PhD. Following specific guidelines, I adapted a cognitive assessment tool for use in Taiwanese clinical settings and conducted an empirical study that involved interview and testing of people with dementia at a hospital in the northern part of Taiwan. I used the adapted T-ACE-III as well as supplementary tests. This study was the first study to validate the T-ACE-III in a population with dementia. I have made the T-ACE-III available for use by clinicians and researchers (FRONTIER Research Group: www.ftdrg.org; Brain and Mind Centre: www.sydney.edu.au/brain-mind/resources-for-clinicians/dementia-test.html).

Chapters 7 and 8 of my thesis used a cross-sectional approach, which is a type of observational study design and involves examining data from an available population of potential relevance to the research question at a specific point in time, without prospective or retrospective follow-up. I adopted nonprobability sampling methods, convenience sampling, where participants were selected based on their attendance in a clinic, their availability, and their willingness to take part in my study. There are many strengths using this method, such as low cost (quick and inexpensive), the collection of data on all variables at a single time point, the examination of multiple outcomes and exposures, the ease with which hypotheses can be formulated, and the suitability for conducting in-depth research (Wang & Cheng, 2020).

In the case of limitations of my thesis, a larger and more representative sample size might have been achieved if I had more time and resources, particularly in terms of recruiting participants from South Taiwan, where Hokkien is the most commonly spoken dialect compared to the North (48.5% vs 18.2%) where I conducted my data collection (Directorate-General of Budget, Accounting and Statistics Executive Yuan, 2021). The influence of demographic factors (language used, education level) on the T-ACE-III scores, which I did consider, could have been examined in greater detail.

9.4 Clinical implications

The findings in my thesis have clinical implications, mainly for clinicians working with different Chinese-speaking populations and using brief cognitive assessments in clinical settings for the detection of dementia, through cultural adaptation of an appropriate tool and providing information on factors that may influence test performance. In addition, I provide suggestions for researchers who intend to validate cognitive assessments in a Taiwanese context.

9.4.1 Is using T-ACE-III alone good enough to measure cognition in patients with suspected dementia?

A key question that I aimed to answer in this thesis was whether the T-ACE-III alone as the cognitive examination part of a clinical assessment of suspected dementia is sufficient to detect dementia in clinical settings. My empirical study of the T-ACE-III pointed to a high level of validity (sensitivity = 89.5%, specificity = 100% at a cut-off point of 73/74), while using a cut-off point of 86/87, it's sensitivity and specificity were 100% and 78.8%. Therefore, it can be a helpful tool in the detection of dementia where this is suspected. In addition, I found that the language used (Hokkien or both Hokkien and Chinese) as well as years of education had an impact on T-ACE-III scores, and this is partly related to the characteristics of the participants who were recruited to my study (all participants who only spoke Hokkien were dementia patients), and partly related to an educational gap caused by the nine-year compulsory education policy introduced in Taiwan in 1968.

It is recommended that the T-ACE-III be used in practice as part of a clinical assessment for the routine cognitive assessment tool. It may be also applied as part of a test battery, when assessing some atypical forms of dementia; however, I lacked enough participants with atypical dementia to evaluate it in my study. When administering the T-ACE-III to individuals who speak different dialects, the examiner should be able to switch to the preferred dialect of the individuals in question, as these individuals may find the official Chinese language to be less familiar, so hindering their test performance. And despite being classified as a

brief cognitive test, the T-ACE-III takes longer than some other brief cognitive tests, such as the MMSE. If necessary, the T-ACE-III can be administered on a different day to that where there are other demanding clinical routines or examinations.

Lastly, timely diagnosis is one of the main aims of Taiwan's dementia policy; an effective brief cognitive assessment, together with additional laboratory tests and a brain scan, can assist in diagnosing dementia patients immediately so that they can receive the appropriate treatment.

9.4.2 What are the factors that may influence diagnostic value of brief cognitive assessment tools across Chinese-speaking populations?

My systematic review and meta-analysis study is a ground-breaking investigation into the brief cognitive instruments used as a diagnostic tool in Chinese-speaking patients with dementia and MCI. Notably, whereas Huo et al. (2021) focused solely on instruments used for screening, our study investigates the diagnostic use of these instruments. My findings revealed that the cut-off scores for brief cognitive tests varied across Chinese-speaking populations, subtypes of dementia and MCI, recruiting settings, age, and level of education. Therefore, it is essential to choose cut-off scores based on these demographic parameters. In addition, I applied meta-regression to determine the possibly impact in performance of factors such as population (China vs. others), subtype of dementia (AD vs. non-AD) and MCI (MCI and others), reference standard of dementia (DSM vs. others) and MCI (Peterson vs. others) and scoring system of MoCA (original vs. revised). My findings showed that the specificity value of MMSE in populations other than China (88%), was significantly higher than in China (70%). The sensitivity value of MMSE was significantly higher in MCI (80%) than other subtypes (60%). And, MoCA's original scoring system exhibited significantly higher sensitivity and specificity (96%, 98%) than the revised scoring system (88%, 89%) in the detection of dementia.

9.5 Policy implications

Based on the findings of my thesis, the T-ACE-III is a sensitive and specific, effective, and highly acceptable (90 out of 91 people completed it) tool in the detection of dementia. However, until now MMSE, CASI, and CDR have been the most frequently used dementia diagnostic routine tests in Taiwan. According to "Drugs Acting on the Nervous System", published by the National Health Insurance Administration, Ministry of Health and Welfare, the medication for treating dementia can be prescribed to a new patient with the following medical histories: Either a computerised tomography (CT) or Magnetic resonance imaging (MRI) scan, or Hachinski lschemic Score report, a complete blood count (CBC), Venereal Disease Research Laboratory (VDRL), Blood Urea Nitrogen **(**BUN), Creatinine, aspartate aminotransferase (GOT), alanine aminotransferase (GPT), thyroxine (T4), thyroid stimulating hormone (TSH) examination, as well as the MMSE, or a CDR report. MMSE and CDR scores also influence the staging and treatment plan for dementia. Therefore, it will be difficult to introduce a new diagnostic routine test or replace previous tests because it has to conform not only to hospital system improvements and staff training, but also to policy. However, timely diagnosis is one of the main aims of Taiwan's Dementia Policy and Action Plan 2.0, which also stipulates that international collaboration should be conducted on the matter of evaluating the validity and reliability of dementia assessment tools in order to provide an increased evidence-base for policy making.

9.6 Future research

9.6.1 Chinese dementia cognitive assessments

The systematic review and meta-analysis study of the overall diagnostic performance of cognitive assessments in Chinese-speaking populations was only published in 2021 (Huo et al., 2021). The study did not directly compare the

performance of every test in different settings, but instead used meta-regression to investigate the subgroups used for each individual test. I performed a meta-analysis of brief cognitive tests using subgroup analysis based on disease classification (dementia and MCI), so as to provide a direct and purposeful comparison. Multiple settings and subtypes should be included in future research in order to guide test selection across a range of settings.

In the case of the substantial heterogeneity of specific tests (CDT, MoCA) that was apparent in the meta-analyses, the risk of overinterpretation is greatest when there are numerous differences across studies, but too few studies are available to look at heterogeneity; therefore, the sources of heterogeneity should be evaluated in future research (Higgins & Li, 2022). In addition, the majority of dementia diagnostic and screening tools were developed in the western world, despite the fact that some cognitive assessments, such as ACE-R and ACE-III, performed exceptionally well after being adapted to Chinese-speaking populations, whereas locally designed assessment tools are rather uncommon in Chinese-speaking population s. Therefore, a systematic review should be conducted on this topic to determine whether locally designed tests are more accurate, relevant, and possess better diagnostic performance than those developed in a foreign culture.

9.6.2 Adapted Taiwanese cognitive assessments

The empirical T-ACE-III study had a number of key findings that deserve additional exploration. The language used (Hokkien and both Hokkien and Chinese) and years of education had an impact on the T-ACE-III scores, however, it did not influence the test accuracy when making a diagnosis of dementia. The participants in this study were only recruited from a northern hospital, and a more diverse representative sample of Taiwanese people should be investigated.

The present adaptation of the T-ACE-III was based on individuals with all-cause, all-stage dementia, with most having a diagnosis of probable Alzheimer’s disease or vascular dementia; future research should investigate its diagnostic accuracy

for less common forms of dementia, such as FTD or dementia with Lewy bodies (DLB). Furthermore, a wider variety of non-dementia controls should also be included in future research.

The face-to-face administration version of the T-ACE-III demonstrated an excellent ability to detect dementia. A parallel version of the T-ACE-III could be constructed for individuals with other Taiwanese dialects based on my current adapted version. In addition, an electronic version of the ACE-III (ACEmobile) has been developed on the iPad for research purposes to help eliminate examiner errors (e.g., administration, scoring); however, only two versions (English and Australian) are currently available (Newman et al., 2018). Future research could devise a Chinese or Taiwanese version of ACEmobile. While up until now, the accuracy of remote cognitive assessm ents by video-link has remained unclear; in view of a large degree of heterogeneity, the scoring roles and cut-off values derived from an in-person test may not be applicable when the same content is converted for remote use (Beishon et al., 2022).

The supplementary tests included in my thesis have not been extensively studied in Taiwan. The CIH test was only validated and compared against the Test of Memory Malingering (TOMM) in a pilot study (Yeh et al., 2019), which included 25 patients with dementia from the clinic, 9 non-dementia individuals, and 19 non-dementia individuals simulating dementia from the community. The results revealed that the CIH test, with a cut-off score of 8, is a quick and practical tool for detecting malingering. In my study, I did not use a group of non-dementia controls to simulate dementia malingering; consequently, such a comparison was not possible. In the case of the Brixton Spatial Anticipation Test, in my study this was tested for the acceptability among patients with dementia and adapted to Taiwan for the first time, and the WAIS-IV speed of processing subtests were also studied for the first time in a study of dementia patients; nevertheless, their completion rates were poor potentially due to a number of factors (e.g., fatigue, reluctance to take part) . The T-ACE-III demonstrated an excellent ability in differentiating patients with all-cause, all-stage dementia from those without dementia. The purpose of these supplementary tests was to explore whether they could assist with diagnosing less common forms of dementia, like frontotemporal

dementia or identify malingering. However, it was only possible to examine acceptability rather than validity as few patients in our study had those dementias, and none as far as we know were malingering.

9.7 Dissemination

I have already disseminated the findings of my thesis in conferences, seminars, and publications (See *11.1 Papers from my PhD* and *11.2 Conferences and other presentations*), and I intend to continue doing so in the future. My validated T-ACE-III questionnaire and scoring sheet have already been put into the public domain by relevant institutions (see *7.3.1 Data collection plan*). I have participated in another international collaborative project between the UK and Taiwan which aims uses the T-ACE-III to assess the participants' cognitive functions. It will enable the further investigation into the influence of demographic characteristics (such as language use) on T-ACE-III performance.

9.8 The potential impact of the COVID-19 pandemic on the administration of cognitive assessments

In 2020, under the threat of COVID-19, citizens of Taiwan were compelled to wear a mask in all indoor spaces, including hospitals and public transportation. During the assessment procedures, both participants and I wore masks, which may have impacted their ability to comprehend instructions or conversations, particularly for specific items, such as language tasks (e.g., asking participants to repeat a sentence, asking them to complete three actions) and attention tasks (e.g., asking participants to repeat three words). Consequently, this may have impacted their performances in T-ACE-III test performance, as well as the slightly longer than 20-minute average time required by each participant from the dementia group to complete the T-ACE-III. In addition to this, personal and environmental factors during the assessment session may also have contributed to variability in time required to complete the T-ACE-III. However, since the T-

ACE-III showed a high level of diagnostic accuracy for dementia, which was superior to those of MMSE and CASI and taking shorter time to administer than CASI, the T-ACE-III would appear to be an effective and useful diagnostic tool for detecting dementia in a Taiwan population.

In addition, because of the Taiwanese government's advice to avoid unnecessary hospital visits and older people’s concern about contracting COVID-19 in a hospital setting, the participants who were able to attend the clinic may have been more mobile, healthier, and less affected by dementia.

Chapter 10. Conclusions

My thesis offers clinicians new alternatives when using brief cognitive evaluations in clinical settings to detect dementia in Chinese-speaking populations.

My systematic review and meta-analysis were, to my knowledge, the first to focus on and meta-analyse data regarding brief cognitive assessments used in clinical settings with Chinese-speaking individuals with dementia and MCI.

My empirical study, the T-ACE-III study, was the first ACE-III made available for use with Taiwanese populations and was the first study to validate the T-ACE-III in a clinical context for the detection of dementia. It was also the first study to examine the acceptability of the Brixton Spatial Anticipation Test and the WAIS-IV speed of processing subtests among people with dementia.

10.1 Chinese brief cognitive assessments in clinical settings

My systematic review showed that cut-off scores for the same diagnostic test can vary depending on the particular Chinese-speaking population. It is essential to select the most appropriate tests and cut-off scores for the population and for the clinical setting. The ACE (ACE-R & ACE-III) demonstrated the most reliable diagnostic accuracy. In most studies, the translation and cultural adaption process for brief cognitive assessments was inadequately described.

10.2 Brief cognitive assessments in Taiwan

The findings of my empirical study highlight the fact that the T-ACE-III is a reliable, effective, and acceptable brief cognitive assessment tool for detecting dementia for a variety of aetiologies and stages in clinical settings in Taiwan. Years of education and language spoken have an impact on test performance.

The CIH test demonstrated excellent acceptability among all participants and a low correlation with T-ACE-III scores, while the Brixton Spatial Anticipation Test and the WASI-IV-PSI had low acceptability among patients with CDR scores of 1 and 2, and a moderate correlation with the T-ACE-III scores.

My work has included translating the Brixton Spatial Anticipation test into traditional Chinese for future researchers to employ and investigate further.

10.3 Conclusions

This thesis highlights the range of options in Chinese-speaking populations for brief cognitive assessments in clinical settings and emphasises the previous lack of a systematic review and meta-analysis of brief diagnostic tests for various settings in Chinese-speaking populations. I have now conducted a careful and extensive review and meta-analysis. It also shows the lack of reporting on translation and cultural adaptation procedures in the majority of studies. Future research should conduct more systematic reviews and meta-analysis studies on Chinese dementia assessment tools based on one specific setting (e.g., only community or clinic), and validation studies should fully report these procedures.

Chapter 11. Other related academic achievements

In this chapter, I list my other academic achievements related to my PhD program from April 2019 till present, including sharing the findings of the work through presentations at conferences, seminars, and other related achievements.

11.1 Papers from my PhD

I have one paper published from my PhD program:

- Yu, R. C., Mukadam, N., Kapur, N., Stott, J., Hu, C. J., Hong, C. T., ... & Livingston, G. (2021). Validation of the Taiwanese Version of ACE-III (T-ACE-III) to Detect Dementia in a Memory Clinic. *Archives of Clinical Neuropsychology*.

11.2 Conferences and other presentations

- Presented a poster at Alzheimer's Association International Conference (AAIC), online (July 2021). *Yu, R.C., Mukadam, N., Kapur, N., Stott, J., Livingston, G. (2021). Validation of the Addenbrooke's cognitive examination III to detect dementia in a Taiwanese population presenting to healthcare professionals for investigation of possible dementia.*
- Presented a poster at Institute of Mental Health (IoMH) International Conference online (September 2020). *Yu, R.C., Lai, J. C., Mukadam, N., Kapur, N., Stott, J., Livingston, G. (2020). Brief cognitive assessments used clinically for Chinese patients in the diagnosis of dementia and MCI– A systematic review.*
- Presented a poster at Alzheimer's Association International Conference (AAIC) online, originally to be held in Amsterdam, The Netherlands (July 2020). *Yu, R.C., Lai, J. C., Mukadam, N., Kapur, N., Stott, J., Livingston, G. (2020). A Systematic review of short cognitive assessments for diagnosing dementia and MCI in clinical settings in Chinese-speaking populations.*

- Presented work from my PhD at the departmental meeting of Mental Health of Older People at UCL, online (April 2020).
- Presented a short presentation at the 7th Meeting of the Federation of the European Societies of Neuropsychology (FESN) in Milan, Italy (September 7th, 2019). *Yu, R.C., Lai, J. C., Mukadam, N., Kapur, N., Livingston, G. (2019). Systematic review of brief cognitive instruments' validity and reliability for evaluating suspected dementia or MCI in clinical settings with Chinese-speaking pat1ients.*

11.3 Other achievements

- Reviewed one study as a journal reviewer on the topic of validation of a cognitive assessment (Addenbrooke's Cognitive Examination-Revised, ACE-R) from western to Korea culture.
- Culturally adapted and finalised the T-ACE-III and test instructions. It is available to download at: www.ftdrg.org or www.sydney.edu.au/brain-mind/resources-for-clinicians/dementia-test.html
- Culturally adapted the Social Functioning in Dementia Scale (SF-DEM) and its instruction into Taiwanese culture and finalised the Taiwan version of SE-FEM (T-SF-DEM).
- Contributed to the start-up of an international research collaboration between Shuang-Ho Hospital, Taiwan and UCL, United Kingdom on the validation of the SF-DEM: Social function in patients with dementia: collaborating with UCL to explore in Taiwanese population through SF-DEM.
- I trained one nurse, one clinical psychologist from Shuang-Ho Hospital in administering T-ACE-III, SF-DEM, AD-QoL, and the Brixton Spatial Anticipation test.
- I delivered a remote lecture to the National Dong-Hua university in Taiwan, for a course of research methods in psychology, on the topic of ethical principles when involving any human participants for research.

References

Abma, I. L., Rovers, M., & van der Wees, P. J. (2016). Appraising convergent validity of patient-reported outcome measures in systematic reviews: Constructing hypotheses and interpreting outcomes. *BMC Research Notes*, *9*(1), 1–5. https://doi.org/10.1186/S13104-016-2034-2/METRICS

ADI. (2014). *Dementia in the Asia Pacific Region*. Technical Report, Alzheimer's Disease International, Dementia. www.daviddesigns.co.uk

Adishesa, M. S., & Halim, M. S. (2015). Diagnostic Utility of the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV) Among Elders with Alzheimer's Dementia. *ANIMA Indonesian Psychological Journal*, *31*(1), 14–21. https://doi.org/10.24123/aipj.v31i1.560

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. https://doi.org/10.1176/APPI.BOOKS.9780890425596

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders, fifth edition, text revision (DSM-5-TR)*. https://doi.org/10.1176/APPI.BOOKS.9780890425787

Appels, B. A., & Scherder, E. (2010). The diagnostic accuracy of dementia-screening instruments with an administration time of 10 to 45 minutes for use in secondary care: a systematic review. *American Journal of Alzheimer's Disease & Other Dementias®*, *25*(4), 301–316. https://doi.org/10.1177/1533317510367485

Arevalo-Rodriguez, I., Smailagic, N., Roquéi Figuls, M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O. L., Bonfill Cosp, X., & Cullum, S. (2015). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews*, *2015*(3), 1–74. https://doi.org/10.1002/14651858.CD010783.pub2

Baddeley, A., & Hitch, G. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*, 485–493.

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and Operationalization of Executive Function. *Mind, Brain, and Education*, *10*(1), 10–33. https://doi.org/10.1111/MBE.12100

Ballard C, Alistar B, Corbett A, Livingston G, & Rasmussen J. (2013). *Helping you to assess cognition: A practical toolkit for clinicians.* London: Alzheimer's Society. https://www.wamhinpc.org.uk/sites/default/files/dementia-practical-toolkit-for-clinicians.pdf

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Beishon, L. C., Elliott, E., Hietamies, T. M., Mc Ardle, R., O'Mahony, A., Elliott, A. R., & Quinn, T. J. (2022). Diagnostic test accuracy of remote, multidomain cognitive assessment (telephone and video call) for dementia. *Cochrane Database of Systematic Reviews*, *2022*(4). https://doi.org/10.1002/14651858.CD013724.PUB2/MEDIA/CDSR/CD013724/IMAGE_N/NCD013724-TST-007.PNG

Berry, J. W. (1979). Culture and cognition style. In A. Marsella, R. G. Tharp, & T. J. Ciborowski (Eds.), *Perspectives in cross-cultural psychology* (pp. 117–135). New York: Academic Press.

Beynon, R., Leeflang, M. M. G., Mcdonald, S., Eisinga, A., Mitchell, R. L., Whiting, P., & Glanville, J. M. (2013). Search strategies to identify diagnostic accuracy studies in MEDLINE and EMBASE. *Cochrane Database of Systematic Reviews*, *2013*(9), 1–101. https://doi.org/10.1002/14651858.MR000022.pub3

Bosnes, I., Bosnes, O., Stordal, E., Nordahl, H. M., Myklebust, T., & Almkvist, O. (2022). Processing speed and working memory are predicted by components of successful aging: a HUNT study. *BMC Psychology*, *10*(1). https://doi.org/10.1186/S40359-022-00718-7

Brodaty, H., Low, L. F., Gibson, L., & Burns, K. (2006). What is the best dementia screening instrument for general practitioners to use? *The American Journal of Geriatric Psychiatry*, *14*(5), 391–400. https://doi.org/10.1097/01.JGP.0000216181.20416.b2

Burgess, P., & Shallice, T. (1997). *The Hayling and Brixton Tests*.

Burgess, P., & Simons, J. (2005). *The effectiveness of rehabilitation for cognitive deficits* (Halligan PW & Wade DT, Eds.). Oxford University Press, USA.

Burgess, P. W., & Shallice, T. (1996). Bizarre responses, rule detection and frontal lobe lesions. *Cortex*, *32*(2), 241–259. https://doi.org/10.1016/S0010-9452(96)80049-9

Cao, L. X., Wang, G., Guo, Q. H., Zhang, W., Bak, T., & Huang, Y. (2022). Utility of Chinese versions of Addenbrooke's Cognitive Examination: A narrative review. *Healthcare*, *10*(10), 1–13. https://doi.org/10.3390/HEALTHCARE10102052

Chan, C., Yung, C., & Pan, P. (2005). Screening of dementia in Chinese elderly adults by the clock drawing test and the time and change test. *Hong Kong Med J*, *11*(1), 13–19.

Chang, C. C., Kramer, J. H., Lin, K. N., Chang, W. N., Wang, Y. L., Huang, C. W., Lin, Y. T., Chen, C., & Wang, P. N. (2010). Validating the Chinese version of the Verbal Learning Test for screening Alzheimer's disease. *Journal of the International Neuropsychological Society*, *16*(2), 244–251. https://doi.org/10.1017/S1355617709991184

Chang, Y. T., Chang, C. C., Lin, H. S., Huang, C. W., Chang, W. N., Lui, C. C., Lee, C. C., Lin, Y. T., Chen, C. H., & Chen, N. C. (2012). Montreal cognitive assessment in assessing clinical severity and white matter hyperintensity in Alzheimer's disease with normal control comparison. *Acta Neurol Taiwan*, *21*(2), 64–73.

Chen, C. H., Wang, L. C., Ma, T. C., & Yang, Y. H. (2014). A walk-in screening of dementia in the general population in Taiwan. *Scientific World Journal*, *2014*. https://doi.org/10.1155/2014/243738

Chen, C. W., Chu, H., Tsai, C. F., Yang, H. L., Tsai, J. C., Chung, M. H., Liao, Y. M., Chi, M. ju, & Chou, K. R. (2015). The reliability, validity, sensitivity, specificity and predictive values of the Chinese version of the Rowland Universal Dementia Assessment Scale. *Journal of Clinical Nursing*, *24*(21–22), 3118–3128. https://doi.org/10.1111/jocn.12941

Chen, J. Y. (1996). Cognitive abilities screening instrument Chinese version 2.0 (CASI C-2.0): Report of additional psychometric properties and a possible dialectal bias. *Chinese Psychiatry*, *10*(4), 325–333.

Chen, K., Xu, Y., Chu, A., Ding, D., Liang, X., Nasreddine, Z. S., Dong, Q., Hong, Z., Zhao, Q., & Guo, Q. (2016). Validation of the Chinese version of Montreal cognitive assessment basic for screening mild cognitive impairment. *Journal of the American Geriatrics Society*, *64*(12), e285–e290. https://doi.org/10.1111/jgs.14530

Chiu, Y. C., Li, C. L., Lin, K. N., Chiu, Y. F., & Liu, H. C. (2008). Sensitivity and specificity of the Clock Drawing Test, incorporating Rouleau scoring system, as a screening instrument for questionable and mild dementia: scale development. *International Journal of Nursing Studies*, *45*(2008), 75–84. https://doi.org/10.1016/j.ijnurstu.2006.09.005

Chong, M. S., Lim, W. S., Chan, S. P., Feng, L., Niti, M., Yap, P., Yeo, D., & Ng, T. P. (2010). Diagnostic performance of the Chinese Frontal Assessment Battery in early cognitive impairment in an Asian population. *Dementia and Geriatric Cognitive Disorders*, *30*(6), 525–532. https://doi.org/10.1159/000321665

Chu, L., Ng, K. H. Y., Law, A. C. K., Lee, A. M., & Kwan, F. (2015). Validity of the Cantonese Chinese Montreal Cognitive Assessment in southern Chinese. *Geriatrics & Gerontology International*, *15*(1), 96–103. https://doi.org/10.1111/ggi.12237

Cochrane, H. J., Baker, G. A., & Meudell, P. R. (1998). Simulating a memory impairment: Can amnesics implicitly outperform simulators? *British Journal of Clinical Psychology*, *37*(1), 31–48. https://doi.org/10.1111/j.2044-8260.1998.tb01277.x

Cole, M. (1975). An ethnographic psychology of cognition. In R. Brislin, S. Bochner, & W. Lonner (Eds.), *Cross-cultural perspectives of learning* (pp. 157–175). Beverly Hills, CA: Sage.

Cortina, J. M. (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology*, *78*(1), 98–104.

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery, and Psychiatry*, *78*(8), 790. https://doi.org/10.1136/JNNP.2006.095414

Custodio, N., Duque, L., Montesinos, R., Alva-Diaz, C., Mellado, M., & Slachevsky, A. (2020). Systematic review of the diagnostic validity of brief cognitive screenings for early dementia detection in Spanish-speaking adults in Latin America. *Frontiers in Aging Neuroscience*, *12*(270), 1–13. https://doi.org/10.3389/fnagi.2020.00270

Dahabreh, I. J., Trikalinos, T. A., Lau, J., & Schmid, C. H. (2017). Univariate and bivariate likelihood-based meta-analysis methods performed comparably when marginal sensitivity and specificity were the targets of inference. *Journal of Clinical Epidemiology*, *83*, 8–17. https://doi.org/10.1016/J.JCLINEPI.2016.12.003

Daugherty, J. C., Querido, L., Quiroz, N., Wang, D., Hidalgo-Ruzzante, N., Fernandes, S., Pérez-García, M., de los Reyes-Aragon, C. J., Pires, R., & Valera, E. (2019). The Coin in Hand–Extended Version: Development and

Validation of a Multicultural Performance Validity Test. *Assessment*, *28*(1), 186–198. https://doi.org/10.1177/1073191119864652

de Faria, C. A., Alves, H. V. D., & Charchat-Fichman, H. (2015). The most frequently used tests for assessing executive functions in aging. *Dementia & Neuropsychologia*, *9*(2), 149. https://doi.org/10.1590/1980-57642015DN92000009

de Vet, H. C. W., Riphagen II, E. A., Aertgeerts, B., & Pewsner, D. (2008). Chapter 7: Searching for Studies. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4*.

Deeks, J., Bossuyt, P., Leeflang, M., Takwoingi, Y., & Flemyng, E. (2022). *Cochrane handbook for systematic reviews of diagnostic test accuracy. (Version 2.0)*. London: The Cochrane Collaboration, 2022.

Deeks, J. J., Macaskill, P., & Irwig, L. (2005). The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology*, *58*(9), 882–893. https://doi.org/10.1016/j.jclinepi.2005.01.016

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135. https://doi.org/10.1146/ANNUREV-PSYCH-113011-143750

Directorate-General of Budget Accounting and Statistics Executive Yuan. (2021). *The preliminary results of the 2020 population and housing census.* https://www.stat.gov.tw/public/Attachment/1831151816OM26MHO7.pdf

Egger, M., Higgins, J. P. T., & Smith, G. D. (2022). *Systematic reviews in health research: Meta-analysis in context*. John Wiley & Sons Ltd.

Elamin, M., Holloway, G., Bak, T. H., & Pal, S. (2016). The Utility of the Addenbrooke's Cognitive Examination Version Three in Early-Onset Dementia. *Dementia and Geriatric Cognitive Disorders*, *41*(1–2), 9–15. https://doi.org/10.1159/000439248

Eng, N., Vonk, J. M. J., Salzberger, M., & Yoo, N. (2018). A cross-linguistic comparison of category and letter fluency: Mandarin and English. *Https://Doi.Org/10.1177/1747021818765997*, *72*(3), 651–660. https://doi.org/10.1177/1747021818765997

Espy, K. A. (2004). Using Developmental, Cognitive, and Neuroscience Approaches to Understand Executive Control in Young Children. *Developmental Neuropsychology*, *26*(1), 379–384. https://doi.org/10.1207/S15326942DN2601_1

Fang, R., Wang, G., Huang, Y., Zhuang, J. P., Tang, H. D., Wang, Y., Deng, Y. L., Xu, W., Chen, S. D., & Ren, R. J. (2014). Validation of the Chinese version of Addenbrooke's cognitive examination-revised for screening mild Alzheimer's disease and mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *37*(3–4), 223–231. https://doi.org/10.1159/000353541

Fernandez, A. L., & Abe, J. (2017). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain 2017 6:1*, *6*(1), 1–35. https://doi.org/10.1007/S40167-017-0050-2

Ferrante Di Ruffano, L., Hyde, C. J., McCaffery, K. J., Bossuyt, P. M. M., & Deeks, J. J. (2012). Assessing the value of diagnostic tests: a framework for designing and evaluating trials. *BMJ*, *344*(7847). https://doi.org/10.1136/BMJ.E686

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.

Journal of Psychiatric Research, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Franzen, S., van den Berg, E., Goudsmit, M., Jurgens, C. K., van de Wiel, L., Kalkisim, Y., Uysal-Bozkir, Ö., Ayhan, Y., Rune Nielsen, T., & Papma, J. M. (2020). A Systematic Review of Neuropsychological Tests for the Assessment of Dementia in Non-Western, Low-Educated or Illiterate Populations. *Journal of the International Neuropsychological Society*, *26*(3), 331–351. https://doi.org/10.1017/S1355617719000894

Ganguli, M., Ratcliff, G., Huff, F. J., Belle, S., Kancel, M. J., Fischer, L., & Kuller, L. H. (2016). Serial Sevens Versus World Backwards: A Comparison of the Two Measures of Attention From the MMSE. *Https://Doi.Org/10.1177/089198879000300405*, *3*(4), 203–207. https://doi.org/10.1177/089198879000300405

Gauthier, S., Rosa-Neto, P., Morais, J. A., & Webster, C. (2021). *World Alzheimer Report 2021: Journey through the diagnosis of dementia.* London, England: Alzheimer's Disease International. https://www.alzint.org/resource/world-alzheimer-report-2021/

Gilbert, R., Logan, S., Moyer, V. A., & Elliott, E. J. (2001). Assessing diagnostic and screening tests: Part 1. Concepts. *The Western Journal of Medicine*, *174*(6), 405–409. https://doi.org/10.1136/EWJM.174.6.405

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *4*, 404–411. https://doi.org/10.1037/h0059831

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*, *46*(12), 1417–1432. https://doi.org/10.1016/0895-4356(93)90142-N

Guo, Q. H., Cao, X. Y., Zhou, Y., Zhao, Q. H., Ding, D., & Hong, Z. (2010). Application study of quick cognitive screening test in identifying mild cognitive impairment. *Neuroscience Bulletin*, *26*(1), 47–54. https://doi.org/10.1007/s12264-010-0816-4

Guo, Q., Hong, Z., Shi, W., Sun, Y.-M., & Lu, C. (2006). Boston naming test in Chinese elderly, patient with mild cognitive impairment and Alzheimer's dementia. *Chinese Mental Health Journal*, *20*(02), 81–84.

Guo, Q., Hong, Z., Yu, H., & Ding, D. (2004). Clinical Validity of the Chinese Version of Mattis Dementia Rating Scale in Differentiating Dementia of Alzheimer Type in Shanghai. *Chinese Journal of Clinical Psychology*, *12*(3), 237–243.

Guo, Q., Jin, L., Hong, Z., & Lu, C. (2007). A specific phenomenon of animal fluency test in Chinese elderly. In *Chinese Mental Health Journal* (Vol. 21, Issue 09).

Hanley, J. R., Baker, G. A., & Ledson, S. (1999). Detecting the faking of amnesia: a comparison of the effectiveness of three different techniques for distinguishing simulators from patients with amnesia. *Journal of Clinical and Experimental Neuropsychology*, *21*(1), 59–69. https://doi.org/10.1076/JCEN.21.1.59.936

Harbord, R. M., Whiting, P., Sterne, J. A. C., Egger, M., Deeks, J. J., Shang, A., & Bachmann, L. M. (2008). An empirical comparison of methods for meta-

analysis of diagnostic accuracy showed hierarchical models are necessary. *Journal of Clinical Epidemiology*, *61*(11), 1095–1103. https://doi.org/10.1016/J.JCLINEPI.2007.09.013

Heneghan, C. (2009). EBM resources on the new CEBM website. *BMJ Evidence-Based Medicine*, *14*(3), 67. https://doi.org/10.1136/ebm.14.3.67

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heo, J. H., Lee, K. M., Park, T. H., Ahn, J. Y., & Kim, M. K. (2012). Validation of the Korean Addenbrooke's Cognitive Examination for diagnosing Alzheimer's dementia and mild cognitive impairment in the Korean elderly. *Applied Neuropsychology: Adult*, *19*(2), 127–131. https://doi.org/10.1080/09084282.2011.643948

Higgins, J. P. T., & Li, T. J. (2022). Chapter 10: Exploring heterogeneity. In M. Egger, J. P. T. Higgins, & G. D. Smith (Eds.), *Systematic reviews in health research: Meta-analysis in context, third edition* (pp. 185–203). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119099369.CH10

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch V A. (2022). Chapter 10: Analysing data and undertaking meta-analyses. In *Cochrane handbook for systematic reviews of interventions version 6.3 (updated February 2022)*. www.training.cochrane.org/handbook.

Hitomi, K., Kawata, S., Hashimoto, R., Nishio, Y., Hayashi, A., Ogawa, N., Kanno, S., Hiraoka, K., Yokoi, K., Iizuka, O., & Mori, E. (2012). A Validation Study of the Japanese Version of the Addenbrooke's Cognitive Examination-Revised. *Dementia and Geriatric Cognitive Disorders Extra*, *2*(1), 29–37. https://doi.org/10.1159/000336909

Hodges, J. R., & Larner, A. J. (2016). Addenbrooke's cognitive examinations: ACE, ACE-R, ACE-III, ACEapp, and M-ACE. *Cognitive Screening Instruments: A Practical Approach*, 109–137. https://doi.org/10.1007/978-3-319-44775-9_6/COVER

Holland, P. W., & Wainer, H. (1993). *Differential item functioning.* . Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Hsieh, S., McGrory, S., Leslie, F., Dawson, K., Ahmed, S., Butler, C. R., Rowe, J. B., Mioshi, E., & Hodges, J. R. (2015). The Mini-Addenbrooke's Cognitive Examination: A New Assessment Tool for Dementia. *Dementia and Geriatric Cognitive Disorders*, *39*(1–2), 1. https://doi.org/10.1159/000366040

Hsieh, S., Schubert, S., Hoon, C., Mioshi, E., & Hodges, J. R. (2013). Validation of the Addenbrooke's Cognitive Examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *36*(3–4), 242–250. https://doi.org/10.1159/000351671

Huang, L., Chen, K. L., Lin, B. Y., Tang, L., Zhao, Q. H., Li, F., & Guo, Q. H. (2019). An abbreviated version of Silhouettes test: A brief validated mild cognitive impairment screening tool. *International Psychogeriatrics*, *31*(6), 849–856. https://doi.org/https://doi.org/10.1017/S1041610218001230

Huang, L., Chen, K. L., Lin, B. Y., Tang, L., Zhao, Q. H., Lv, Y. R., & Guo, Q. H. (2018). Chinese version of Montreal Cognitive Assessment Basic for discrimination among different severities of Alzheimer's disease.

Neuropsychiatric Disease and Treatment, *14*, 2133–2140. https://doi.org/10.2147/NDT.S174293

Hung, Y. H., Wang, T. L., & Yang, C. C. (2016). Chinese character fluency test: An investigation of psychometric properties. *The Clinical Neuropsychologist*, *30*(sup1), 1–14. https://doi.org/10.1080/13854046.2016.1166266

Huo, Z., Lin, J., Bat, B. K. K., Chan, J. Y. C., Tsoi, K. K. F., & Yip, B. H. K. (2021). Diagnostic accuracy of dementia screening tools in the Chinese population: A systematic review and meta-analysis of 167 diagnostic studies. *Age and Ageing*, *50*(4), 1093–1101. https://doi.org/10.1093/ageing/afab005

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Silverberg, N. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*(4), 535–562. https://doi.org/10.1016/J.JALZ.2018.02.018

Jessen, F., Amariglio, R. E., van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., van der Flier, W. M., Glodzik, L., van Harten, A. C., de Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., … Wagner, M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, *10*(6), 844–852. https://doi.org/10.1016/J.JALZ.2014.01.001

Junquera, A., Garcia-Zamora, E., Olazaran, J., Parra, M. A., & Fernandez-Guinea, S. (2020). Role of Executive Functions in the Conversion from Mild Cognitive Impairment to Dementia. *Journal of Alzheimer's Disease*, *77*(2), 641–653. https://doi.org/10.3233/JAD-200586

Kapur, N. (1994). The coin-in-the-hand test: a new "bed-side" test for the detection of malingering in patients with suspected memory disorder. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*(3), 385. https://doi.org/10.1136/JNNP.57.3.385

Karimi, L., Mahboub–Ahari, A., Jahangiry, L., Sadeghi-Bazargani, H., & Farahbakhsh, M. (2022). A systematic review and meta-analysis of studies on screening for mild cognitive impairment in primary healthcare. *BMC Psychiatry*, *22*(1), 1–16. https://doi.org/10.1186/S12888-022-03730-8/FIGURES/6

Kelly, P. J., Baker, G. A., van den Broek, M. D., Jackson, H., & Humphries, G. (2005). The detection of malingering in memory performance: The sensitivity and specificity of four measures in a UK population. *British Journal of Clinical Psychology*, *44*(3), 333–341. https://doi.org/10.1348/014466505X35687

Kim, S., Kang, Y., Yu, K. H., & Lee, B. C. (2016). Disproportionate decline of executive functions in early mild cognitive impairment, late mild cognitive impairment, and mild Alzheimer's disease. *Dementia and Neurocognitive Disorders*, *15*(4), 159–164. https://doi.org/10.12779/DND.2016.15.4.159

Kuhn, E., Moulinet, I., Perrotin, A., la Joie, R., Landeau, B., Tomadesso, C., Bejanin, A., Sherif, S., de La Sayette, V., Desgranges, B., Vivien, D.,

Poisnel, G., & Chételat, G. (2019). Cross-sectional and longitudinal characterization of SCD patients recruited from the community versus from a memory clinic: Subjective cognitive decline, psychoaffective factors, cognitive performances, and atrophy progression over time. *Alzheimer's Research & Therapy*, *11*(1), 1–16. https://doi.org/10.1186/S13195-019-0514-Z/FIGURES/4

Kwak, Y. T., Yang, Y., & Kim, G. W. (2010). Korean Addenbrooke's Cognitive Examination Revised (K-ACER) for differential diagnosis of Alzheimer's disease and subcortical ischemic vascular dementia. *Geriatrics & Gerontology International*, *10*(4), 295–301. https://doi.org/10.1111/J.1447-0594.2010.00624.X

Lamar, M., Swenson, R., Kaplan, E., & Libon, D. J. (2010). Characterizing Alterations in Executive Functioning Across Distinct Subtypes of Cortical and Subcortical Dementia. *Https://Doi.Org/10.1080/13854040490507127*, *18*(1), 22–31. https://doi.org/10.1080/13854040490507127

Lau, C. W., & Hoosain, R. (1999). Working memory and language difference in sound duration: A comparison of mental arithmetic in Chinese, Japanese, and English. *Psychologia*, *42*(3), 139–144.

Lee, M. T., Chang, W. Y., & Jang, Y. (2018). Psychometric and diagnostic properties of the Taiwan version of the Quick Mild Cognitive Impairment screen. *PloS One*, *13*(12), e0207851. https://doi.org/10.1371/journal.pone.0207851

Lee, T. M. C., Cheung, C. C. Y., Chan, J. K. P., & Chan, C. C. H. (2000). Trail making across languages. *Journal of Clinical and Experimental Neuropsychology*, *22*(6), 772–778. https://doi.org/10.1076/JCEN.22.6.772.954

Li, H., Jia, J., Yang, Z., & Moreau, N. (2016). Mini-Mental State Examination in Elderly Chinese: A Population-Based Normative Study. *Journal of Alzheimer's Disease : JAD*, *53*(2), 487–496. https://doi.org/10.3233/JAD-160119

Li, X. D., Jia, S. H., Zhou, Z., Jin, Y., Zhang, X. F., Hou, C. L., Zheng, W. J., Rong, P., & Jiao, J. S. (2018). The role of the Montreal Cognitive Assessment (MoCA) and its memory tasks for detecting mild cognitive impairment. *Neurological Sciences*, *39*(6), 1029–1034. https://doi.org/10.1007/s10072-018-3319-0

Li, X., Shen, M., Jin, Y., Jia, S., Zhou, Z., Han, Z., Zhang, X., Tong, X., & Jiao, J. (2021). Validity and reliability of the new Chinese version of the Frontal Assessment Battery-Phonemic. *Journal of Alzheimer's Disease*, *80*(1), 371–381. https://doi.org/10.3233/JAD-201028

Li, X. Y., Dai, J., Zhao, S. S., Liu, W. G., & Li, H. M. (2018). Comparison of the value of Mini-Cog and MMSE screening in the rapid identification of Chinese outpatients with mild cognitive impairment. *Medicine*, *97*(22), e10966. https://doi.org/10.1097/MD.0000000000010966

Li, X., Yang, L., Yin, J., Yu, N., & Ye, F. (2019). Validation study of the Chinese version of Addenbrooke's Cognitive Examination III for diagnosing mild cognitive impairment and mild dementia. *Journal of Clinical Neurology*, *15*(3), 313. https://doi.org/10.3988/JCN.2019.15.3.313

Liew, T. M., Feng, L., Gao, Q., Ng, T. P., & Yap, P. (2015). Diagnostic utility of Montreal cognitive assessment in the fifth edition of diagnostic and

statistical manual of mental disorders: major and mild neurocognitive disorders. *Journal of the American Medical Directors Association*, *16*(2), 144–148. https://doi.org/10.1016/j.jamda.2014.07.021

Lin, J. S., O'Connor, E., Rossom, C., Perdue, L. A., & Eckstrom, E. (2013). Screening for cognitive impairment in older adults: A systematic review for the U.S. preventive services task force. *Annals of Internal Medicine*, *159*(9), 601–612. https://doi.org/10.7326/0003-4819-159-9-201311050-00730/SUPPL_FILE/AIME201311050-00730_SUPPLEMENT.PDF

Lin, J. Y., Hsu, H. C., Wang, J. Y., & Ouyang, W. C. (2019). Comparison of the Taiwanese versions of the Addenbrooke's Cognitive Examination-III (ACE-III), Mini-Mental State Examination (MMSE), and Montreal Cognitive Assessment (MOCA) for screening mild cognitive impairment among older Taiwanese people. *International Journal of Gerontology*, *13*(2), 162–166. https://doi.org/10.6890/IJGE.201906_13(2).0013

Lin, K. N., Wang, P. N., Chen, C., Chiu, Y. H., Kuo, C. C., Chuang, Y. Y., & Liu, H. C. (2003). The three-item clock-drawing test: A simplified screening test for Alzheimer's disease. *European Neurology*, *49*(1), 53–58. https://doi.org/10.1159/000067026

Lin, K. N., Wang, P. N., Liu, C. Y., Chen, W. T., Lee, Y. C., & Liu, H. C. (2002). Cutoff scores of the Cognitive Abilities Screening Instrument, Chinese version in screening of dementia. *Dementia and Geriatric Cognitive Disorders*, *14*(4), 176–182. https://doi.org/10.1159/000066024

Lin SF. (2011). *The Performance of Psychiatric Patients Claiming for Labor Insurance Disability Allowance on Cognitive Tests for Detecting Simulated Cognitive Impairments.* .

Lisy, K., & Porritt, K. (2016). Narrative synthesis: Considerations and challenges. *JBI Evidence Implementation*, *14*(4), 201–201. https://doi.org/10.1097/01.XEB.0000511348.97198.8C

Liu, E., Nisenblat, V., Farquhar, C., Fraser, I., Bossuyt, P. M. M., Johnson, N., & Hull, M. L. (2015). Urinary biomarkers for the non-invasive diagnosis of endometriosis. *Cochrane Database of Systematic Reviews*, *2015*(12), 1–89. https://doi.org/10.1002/14651858.CD012019

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., … Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Lonie, J. A., Tierney, K. M., & Ebmeier, K. P. (2009). Screening for mild cognitive impairment: a systematic review. *International Journal of Geriatric Psychiatry*, *24*(9), 902–915. https://doi.org/10.1002/GPS.2208

Low, A., Lim, L., Lim, L., Wong, B., Silva, E., Ng, K. P., & Kandiah, N. (2020). Construct validity of the Visual Cognitive Assessment Test (VCAT) - A cross-cultural language-neutral cognitive screening tool. *International Psychogeriatrics*, *32*(1), 141–149. https://doi.org/https://doi.org/10.1017/S1041610219000504

Lu, H., Chan, S. S. M., & Lam, L. C. W. (2017). 'Two-level' measurements of processing speed as cognitive markers in the differential diagnosis of DSM-

5 mild neurocognitive disorders (NCD). *Scientific Reports 2017 7:1*, *7*(1), 1–8. https://doi.org/10.1038/s41598-017-00624-8

Macaskill, P., Gatsonis, C., Deeks, J., Harbord, R., & Takwoingi, Y. (2010). Chapter 10: Analysing and Presenting Results. In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.* http://srdta.cochrane.org/

Malhotra, C., Chan, A., Matchar, D., Seow, D., Chuo, A., & Do, Y. K. (2013). Diagnostic performance of short portable mental status questionnaire for screening dementia among patients attending cognitive assessment clinics in Singapore. *Annals of the Academy of Medicine, Singapore*, *42*(7), 315–319. https://doi.org/10.47102/annals-acadmedsg.V42N7p315

Mathuranath, P. S., Nestor, P. J., Berrios, G. E., Rakowicz, W., & Hodges, J. R. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*, *55*(11), 1613–1620. https://doi.org/10.1212/01.WNL.0000434309.85312.19

Matias-Guiu, J. A., Cortés-Martínez, A., Valles-Salgado, M., Rognoni, T., Fernández-Matarrubia, M., Moreno-Ramos, T., & Matías-Guiu, J. (2017). Addenbrooke's cognitive examination III: diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests. *International Psychogeriatrics*, *29*(1), 105–113. https://doi.org/10.1017/S1041610216001496

Matias-Guiu, J. A., Valles-Salgado, M., Rognoni, T., Hamre-Gil, F., Moreno-Ramos, T., & Matías-Guiu, J. (2017). Comparative diagnostic accuracy of the ACE-III, MIS, MMSE, MoCA, and RUDAS for screening of Alzheimer disease. *Dementia and Geriatric Cognitive Disorders*, *43*(5–6), 237–246.

McInnes, M. D. F., Moher, D., Thombs, B. D., McGrath, T. A., Bossuyt, P. M., Clifford, T., Cohen, J. F., Deeks, J. J., Gatsonis, C., Hooft, L., Hunt, H. A., Hyde, C. J., Korevaar, D. A., Leeflang, M. M. G., Macaskill, P., Reitsma, J. B., Rodin, R., Rutjes, A. W. S., Salameh, J. P., … Willis, B. H. (2018). Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. *JAMA*, *319*(4), 388–396. https://doi.org/10.1001/JAMA.2017.19163

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202.

Milne, A., Culverwell, A., Guss, R., Tuppen, J., & Whelton, R. (2008). Screening for dementia in primary care: A review of the use, efficacy and quality of measures. *International Psychogeriatrics*, *20*(5), 911–926. https://doi.org/10.1017/S1041610208007394

Ministry of Health and Welfare in Taiwan. (2020). *The political framework and action on dementia care and prevention 2.0 (including working items).* https://1966.gov.tw/LTC/cp-4020-42469-201.html

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, *21*(11), 1078–1085. https://doi.org/10.1002/GPS.1610

Mirza, N., Panagioti, M., Waheed, M. W., & Waheed, W. (2017). Reporting of the translation and cultural adaptation procedures of the Addenbrooke's Cognitive Examination version III (ACE-III) and its predecessors: A

systematic review. *BMC Medical Research Methodology*, *17*(1), 1–10. https://doi.org/10.1186/S12874-017-0413-6

Mitchell, A. J., & Malladi, S. (2010a). Screening and case finding tools for the detection of dementia. Part I: Evidence-based meta-analysis of multidomain tests. *The American Journal of Geriatric Psychiatry*, *18*(9), 759–782. https://doi.org/10.1097/JGP.0b013e3181cdecb8

Mitchell, A. J., & Malladi, S. (2010b). Screening and Case-Finding Tools for the Detection of Dementia. Part II: Evidence-Based Meta-Analysis of Single-Domain Tests. *The American Journal of Geriatric Psychiatry*, *18*(9), 783–800. https://doi.org/10.1097/JGP.0B013E3181CDECD6

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414. https://doi.org/10.1212/WNL.43.11.2412-A

Mukadam, N., Marston, L., Lewis, G., Mathur, R., Rait, G., & Livingston, G. (2022). Incidence, age at diagnosis and survival with dementia across ethnic groups in England: A longitudinal study using electronic health records. *Alzheimer's & Dementia*. https://doi.org/10.1002/ALZ.12774

Newman, C. G. J., Bevins, A. D., Zajicek, J. P., Hodges, J. R., Vuillermoz, E., Dickenson, J. M., Kelly, D. S., Brown, S., & Noad, R. F. (2018). Improving the quality of cognitive screening assessments: ACEmobile, an iPad-based version of the Addenbrooke's Cognitive Examination-III. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, *10*, 182–187. https://doi.org/10.1016/J.DADM.2017.12.003

Ng, K. P., Chiew, H. J., Lim, L., Rosa-Neto, P., Kandiah, N., & Gauthier, S. (2018). The influence of language and culture on cognitive assessment tools in the diagnosis of early cognitive impairment and dementia. *Expert Review of Neurotherapeutics*, *18*(11), 859–869. https://doi.org/10.1080/14737175.2018.1532792

Ng, T. P., Niti, M., Chiam, P. C., & Kua, E. H. (2007). Ethnic and educational differences in cognitive test performance on mini-mental state examination in Asians. *The American Journal of Geriatric Psychiatry : Official Journal of the American Association for Geriatric Psychiatry*, *15*(2), 130–139. https://doi.org/10.1097/01.JGP.0000235710.17450.9A

Nyaga, V. N., & Arbyn, M. (2022). Metadta: A Stata command for meta-analysis and meta-regression of diagnostic test accuracy data – A tutorial. *Archives of Public Health*, *80*(1), 1–15. https://doi.org/10.1186/S13690-021-00747-5/FIGURES/5

O'Callaghan, C., Hodges, J. R., & Hornberger, M. (2013). Inhibitory dysfunction in frontotemporal dementia: A review. *Alzheimer Disease & Associated Disorders*, *27*(2), 102–108. https://doi.org/10.1097/WAD.0B013E318265BBC9

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pan, F. F., Cui, L., Li, Q. J., & Guo, Q. H. (2022). Validation of a modified Chinese version of Mini-Addenbrooke's Cognitive Examination for detecting

mild cognitive impairment. *Brain and Behavior*, *12*(1). https://doi.org/10.1002/BRB3.2418

Pan, F. F., Wang, Y., Huang, L., Huang, Y., & Guo, Q. H. (2021). Validation of the Chinese version of Addenbrooke's Cognitive Examination III for detecting mild cognitive impairment. *Aging & Mental Health*, *26*(2), 384–391. https://doi.org/10.1080/13607863.2021.1881757

Parra, M. A., Ascencio, L. L., Urquina, H. F., Manes, F., & Ibáñez, A. M. (2012). P300 and neuropsychological assessment in mild cognitive impairment and Alzheimer dementia. *Frontiers in Neurology*, *3*. https://doi.org/10.3389/FNEUR.2012.00172

Perone, S., Almy, B., & Zelazo, P. D. (2018). Toward an Understanding of the Neural Basis of Executive Function Development. *The Neurobiology of Brain and Behavioral Development*, 291–314. https://doi.org/10.1016/B978-0-12-804036-2.00011-X

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/J.1365-2796.2004.01388.X

Phillips, R., Cheung, Y. B., Collinson, S. L., Lim, M. L., Ling, A., Feng, L., & Ng, T. P. (2015). The equivalence and difference between the English and Chinese language versions of the repeatable battery for the assessment of neuropsychological status. *The Clinical Neuropsychologist*, *29*(sup1), 1–18. https://doi.org/10.1080/13854046.2015.1034182

Plana, M. N., Pérez, T., & Zamora, J. (2021). New measures improved the reporting of heterogeneity in diagnostic test accuracy reviews: a metaepidemiological study. *Journal of Clinical Epidemiology*, *131*, 101–112. https://doi.org/10.1016/j.jclinepi.2020.11.011

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews A Product from the ESRC Methods Programme Peninsula Medical School, Universities of Exeter and Plymouth*.

Potts, C., Richardson, J., Bond, R. B., Price, R. K., Mulvenna, M. D., Zvolsky, P., Harvey, M., Hughes, C. F., & Duffy, F. (2021). Reliability of Addenbrooke's Cognitive Examination III in differentiating between dementia, mild cognitive impairment and older adults who have not reported cognitive problems. *European Journal of Ageing*, *19*(3), 495–507. https://doi.org/10.1007/S10433-021-00652-4/FIGURES/8

Power, M., Fell, G., & Wright, M. (2013). Principles for high-quality, high-value testing. *BMJ Evidence-Based Medicine*, *18*(1), 5–10. https://doi.org/10.1136/EB-2012-100645

Powsner, S., & Powsner, D. (2005). Cognition, copyright, and the classroom. *The American Journal of Psychiatry*, *162*(3), 627–628. https://doi.org/10.1176/APPI.AJP.162.3.627-A

Qian, S. X., Chen, K. L., Guan, Q. B., & Guo, Q. H. (2021). Performance of Mattis dementia rating scale-Chinese version in patients with mild cognitive impairment and Alzheimer's disease. *BMC Neurology*, *21*(1), 1–5. https://doi.org/10.1186/s12883-021-02173-0

Ramsay, G. (2017). Taiwanese Stories of Dementia. *Modern China*, *43*(2), 217–243. https://doi.org/10.1177/0097700416664486

Ranson, J. M., Kuźma, E., Hamilton, W., Muniz-Terrera, G., Langa, K. M., & Llewellyn, D. J. (2019). Predictors of dementia misclassification when using brief cognitive assessments. *Neurology: Clinical Practice*, *9*(2), 109–117. https://doi.org/10.1212/CPJ.0000000000000566

Reid, M. C., Lachs, M. S., & Feinstein, A. R. (1995). Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA*, *274*(8), 645–651. https://doi.org/10.1001/JAMA.1995.03530080061042

Reitsma, J. B., Glas, A. S., Rutjes, A. W. S., Scholten, R. J. P. M., Bossuyt, P. M., & Zwinderman, A. H. (2005). Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*, *58*(10), 982–990. https://doi.org/10.1016/J.JCLINEPI.2005.02.022

Riello, M., Rusconi, E., & Treccani, B. (2021). The Role of Brief Global Cognitive Tests and Neuropsychological Expertise in the Detection and Differential Diagnosis of Dementia. *Frontiers in Aging Neuroscience*, *13*, 219. https://doi.org/10.3389/FNAGI.2021.648310/BIBTEX

Rizzo, S., Venneri, A., & Papagno, C. (2002). Famous face recognition and naming test: a normative study. *Neurological Sciences 2002 23:4*, *23*(4), 153–159. https://doi.org/10.1007/S100720200056

Röhr, S., Pabst, A., Riedel-Heller, S. G., Jessen, F., Turana, Y., Handajani, Y. S., Brayne, C., Matthews, F. E., Stephan, B. C. M., Lipton, R. B., Katz, M. J., Wang, C., Guerchet, M., Preux, P. M., Mbelesso, P., Ritchie, K., Ancelin, M. L., Carrière, I., Guaita, A., … Sachdev, P. S. (2020). Estimating prevalence of subjective cognitive decline in and across international cohort studies of aging: A COSMIC study. *Alzheimer's Research & Therapy*, *12*(167), 1–14. https://doi.org/10.1186/s13195-020-00734-y

Rosen, W. G., Mohs, R. C., & Davis, K. L. (2006). A new rating scale for Alzheimer's disease. *Https://Doi.Org/10.1176/Ajp.141.11.1356*, *141*(11 (1356-1364),), 41. https://doi.org/10.1176/AJP.141.11.1356

Rosli, R., Tan, M. P., Gray, W. K., Subramanian, P., & Chin, A. V. (2016). Cognitive assessment tools in Asia: a systematic review. *International Psychogeriatrics*, *28*(2), 189–210. https://doi.org/10.1017/S1041610215001635

Rutter, C. M., & Gatsonis, C. A. (2001). A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine*, *20*(19), 2865–2884. https://doi.org/10.1002/SIM.942

Sandel, T. L. (2003). Linguistic capital in Taiwan: The KMT's Mandarin language policy and its perceived impact on language practices of bilingual Mandarin and Tai-gi speakers. *Language in Society*, *32*(4), 523–551. https://doi.org/10.1017/S0047404503324030

Schroeder, R. W., Peck, C. P., Buddin, W. H., Heinrichs, R. J., & Baade, L. E. (2012). The coin-in-the-hand test and dementia: More evidence for a screening test for neurocognitive symptom exaggeration. *Cognitive and Behavioral Neurology*, *25*(3), 139–143. https://doi.org/10.1097/WNN.0B013E31826B71C1

Senda, M., Terada, S., Takenoshita, S., Hayashi, S., Yabe, M., Imai, N., Horiuchi, M., & Yamada, N. (2020). Diagnostic utility of the Addenbrooke's Cognitive Examination - III (ACE-III), Mini-ACE, Mini-Mental State

Examination, Montreal Cognitive Assessment, and Hasegawa Dementia Scale-Revised for detecting mild cognitive impairment and dementia. *Psychogeriatrics : The Official Journal of the Japanese Psychogeriatric Society*, *20*(2), 156–162. https://doi.org/10.1111/PSYG.12480

Seshadri, M., & Mazi-Kotwal, N. (2012). A copyright-free alternative to the Mini-Mental State Examination is needed. *BMJ*, *345*, e8589. https://doi.org/10.1136/BMJ.E8589

Shi, J., Tian, J., Wei, M., Miao, Y., & Wang, Y. (2012). The utility of the Hopkins Verbal Learning Test (Chinese version) for screening dementia and mild cognitive impairment in a Chinese population. *BMC Neurology*, *12*(1), 1–8. https://doi.org/10.1186/1471-2377-12-136

Shim, Y. S., Yang, D. W., Kim, H. J., Park, Y. H., & Kim, S. Y. (2017). Characteristic differences in the Mini-Mental State Examination used in Asian countries. *BMC Neurology*, *17*(1), 1–9. https://doi.org/10.1186/S12883-017-0925-Z

Shulman, K. I., Herrmann, N., Brodaty, H., Chiu, H., Lawlor, B., Ritchie, K., & Scanlan, J. M. (2006). IPA survey of brief cognitive screening instruments. In *International Psychogeriatrics* (Vol. 18, Issue 2, pp. 281–294). https://doi.org/10.1017/S1041610205002693

Simel, D. L., & Bossuyt, P. M. M. (2009). Differences between univariate and bivariate models for summarizing diagnostic accuracy may not be large. *Journal of Clinical Epidemiology*, *62*(12), 1292–1300. https://doi.org/10.1016/J.JCLINEPI.2009.02.007

Simpson, J. R. (2014). DSM-5 and neurocognitive disorders. *Journal of the American Academy of Psychiatry and the Law*, *42*(2), 159–164.

Slot, R. E. R., Sikkes, S. A. M., Berkhof, J., Brodaty, H., Buckley, R., Cavedo, E., Dardiotis, E., Guillo-Benarous, F., Hampel, H., Kochan, N. A., Lista, S., Luck, T., Maruff, P., Molinuevo, J. L., Kornhuber, J., Reisberg, B., Riedel-Heller, S. G., Risacher, S. L., Roehr, S., … van der Flier, W. M. (2019). Subjective cognitive decline and rates of incident Alzheimer's disease and non-Alzheimer's disease dementia. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *15*(3), 465. https://doi.org/10.1016/J.JALZ.2018.10.003

Smith, E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657–1661.

Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*(2), 268–274. https://doi.org/10.1111/J.1365-2753.2010.01434.X

Spencer, R. J., Wendell, C. R., Giggey, P. P., Seliger, S. L., Katzel, L. I., & Waldstein, S. R. (2013). Judgment of line orientation: An examination of eight short forms. *Journal of Clinical and Experimental Neuropsychology*, *35*(2), 160–166. https://doi.org/10.1080/13803395.2012.760535

Statista. (2021). *The world's most spoken languages*. https://www.statista.com/chart/12868/the-worlds-most-spoken-languages/

Stegeman, I., & Leeflang, M. M. G. (2020). Meta-analyses of diagnostic test accuracy could not be reproduced. *Journal of Clinical Epidemiology*, *127*, 161–166. https://doi.org/10.1016/J.JCLINEPI.2020.06.033

Takenoshita, S., Terada, S., Yoshida, H., Yamaguchi, M., Yabe, M., Imai, N., Horiuchi, M., Miki, T., Yokota, O., & Yamada, N. (2019). Validation of Addenbrooke's cognitive examination III for detecting mild cognitive impairment and dementia in Japan. *BMC Geriatrics*, *19*(1), 1–8. https://doi.org/10.1186/S12877-019-1120-4/TABLES/4

Takwoingi, Y., & Deeks, J. J. (2022). Systematic Reviews of Diagnostic Accuracy. *Systematic Reviews in Health Research*, 296–323. https://doi.org/10.1002/9781119099369.CH16

Tanzer, N. (2005). Developing test for use in multiple languages and cultures: A plea for simultaneous development. In R. M. Hambleton (Ed.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 235–263). Mahwah, NJ: Erlbaum.

Teresi, J. A., Golden, R. R., Cross, P., Gurland, B., Kleinman, M., & Wilder, D. (1995). Item bias in cognitive screening measures: Comparisons of elderly white, Afro-American, Hispanic and high and low education subgroups. *Journal of Clinical Epidemiology*, *48*(4), 473–483. https://doi.org/10.1016/0895-4356(94)00159-N

The University of Sydney. (2023). *Diagnostic dementia tests - ACE-III.* Australia: Brain and Mind Centre, the University of Sydney (online). https://www.sydney.edu.au/brain-mind/resources-for-clinicians/dementia-test.html

Thomas, R. L., Zidan, M. A., & Slovis, T. L. (2015). What you need to know about statistics Part I: validity of diagnostic and screening tests. *Pediatric Radiology*, *45*(2), 146–152. https://doi.org/10.1007/S00247-014-2882-7/TABLES/2

Tsai, C. F., Lee, W. J., Wang, S. J., Shia, B. C., Nasreddine, Z., & Fuh, J. L. (2012). Psychometrics of the Montreal Cognitive Assessment (MoCA) and its subscales: Validation of the Taiwanese version of the MoCA and an item response theory analysis. *International Psychogeriatrics*, *24*(4), 651–658. https://doi.org/10.1017/S1041610211002298

Tsai, J. C., Chen, C. W., Chu, H., Yang, H. L., Chung, M. H., Liao, Y. M., & Chou, K. R. (2016). Comparing the sensitivity, specificity, and predictive values of the Montreal Cognitive Assessment and Mini-Mental State Examination when screening people for mild cognitive impairment and dementia in Chinese population. *Archives of Psychiatric Nursing*, *30*(4), 486–491. https://doi.org/10.1016/j.apnu.2016.01.015

Tsai, P. H., Liu, J. L., Lin, K. N., Chang, C. C., Pai, M. C., Wang, W. F., Huang, J. P., Hwang, T. J., & Wang, P. N. (2018). Development and validation of a dementia screening tool for primary care in Taiwan: Brain Health Test. *PloS One*, *13*(4), 1–14. https://doi.org/10.1371/journal.pone.0196214

Tsao F. F., & Lu F. H. (1999). *Exploration of the communication between Taiwan physicians and elderly patients from a perspective of discourse analysis*.

Tsoi, K. K. F., Chan, J. Y. C., Hirai, H. W., Wong, S. Y. S., & Kwok, T. C. Y. (2015). Cognitive tests to detect dementia: A systematic review and meta-analysis. *JAMA Internal Medicine*, *175*(9), 1450–1458. https://doi.org/10.1001/JAMAINTERNMED.2015.2152

Uzzell, B. P., Pontón, M. O., & Ardila, Alfredo. (2007). *International handbook of cross-cultural neuropsychology* (B. P. Uzzell, M. O. Pontón, & Alfredo. Ardila, Eds.). Mahwah, NJ: Lawrence Erlbaum Associates.

Valcour, V. G., Masaki, K. H., & Blanchette, P. L. (2002). The phrase: “no ifs, ands, or buts” and cognitive testing. Lessons from an Asian-American community. *Hawaii Medical Journal*, *61*(4), 72–74. http://evols.library.manoa.hawaii.edu/handle/10524/53707

van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: an overview. *European Review of Applied Psychology*, *54*(2), 119–135. https://doi.org/10.1016/J.ERAP.2003.12.004

van den Berg, E., Nys, G. M. S., Brands, A. M. A., Ruis, C., van Zandvoort, M. J. E., & Kessels, R. P. C. (2009). The Brixton spatial anticipation test as a test for executive function: Validity in patient groups and norms for older adults. *Journal of the International Neuropsychological Society*, *15*(5), 695–703. https://doi.org/10.1017/S1355617709990269

van der Flier, W. M., & Scheltens, P. (2005). Epidemiology and risk factors of dementia. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(Suppl5), v2–v7. https://doi.org/10.1136/JNNP.2005.082867

Velayudhan, L., Ryu, S. H., Raczek, M., Philpot, M., Lindesay, J., Critchfield, M., & Livingston, G. (2014). Review of brief cognitive tests for patients with suspected dementia. *International Psychogeriatrics*, *26*(8), 1247–1262. https://doi.org/10.1017/S1041610214000416

Waheed, W., Mirza, N., Waheed, M. W., Malik, A., & Panagioti, M. (2020). Developing and implementing guidelines on culturally adapting the Addenbrooke’s cognitive examination version III (ACE-III): A qualitative illustration. *BMC Psychiatry*, *20*(1), 1–13. https://doi.org/10.1186/S12888-020-02893-6/TABLES/5

Wang, B., Ou, Z., Gu, X., Wei, C., Xu, J., & Shi, J. (2017). Validation of the Chinese version of Addenbrooke’s Cognitive Examination III for diagnosing dementia. *International Journal of Geriatric Psychiatry*, *32*(12), e173–e179. https://doi.org/10.1002/gps.4680

Wang, B. R., Zheng, H. F., Xu, C., Sun, Y., Zhang, Y. D., & Shi, J. Q. (2019). Comparative diagnostic accuracy of ACE-III and MoCA for detecting mild cognitive impairment. *Neuropsychiatric Disease and Treatment*, *15*, 2647–2653. https://doi.org/10.2147/NDT.S212328

Wang, C. S. M., Pai, M. C., Chen, P. L., Hou, N. T., Chien, P. F., & Huang, Y. C. (2013). Montreal Cognitive Assessment and Mini-Mental State Examination performance in patients with mild-to-moderate dementia with Lewy bodies, Alzheimer’s disease, and normal participants in Taiwan. *International Psychogeriatrics*, *25*(11), 1839–1848. https://doi.org/10.1017/S1041610213001245

Wang, J., Halffman, W., & Zwart, H. (2003). Diagnosing tests: Using and misusing diagnostic and screening tests. *Journal of Personality Assessment*, *81*(3), 209–219. https://doi.org/10.1207/S15327752JPA8103_03

Wang, J., Halffman, W., & Zwart, H. (2021). The Chinese scientific publication system: Specific features, specific challenges. *Learned Publishing*, *34*(2), 105–115. https://doi.org/10.1002/LEAP.1326

Wang, X., & Cheng, Z. (2020). Cross-sectional studies: Strengths, weaknesses, and recommendations. *Chest*, *158*(1), S65–S71. https://doi.org/10.1016/J.CHEST.2020.03.012

Webster, J. P., Lamberton, P. H. L., & McConkey, G. A. (2016). The Toxoplasma gondii Model of Schizophrenia. *Handbook of Behavioral Neuroscience*, *23*, 225–241. https://doi.org/10.1016/B978-0-12-800981-9.00014-6

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV)*. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Wechsler%2C+D.+%282008%29.Wechsler+Adult+Intelligence+Scale%E2%80%93Fourth+Edition+%28WAIS%E2%80%93IV%29.+San+Antonio%2C+TX%3A+2008%3A+Pearson+Assessment.&btnG=

Wechsler, D. (2015a). *Wechsler Adult Intelligence Scale–Fourth edition (WAIS–IV): Manual for Taiwan.* Taipei, Taiwan: Chinese Behavioral Science Corporation.

Wechsler, D. (2015b). *Wechsler Adult Intelligence Scale–Fourth edition (WAIS–IV): Technical and interpretative manual for Taiwan*. Taipei, Taiwan: Chinese Behavioral Science Corporation.

Wei, M., Shi, J., Li, T., Ni, J., Zhang, X., Li, Y., Kang, S., Ma, F., Xie, H., & Qin, B. (2018). Diagnostic accuracy of the Chinese version of the trail-making test for screening cognitive impairment. *Journal of the American Geriatrics Society*, *66*(1), 92–99. https://doi.org/10.1111/jgs.15135

Wen, H. B., Zhang, Z. X., Niu, F. S., & Li, L. (2008). The application of Montreal cognitive assessment in urban Chinese residents of Beijing. *Zhonghua Nei Ke Za Zhi*, *47*(1), 36–39.

Werner, P., Heinik, J., Lin, R., & Bleich, A. (1999). “Yes” ifs, ands or buts: Examining performance and correlates of the repetition task in the Mini-Mental State Examination. *International Journal of Geriatric Psychiatry*, *14*(9), 719-25. https://doi.org/10.1002/(sici)1099-1166(199909)14:9<719::aid-gps2>3.0.co;2-k

WHO. (2001). *International Classification of Functioning, Disability and Health (ICF)*. https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health

WHO. (2003). *ICF Checklist Version 2.1a, Clinician Form. 2003*. https://sid-inico.usal.es/wp-content/uploads/2021/02/icfchecklist.pdf

WHO. (2004). *ICD-10: International statistical classification of diseases and related health problems: Tenth revision, 2nd ed.* World Health Organization. https://apps.who.int/iris/handle/10665/42980

WHO. (2012). *Dementia: A public health priority*. World Health Organization. https://apps.who.int/iris/handle/10665/75263

WHO. (2017). *Global action plan on the public health response to dementia 2017 - 2025*. World Health Organization. http://www.who.int/mental_health/neurology/dementia/action_plan_2017_2025/en/

Williams, C. H. (2016). *Teaching English reading in the Chinese-speaking world: Building strategies across scripts*. Springer: Science and Business Media.

Wilson, J. M. G., Jungner, G., & World Health Organization. (1968). *Principles and practice of screening for disease*. World Health Organization. https://apps.who.int/iris/handle/10665/37650

Wong, K. F., & Xiao, Y. (2010). Diversity and difference: Identity issues of Chinese heritage language learners from dialect backgrounds. *Heritage Language Journal*, *7*(2), 314–348. https://doi.org/10.46538/HLJ.7.2.8

Wong, L. L., Chan, C. C., Leung, J. L., Yung, C. Y., Wu, K. K., Cheung, S. Y. Y., & Lam, C. L. M. (2013). A validation study of the Chinese-Cantonese Addenbrooke's Cognitive Examination Revised (C-ACER). *Neuropsychiatric Disease and Treatment*, *9*, 731–737. https://doi.org/10.2147/NDT.S45477

Xu, G. L., Meyer, J. S., Huang, Y. G., Du, F., Chowdhury, M., & Quach, M. (2003). Adapting Mini-Mental State Examination for dementia screening among illiterate or minimally educated elderly Chinese. *International Journal of Geriatric Psychiatry*, *18*(7), 609–616. https://doi.org/10.1002/gps.890

Xu, S. H., Chen, L. H., Jin, X. Q., Yan, J., Xu, S. Z., Xu, Y., Liu, C. X., & Jin, Y. (2019). Effects of age and education on clock-drawing performance by elderly adults in China. *The Clinical Neuropsychologist*, *33*(sup1), 96–105. https://doi.org/10.1080/13854046.2019.1640285

Yang, L. L., Li, X. J., Yin, J., Yu, N. W., Liu, J., & Ye, F. (2019). A validation study of the Chinese version of the Mini-Addenbrooke's Cognitive Examination for screening mild cognitive impairment and mild dementia. *Journal of Geriatric Psychiatry and Neurology*, *32*(4), 205–210. https://doi.org/10.1177/0891988719841726

Yang, Y. (2016). Dementia in Taiwan area. *Translational Neuroscience and Clinics*, *2*(1), 38–45. https://doi.org/10.18679/CN11-6030_R.2016.011

Yeh, T. C., Chou, Y. C., Weng, J. P., Yeh, H. W., Kao, Y. C., Chiang, W. S., Chen, T. Y., Hu, J. M., & Tzeng, N. S. (2019). Detection of malingering in the memory of patients with dementia: A pilot study on coin-in-the-hand test in a Northern Taiwan Memory Clinic. *J Med Sci*, *39*(2), 81–89. https://doi.org/10.4103/JMEDSCI.JMEDSCI_100_18

Yeung, P. Y., Wong, L. L., Chan, C. C., Leung, J. L., & Yung, C. Y. (2014). A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. *Hong Kong Med J*, *20*(6), 504–510. https://doi.org/10.12809/hkmj144219

Yoshida, H., Terada, S., Honda, H., Ata, T., Takeda, N., Kishimoto, Y., Oshima, E., Ishihara, T., & Kuroda, S. (2011). Validation of Addenbrooke's cognitive examination for detecting early dementia in a Japanese population. *Psychiatry Research*, *185*(1–2), 211–214. https://doi.org/10.1016/J.PSYCHRES.2009.06.012

Yoshida, H., Terada, S., Honda, H., Kishimoto, Y., Takeda, N., Oshima, E., Hirayama, K., Yokota, O., & Uchitomi, Y. (2012). Validation of the revised Addenbrooke's Cognitive Examination (ACE-R) for detecting mild cognitive impairment and dementia in a Japanese population. *International Psychogeriatrics*, *24*(1), 28–37. https://doi.org/10.1017/S1041610211001190

Yu, R. C., Mukadam, N., Kapur, N., Stott, J., Hu, C. J., Hong, C. T., Yang, C. C., Chan, L., Huang, L. K., & Livingston, G. (2021). Validation of the

Taiwanese version of ACE-III (T-ACE-III) to detect dementia in a memory clinic. *Archives of Clinical Neuropsychology*, *37*(3), 692–703. https://doi.org/10.1093/arclin/acab089

Zhang, X. N., Lv, L. L., Min, G. W., Wang, Q. Y., Zhao, Y. R., & Li, Y. (2021). Overview of the Complex Figure Test and its clinical application in neuropsychiatric disorders, including copying and recall. *Frontiers in Neurology*, *12*, 680474. https://doi.org/10.3389/fneur.2021.680474

Zhang, Y. R., Ding, Y. L., Chen, K. L., Liu, Y., Wei, C., Zhai, T. T., Wang, W. J., & Dong, W. L. (2019). The items in the Chinese version of the Montreal cognitive assessment basic discriminate among different severities of Alzheimer's disease. *BMC Neurology*, *19*(1), 1–7. https://doi.org/10.1186/s12883-019-1513-1

Zhao, Q. H., Guo, Q. H., Li, F., Zhou, Y., Wang, B., & Hong, Z. (2013). The Shape Trail Test: Application of a new variant of the Trail making test. *PloS One*, *8*(2), e57333. https://doi.org/10.1371/journal.pone.0057333

Appendices

Appendix 1: Registered protocol on PROSPERO

NIHR | National Institute for Health Research

PROSPERO
International prospective register of systematic reviews

Home | About PROSPERO | How to register | Service information Search | My PROSPERO | Logout: **Ruan-Ching Yu**

Register your review now Edit your details

You have 1 records

My other records

These are records that have either been published or rejected and are not currently being worked on.

ID	Title	Status	Last edited
CRD42019134092	Systematic review and meta-analysis of the validity and reliability of brief cognitive instruments used in clinical settings with Chinese-speaking patients to evaluate suspected dementia or MCI	Registered	28/11/2022

Systematic Review Protocol

Research Question: A systematic review of the validity and reliability of brief cognitive instruments used in clinical settings with Chinese-speaking patients to evaluate suspected dementia or MCI

Database:
1. Embase 1974 to 2019 April 12
2. Ovid MEDLINE(R) 1946 to April Week 1 2019
3. PsycINFO 1806 to April Week 2 2019
4. PsycTESTS 1910 to March 2019
5. Web of Science core collection
6. Airiti Library (Chinese database)
7. CNKI (Chinese database)
8. The Cochrane Library (Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Methodology Register)

Search strategy:
1. Title, keywords or abstract: Chinese OR Mandarin OR Hokkien OR Hoklo OR Cantonese OR Hakka OR Taiwan* OR China OR Hong Kong* or Singapore* or Macao OR Malaysia
2. And Title, keywords or abstract: Alzheimer* OR AD OR dement* OR VaD OR FTD OR Mild cognitive impairment OR MCI OR memory loss
3. And all fields: assessment OR evaluation OR scale OR test OR tool OR Instrument OR battery OR measure* OR screen* OR diagnos* OR inventory* OR validat*

Inclusion:
1.original peer reviewed research
2.human subjects

Further detail:
Assessments used in studies should meet the criteria as follows---
1. the measure is validated as part of the study,
2. diagnosis purpose,
3. in Chinese speaking populations
4. cognitive assessments (other kinds of assessments will be excluded, e.g. behavioural assessments or functional assessments),
5. in a memory clinic or similar setting,
6. taking <20 minutes,
7. assessing the patient directly, not through an informant, and
8. face-to-face

Exclusion:
The study does not meet those criteria mentioned

Appendix 2: Operational criteria for the scoring of the Centre for Evidence-based Medicine (CEBM) diagnostics criteria.

Question	In this paper		
	Yes (2 points)	No (0 point)	Unclear (1 point)
1. Was the diagnostic test evaluated in a Representative spectrum of patients (like those in whom it would be used in practice)?	☐	☐	☐
2. Was the reference standard applied regardless of the index test result?	☐	☐	☐
3. Was there an independent, blind comparison between the index test and an appropriate reference (‘gold’) standard of diagnosis?	☐	☐	☐
4. What were the results? (Are test characteristics presented?)	☐	☐	☐
5. Were the methods for performing the test described in sufficient detail to permit replication?	☐	☐	☐

Appendix 3: PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	Page 40 (chapter 4)
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	Page 5 (Abstract)
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Page 42 (4.1.2 Differences between screening tests and diagnostic tests)
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	Page 40 (4.1.1 A lack of systematic reviews of brief cognitive assessments to be used with dementia and MCI in Chinese-speaking populations)
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	Page 42 (4.1.3 Aims)
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Page 43 (4.2.2 Data sources)
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Page 44 (4.2.4 Study selection)
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Page 43 (4.2.2 Data sources)
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Page 43 (4.2.3 Search strategy)
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Page 44 (4.2.4 Study selection) Page 45 (4.2.5 Data screening and selection) Page 47 (4.2.8 Meta-analysis)
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Page 45 (4.2.6 Data extraction and definition)
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g., study design, clinical setting).	Page 45 (4.2.6 Data extraction and definition)

Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	Page 46 (4.2.7 Quality assessment)
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g., sensitivity, specificity) and state the unit of assessment (e.g., per-patient, per-lesion).	Page 45 (4.2.6 Data extraction and definition) Page 190 (Appendix 4) Page 194 (Appendix 5) Appendix 5: Brief cognitive tests for MCI: Diagnostic performance (in alphabetical order)
Synthesis of results	14	Describe methods of handling data, combining results of studies, and describing variability between studies. This could include but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	Page 47 (4.2.8 Meta-analysis)
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	Page 47 (4.2.8 Meta-analysis)
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Page 47 (4.2.8 Meta-analysis)
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	Page 51 (4.3.1 Study selection, *Figure 1:* PRISMA diagram)
Study characteristics	18	For each included study provide citations and present key characteristics including a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	Page 52 (4.3.2 Study characteristics and quality analysis) Page 190 (Appendix 4) Page 194 (Appendix 5)
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	Page 52 (4.3.2 Study characteristics and quality analysis)
Results of individual studies	20	For each analysis in each study (e.g., unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	Page 57 (4.3.3 Instruments in meta-analyses) Page 191 (Appendix 4) Page 195 (Appendix 5) Page 198 (Appendix 6)
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	Page 57 (4.3.3 Instruments in meta-analyses)
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	Page 64-65 (4.3.3.4 Sensitivity analysis of univariate and bivariate analysis, 4.3.3.5 Meta-regression, 4.3.3.6 Publication bias)

DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	Page 65 (4.4.1 Summary of findings) Page 67 (4.4.3.1 Strengths)
Limitations	25	Discuss limitations from included studies (e.g., risk of bias and concerns regarding applicability) and from the review process (e.g., incomplete retrieval of identified research).	Page 66 (4.4.3 Heterogeneity) Page 68 (4.4.3.2 Weaknesses)
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g., the intended use and clinical role of the index test).	Page 69 (4.5 Conclusions)
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	Page 3 (Acknowledgements) Page 23 (1.3 Start of the project)

Appendix 4: Brief cognitive tests for dementia: Diagnostic performance (in alphabetical order)

SETTINGS, TEST CODE & AUTHOR			ILLINESS	REFERENCE STANDARD	NUMBERS (DEMENTIA/ND)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
ACE-R	**A**	C10	mild AD	NINCDS-ADRDA, CDR	25/51	67/68	92.0	85.7	6.43	0.09
	B	H3	dementia	DSM-IV	54/43	73/74	93.0	95.0	18.60	0.07
ACE-III	**B**	C9	dementia	VaD = NINDS-AIREN; AD = NINCDS-ADRDA	177/180	82/83	91.1	83.1	5.39	0.11
AFT	**B**	C11	AD	aMCI = NINCDS-ADRDA, Petersen 1999; AD = NINCDS-ADRDA, CDR	124/512	N/A	81.0	81.0	4.26	0.23
non-zodiac animal							85.0	81.0	4.47	0.19
BHT- cog	**B**	T5	dementia	NIAAA, CDR	422/166	9/10	91.5	87.3	7.20	0.10
BNT	**B**	C12	Mild AD	NINCDS-ADRDA, CDR	34/100	22/23	79.0	81.0	4.16	0.26
			Moderate AD	NINCDS-ADRDA, CDR	38/100	22/23	95.0	81.0	5.00	0.06
		C19	AD	NIAAA	139/211(MCI)	22	77.0	49.3	1.52	0.47
CDT	**A**	H1	dementia	DSM-IV	51/34	3/4	89.4	47.1	1.69	0.23
		T1	AD	DSM-IV, NINCDS-ADRDA, CDR	144/259	2/3	73.0	66.0	2.15	0.41
Command			AD	DSM-IV, NINCDS-ADRDA, CDR	144/259	10/11	67.0	75.0	2.68	0.44
		T2	Mild AD	DSM-IV, CDR, NINCDS-ADRDA,	42/40	8.5/11	60.0	72.0	2.14	0.56
			QD	CDR	34/40	9.5/11	74.2	56.4	1.70	0.46
			QD		34/42(mild AD)	8.5/11	60.0	39.0	0.98	1.03
Copy		T1	AD	DSM-IV, NINCDS-ADRDA, CDR	144/259	12/13	51.0	74.0	1.96	0.66
		T2	Mild AD		42/40	9.5/10	57.5	85.0	3.83	0.50
			QD		34/40	9.5/11	32.3	84.6	2.10	0.80
			QD		34/42(mild AD)	9.5/10	57.5	68.0	1.80	0.63
CFT-C	**B**	C19	AD	NIAAA	139/211(MCI)	29	52.5	83.9	3.26	0.57
CVVLT	**B**	T6	AD	NINCDS-ADRDA, DSM-IV	232/185	age<75: 1st trial = 4/5	81.0	77.0	3.52	0.25
						age<75: Total = 22/23	91.0	92.0	11.38	0.10
						age<75: 10m recall = 4/5	95.0	97.0	31.67	0.05
						age>75: 1st trial = 3/4	81.0	74.0	3.12	0.26
						age>75: Total = 18/19	86.0	92.0	10.75	0.15
						age>75: 10m recall = 4/5	96.0	92.0	12.00	0.04
						All: 1st trial = 3/4	77.0	80.0	3.85	0.29

SETTINGS, TEST & AUTHOR	CODE		ILLINESS	REFERENCE STANDARD	NUMBERS (DEMENTIA/ND)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
						All: Total = 20/21	92.0	91.0	10.22	0.09
						All: 10m recall = 3/4	93.0	97.0	31.00	0.07
DRS/ MDRS	**B**	C13	Mild AD	NIAAA, CDR	116/167 (MCI)	120	84.5	85.0	5.63	0.18
			Moderate AD	NIAAA, CDR	64/116 (mild AD)	103	79.7	78.4	3.69	0.26
		C14	Mild AD	NINCDS-ADRDA, DSM-IV	5/16	Illiterate = 90/91	81.0	86.0	5.79	0.22
					4/24	Primary school = 115/116	88.0	88.0	7.33	0.14
					23/65	Secondary school = 120/121	88.0	86.0	6.29	0.14
FAB-Phonemic	**A**	C1	AD	NIA-AA, CDR	76/123	12/13	93.4	82.9	5.46	0.08
					76/37 (naMCI)	12/13	86.1	82.7	4.98	0.17
					76/107 (aMCI)	11/12	77.5	70.7	2.65	0.32
HVLT (learning)	**A**	C2	AD	DSM-IV, NINCDS, ADRDA	97/249	15/16 (total learning, tl)	94.7	92.5	12.63	0.06
						18/19 (tl: age 50-64)	95.5	92.1	12.09	0.05
						14/15 (tl: age 65-80)	94.8	92.5	12.64	0.06
			dementia			15/16 (tl)	94.7	93.4	14.34	0.06
JLO	**B**	C19	AD	NIAAA	139/211 (MCI)	17	64.0	65.9	1.88	0.55
M-ACE	**A**	C3	Mild dementia	DSM-5, CDR	54/51	21/22	96.0	87.0	7.38	0.05
MMSE	**A**	C3	Mild dementia	DSM-5, CDR	54/51	25/26	88.0	87.0	6.77	0.14
		C7	dementia	AD: NINCDS-ADRDA, VaD: NINDS-AIREN	93/277	literates:22 illiterates:20	83.9	84.5	5.41	0.19
		C10	mild AD	NINCDS-ADRDA, CDR	25/51	23/24	100.0	93.7	15.87	0.00
		H2	dementia	DSM-IV	130/49	24/25	95.4	89.8	9.35	0.05
		T3	dementia	NIAAA, DSM-IV-TR	57/26	24	84.0	86.0	6.00	0.19
	B	H3	dementia	DSM-IV	54/43	25/26	96.0	88.0	8.00	0.05
		H4	AD	Dementia = DSM-IV, AD = NINCDS-ADRDA	64/115	24/25	94.0	98.0	47.00	0.06
		T7	Very mild AD	NINCDS- ADRDA, DSM-IV	52/97	26/27	94.2	83.5	4.78	0.21
		T9	AD	DSM-IV, NINCDS-ADRDA, NIAAA	31/36(MCI)	18/19	77.0	89.0	7.00	0.26

SETTINGS, TEST & AUTHOR	CODE		ILLINESS	REFERENCE STANDARD	NUMBERS (DEMENTIA/ND)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
MoCA	**A**	H2	dementia	DSM-IV	130/49	18/19	92.3	91.8	11.26	0.08
		T3	dementia	NIAAA, DSM-IV-TR	57/26	20	79.0	80.0	3.95	0.26
	B	C19	AD	NIAAA	139/211(MCI)	19	81.3	76.8	3.50	0.24
		S3	Major NCD	DSM-5	64/146	Overall = 21/22	92.0	96.0	23.00	0.08
					31/93	edu<6 = 20/21	94.0	100.0	∞	0.06
					33/53	edu>6 = 22/23	94.0	98.0	47.00	0.06
		H4	AD	Dementia = DSM-IV, AD = NINCDS-ADRDA	64/115	19/20	94.0	92.0	11.75	0.07
		T7	Very-mild AD	NINCDS- ADRDA, DSM-IV	52/97	22/23	82.7	87.6	6.67	0.20
		T8	AD	NINCDS-ADRDA	98/38	21/22	98.0	95.0	19.60	0.02
		T9	AD	DSM-IV, NINCDS-ADRDA, NIAAA	31/36(MCI)	11/12	77.0	84.0	4.81	0.27
MoCA-BC	**B**	C16	Mild AD	NIAAA, CDR	80/96 (MCI)	Low-level edu = 13	77.4	79.4	3.76	0.28
					180/379 (MCI)	Mid-level edu = 15	79.0	88.9	7.12	0.24
					85/188 (MCI)	Mid-level edu = 16	78.7	86.7	5.92	0.25
			Moderate AD	NIAAA, CDR	132/80 (mild AD)	Low-level edu = 10	70.5	81.2	3.75	0.36
					225/180 (mild AD)	Mid-level edu = 11	72.9	82.8	4.24	0.33
					84/85 (mild AD)	Mid-level edu = 13	76.2	69.4	2.49	0.34
Verbal fluency		C17	AD	NAI-AA	604/329	9/10	85.4	77.6	3.81	0.19
Orientation						5/6	91.5	83.8	5.65	0.10
Visual perception						6/7	75.1	82.6	4.32	0.30
Immediate recall						7/8	77.5	76.0	3.23	0.30
Delayed recall						4/5	93.6	77.2	4.11	0.08
Qmci	**B**	T9	AD	DSM-IV, NINCDS-ADRDA, NIAAA	31/36(MCI)	31/32	94.0	78.0	4.27	0.08
RUDAS	**A**	T4	dementia	NIAAA, DSM-IV-TR	53/22	22	76.0	81.0	4.00	0.30
Silhouettes test	**B**	C19	AD	NIAAA	139/211(MCI)	8/9	78.4	46.4	1.46	0.47
SPMSQ	**A**	S1	dementia	-	103/24 (NC+MCI)	4/5	78.0	75.0	3.12	0.29
						edu<6 = 5/6	72.0	43.0	1.26	0.65
						edu≥6 = 3/4	79.0	76.0	3.29	0.28
STT(A)	**B**	C20	AD	NINCDS-ADRDA	86/336	age<65, edu<12 = 80/81	91.7	72.1	3.29	0.12
					72/313	age<65; edu>12 = 70/71	87.2	77.8	3.93	0.16
					138/201	age>65, edu<12 = 90/91	84.6	66.7	2.54	0.23

SETTINGS, TEST & AUTHOR	CODE		ILLINESS	REFERENCE STANDARD	NUMBERS (DEMENTIA/ND)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
					125/301	age>65, edu>12 = 80/81	88.4	66.4	2.63	0.17
STT(B)	**B**	C20	AD	NINCDS-ADRDA	86/336	age<65, edu<12 = 220/221	92.4	75.0	3.70	0.10
					72/313	age<65; edu>12 = 200/201	90.7	72.5	3.30	0.13
					138/201	age>65, edu<12 = 240/241	76.4	69.9	2.54	0.34
					125/301	age>65, edu>12 = 220/221	89.5	67.0	2.71	0.16
		C19	AD	NIAAA	139/211(MCI)	203	66.9	72.5	2.43	0.46
T&C	**A**	H1	dementia	DSM-IV	51/34	45 sec	74.5	88.2	6.31	0.29
TMT (A)	**A**	C8	AD	NIAAA	108/1026	98/99	77.8	92.0	9.73	0.24
			VaD	NINDS Workshop	122/1026	77/78	85.7	81.6	4.66	0.18
TMT (B)	**A**	C8	AD	NIAAA	108/1026	188/189	83.3	91.8	10.16	0.18
			VaD	NINDS Workshop	122/1026	147/148	81.6	83.9	5.07	0.22
VCAT	**B**	S4	Mild AD	NIA-AA	121/117(MCI)	19/20	68.3	84.8	4.49	0.37

Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. A and B refer to the controls were from the clinical and community-based setting, respectively. AD = Alzheimer's disease; QD = questionable dementia; VaD = vascular dementia; aMCI = amnestic mild cognitive impairment; naMCI = nonamnestic MCI; NCD = neurocognitive disorder; ND=non-dementia; SN = sensitivity; SP = specificity; +LR = positive likelihood ratio; -LR = negative likelihood ratio. edu = years of education; NINDS-AIREN = the National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; CDR = Clinical Dementia Rating; DSM = Diagnostic and Statistical Manual of Mental Disorders; NIAAA = National Institute on Aging/Alzheimer's Association. ACE-R = Addenbrooke's cognitive Examination Revised; AFT = Animal Fluency Test; BHT-cog = Brain Health Test-Cog part; BNT = Boston Naming Test; CDT = Clock Drawing Test; CFT-C = Rey-Osteriche Complex Figure Test-Copy; CVVLT = Chinese version of the Verbal Learning Test; DRS/MDRS = Mattis dementia rating scale; FAB = Frontal Assessment Battery; HVLT = Hopkins Verbal Learning Test; JLO = Judgment of Line Orientation; M-ACE = Mini-Addenbrooke's Cognitive Examination; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; MoCA-BC = Montreal Cognitive Assessment Basic; Qmci = Quick Mild Cognitive Impairment Screen; RUDAS = Rowland Universal Dementia Assessment Scale; SPMSQ = Short Portable Mental Status Questionnaire; STT = Shape Trail Test; T&C = Time and Change Test; TMT = Trail-Making Test; VCAT = Visual Cognitive Assessment Test.

Appendix 5: Brief cognitive tests for MCI: Diagnostic performance (in alphabetical order)

SETTINGS, TEST CODE & AUTHOR			ILLINESS	REFERENCE STANDARD	NUMBERS (MCI/NMCI)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
ACE-R	**B**	C10	aMCI	Petersen et al 1999, CDR	75/51	85/86	86.7	70.6	2.95	0.19
		H3	MCI	Peterson's criteria	50/43	79/80	74.0	84.0	4.63	0.31
			CI		104/43	79/80	88.0	84.0	5.50	0.14
BHT- cog	**B**	T5	MCI	NIAAA, CDR	225/422 (dementia)	9/10	91.5	64.9	2.61	0.13
BNT	**B**	C12	aMCI	Petersen criteria	38/100	22/23	61.0	81.0	3.21	0.48
		C19	MCI	Petersen et al., 1999	211/241	24	70.6	55.2	1.58	0.53
CFT-C	**B**	C19	MCI	Petersen et al., 1999	211/241	32	46.9	76.8	2.02	0.69
DRS	**B**	C13	MCI	Portet et al., 2006, NIAAA	167/136	131	65.3	67.6	2.02	0.51
	A	C1	aMCI	NIA-AA, CDR, Petersen's criteria	106/123	14/15	77.0	64.2	2.15	0.36
FAB-			naMCI		37/123	15/16	62.3	58.3	1.49	0.65
Phonemic			aMCI/naMCI		106/37	14/15	56.6	64.2	1.58	0.68
FAB	**B**	S2	CI	MCI: Peterson 2004, DSM-IV; dementia: DSM-IV, CDR	80/100	Unadjusted = 12/13	92.0	78.7	4.32	0.10
						Age <75 years = 12/13	92.6	76.5	3.94	0.10
						Age ≥75 years = 12/13	83.3	81.8	4.58	0.20
						Edu <6 years = 12/13	77.8	95.2	16.21	0.23
						Edu ≥6 years = 13/14	91.8	70.3	3.09	0.12
HVLT (learning)	**A**	C2	aMCI	CDR, Folstein and Petersen's criteria	134/249	21/22 (total learning, tl)	69.1	70.7	2.36	0.44
						23/24 (tl: age 50-64)	70.0	71.8	2.48	0.42
						18/19 (tl: age 65-80)	77.6	56.2	1.77	0.40
						11/12 (recognition)	58.9	69.9	1.96	0.59
JLO	**B**	C19	MCI	Petersen et al., 1999	211/241	21	59.7	53.2	1.28	0.76
M-ACE	**A**	C3	MCI	Petersen's criteria, CDR	64/51	25/26	88.0	72.0	3.14	0.17
Mini-Cog	**A**	C4	MCI	Petersen's criteria	119/110	N/A	85.7	79.4	4.16	0.18
MMSE	**A**	C3	MCI	Petersen's criteria, CDR	64/51	27/28	82.0	44.0	1.46	0.41
		C6	amMCI	CDR, MMSE, ADL, RAVLT, ROCF	56/53	27/28	74.0	77.0	3.22	0.34
			asMCI		32/53	28/29	44.8	77.0	1.95	0.72
		C4	MCI	Petersen's criteria	119/110	N/A	64.8	71.6	2.28	0.49
		H2	MCI	Peterson's criteria	93/49	26/27	78.5	81.6	4.27	0.26
			CI	DSM-IV, Peterson's criteria	223/49	26/27	91.5	75.5	3.73	0.11
		T3	MCI	NIAAA, DSM-IV-TR	59/26	27	88.0	70.0	2.93	0.17

SETTINGS, TEST CODE & AUTHOR			ILLINESS	REFERENCE STANDARD	NUMBERS (MCI/NMCI)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
	B	C10	aMCI	Petersen et al 1999, CDR	75/51	27/28	52.0	86.3	3.80	0.56
		C15	MCI	CDR, MMSE, Petersen et al, 1999	63/58	edu ≤6 = 26	86.2	60.3	2.17	0.23
					113/112	edu 7-12 = 27	78.6	52.2	1.64	0.41
					88/110	Edu >12 = 28	76.4	53.4	1.64	0.44
		C18	MCI	Petersen et al. 1999	121/186	26	83.3	38.3	1.35	0.44
		H3	MCI	Peterson's criteria	50/43	26/27	76.0	81.0	4.00	0.30
			CI	Dementia = DSM-IV, MCI = Peterson's criteria	104/43	25/26	82.7	88.4	7.13	0.20
		H4	aMCI	Petersen et al.1999	87/115	27/28	67.0	83.0	3.94	0.40
		T9	MCI	DSM-IV, NINCDS-ADRDA, NIAAA	36/35	26/27	69.0	97.0	23.00	0.32
MoCA	**A**	C5	MCI	Petersen's criteria	66/215	25/26	92.4	88.4	7.97	0.09
		C6	amMCI	CDR, MMSE, ADL, RAVLT, ROCF	56/53	24/25	88.0	66.7	2.64	0.18
			naMCI		33/53	25/26	65.5	56.3	1.50	0.61
Delayed free recall			amMCI		56/53	2/3	83.3	66.0	2.45	0.25
			asMCI		32/53	2/3	55.2	66.0	1.62	0.68
Category prompted recall			amMCI		56/53	3/4	85.4	66.0	2.51	0.22
			asMCI		32/53	3/4	51.7	66.0	1.52	0.73
Multiple choice recognition			naMCI		33/53	4/5	44.8	89.6	4.31	0.62
		H2	MCI	Peterson's criteria	93/49	21/22	82.8	73.5	3.12	0.23
			CI	DSM-IV, Peterson's criteria	223/49	21/22	92.8	73.5	3.50	0.10
		T3	MCI	NIAAA, DSM-IV-TR	59/26	24	88.0	74.0	3.38	0.16
	B	C18	MCI	Petersen et al. 1999	121/186	23	79.6	72.7	2.92	0.28
		C19	MCI	Petersen et al. 1999	211/241	24	81.5	65.1	2.34	0.28
		H4	aMCI	Petersen et al.1999	87/115	22/23	78.0	73.0	2.89	0.30
		S3	Mild NCD	DSM-5	41/146	24/25	78.0	62.0	2.05	0.35
					22/93	Edu<6 = 22/23	68.0	85.0	4.53	0.38
					19/53	Edu>6 = 22/23	37.0	93.0	5.29	0.68
		T8	MCI	Petersen et al. 2001	71/ 38	23/24	92.0	78.0	4.18	0.10
		T9	MCI	DSM-IV, NINCDS-ADRDA, NIAAA	36/35	23/24	94.0	85.0	6.27	0.07
MoCA-BC	**B**	C15	MCI	CDR, MMSE, Petersen et al, 1999	63/58	edu ≤6 = 19	87.9	81.0	4.63	0.15
					113/112	edu 7-12 = 22	92.9	91.2	10.56	0.08
					88/110	edu>12 = 24	89.8	90.9	9.87	0.11
		C16	MCI	Petersen et al 1999	96/82	Low-level edu = 19	79.4	70.6	2.70	0.29

SETTINGS, TEST CODE & AUTHOR			ILLINESS	REFERENCE STANDARD	NUMBERS (MCI/NMCI)	CUT-OFF SCORE	SN (%)	SP (%)	+LR	-LR
					379/285	Mid-level edu = 22	77.7	83.0	4.57	0.27
					188/153	high-level edu = 24	89.9	68.6	2.86	0.15
Verbal fluency		C17	MCI	Petersen et al, 1999	456/329	9/10	72.3	55.4	1.62	0.50
Orientation						5/6	91.5	24.3	1.21	0.35
Visual perception						7/8	75.1	51.7	1.55	0.48
Immediate recall						8/9	54.7	71.2	1.90	0.64
Delay recall						9/10	63.2	83.1	3.74	0.44
QCST	**B**	C18	MCI	Petersen et al. 1999	121/186	edu5-8 = 63/64	89.4	91.0	9.93	0.12
						edu9-12 = 65/66	89.3	94.3	15.67	0.11
						edu≥13 = 68/69	86.7	78.2	3.98	0.17
						NA	87.6	84.3	5.58	0.15
Qmci	**B**	T9	MCI	DSM-IV, NINCDS-ADRDA, NIAAA	36/35	51/52	69.0	97.0	23.00	0.32
RUDAS	**A**	T4	MCI	NIAAA, DSM-IV-TR	55/22	23/24	79.0	91.0	8.78	0.23
silhouettes test	**B**	C19	MCI	Petersen et al., 1999	211/241	10	79.6	65.1	2.28	0.31
STT(B)	**B**	C19	MCI	Petersen et al., 1999	211/241	169	50.7	80.0	2.54	0.62
TMT (A)	**A**	C8	MCI	MCI Working Group of EADC	462/1026	72/73	48.4	78.4	2.24	0.66
			VaMCI	Gorelick et al., 2011 guideline	113/1026	63/64	70.5	67.7	2.18	0.44
TMT (B)	**A**		MCI	MCI Working Group of EADC	462/1026	135/136	51.8	80.2	2.62	0.60
			VaMCI	Gorelick et al., 2011 guideline	113/1026	126/127	62.9	75.9	2.61	0.49
VCAT	**B**	S4	CI (MCI, AD)	NIA-AA	238/233	24/25	75.4	71.1	2.61	0.35

Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively. A and B refer to the controls were from the clinical and community-based setting, respectively. CI = cognitive impairment; aMCI = amnestic MCI; amMCI = aMCI-multiple domains; asMCI = aMCI-single domain; naMCI = nonamnestic MCI; VaMCI = vascular MCI; NMCI=non-MCI; SN = sensitivity; SP = specificity; +LR = positive likelihood ratio; -LR = negative likelihood ratio. edu = years of education; NINDS-AIREN = the National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l'Enseignement en Neurosciences; NINCDS-ADRDA = National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; CDR = Clinical Dementia Rating; DSM = Diagnostic and Statistical Manual of Mental Disorders; NIAAA = National Institute on Aging/Alzheimer's Association. ACE-R = Addenbrooke's cognitive Examination Revised; BNT = Boston Naming Test; CFT-C = Rey-Osteriche Complex Figure Test-Copy; DRS/MDRS = Mattis dementia rating scale-Chinese version; FAB = Frontal Assessment Battery; HVLT = Hopkins Verbal Learning Test; JLO = Judgment of Line Orientation; M-ACE = Mini-Addenbrooke's Cognitive Examination; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; MoCA-BC = Montreal Cognitive Assessment Basic; QCST = Quick Cognitive Screening Test; Qmci = Quick Mild Cognitive Impairment Screen; RUDAS = Rowland Universal Dementia Assessment Scale; STT = Shape Trail Test; TMT = Trail-Making Test; VCAT = Visual Cognitive Assessment Test

Appendix 6: 2x2 data table

Test for dementia		TP	FN	TN	FP
ACE	C9- Wang 2017	161	16	150	30
	C10- Fang 2014	23	2	44	7
	H3- Wong 2013	50	4	41	2
		234	**22**	**235**	**39**
CDT	H1- Chan 2005	46	5	16	18
	T1- Lin 2003	105	39	171	88
		151	**44**	**187**	**106**
DRS	C14- Guo 2004	4	1	14	4
		4	0	21	4
		20	3	56	20
		28	**4**	**91**	**14**
MMSE	C3- Yang 2019	48	6	44	7
	C7- Xu 2003	78	15	234	43
	H2- Yeung 2014	124	6	44	5
	T3- Tsai 2016	48	9	22	4
	C10- Fang 2014	25	0	48	3
	H3- Wong 2013	52	2	38	5
	H4- Chu 2015	60	4	113	2
	T7- Chang 2012	49	3	81	16
		484	**45**	**624**	**85**
MoCA	H2- Yeung 2014	120	10	45	4
	T3- Tsai 2016	45	12	21	5
	H4- Chu 2015	60	4	106	9
	S3- Liew 2015	29	2	93	0
		31	2	52	1
	T7- Chang 2012	43	9	85	12
	T8- Tsai 2012	96	2	36	2
		424	**41**	**438**	**33**
STT-A	C20- Zhao 2013	79	7	242	94
		63	9	244	69
		117	21	134	67
		111	15	200	101
		370	**52**	**820**	**331**
STT-B	C20- Zhao 2013	79	7	252	84
		65	7	227	86
		105	33	140	61
		112	13	202	99
		361	**60**	**821**	**330**

Test for MCI		TP	FN	TN	FP
ACE	C10- Fang 2014	65	10	36	15
	H3- Wong 2013	37	13	36	7
		102	**23**	**72**	**22**
MMSE	C3- Yang 2019	52	12	22	29
	H2- Yeung 2014	73	20	40	9
	T3- Tsai 2016	52	7	18	8
	C10- Fang 2014	39	36	44	7
	C15- Chen 2016	54	9	35	23
		89	24	58	54
		67	21	59	51
	C18- Guo 2010	101	20	71	115
	H3- Wong 2013	38	12	35	8
	H4- Chu 2015	58	29	95	20
	T9- Lee 2018	25	11	34	1
		648	**201**	**511**	**325**
MoCA	C5- Wen 2008	61	5	190	25
	H2- Yeung 2014	77	16	36	13
	T3- Tsai 2016	52	7	19	7
	C18- Guo 2010	96	25	135	51
	C19- Huang 2019	172	39	157	84
	H4- Chu 2015	68	19	84	31
	S3- Liew 2015	15	7	79	14
		7	12	49	4
	T8- Tsai 2012	65	6	30	8
	T9- Lee 2018	34	2	30	5
		647	**138**	**809**	**242**
MoCA-BC	C15- Chen 2016	55	8	47	11
		105	8	102	10
		79	9	90	20
	C16- Huang 2018	76	20	58	24
		295	85	237	49
		169	19	105	48
		778	**149**	**649**	**151**

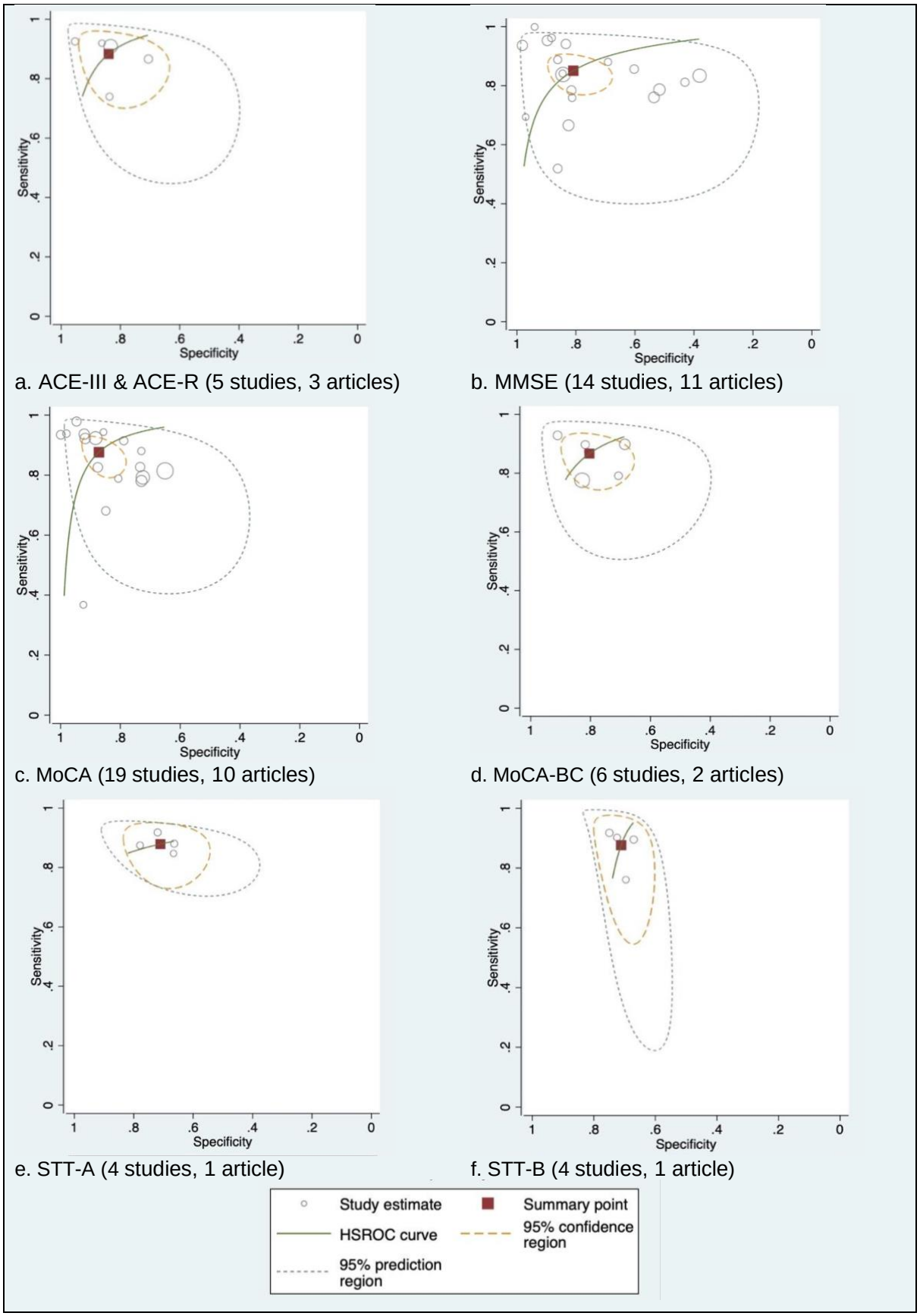
Note. The codes starting with syllables of C, H, S, T refer to China, Hong Kong, Singapore, and Taiwan, respectively; TP = true positive; FN = false negative, TN = true negative, FP = false positive; ACE = Addenbrooke's Cognitive Examination III & Revised; CDT = Clock Drawing Test; DRS = Mattis Dementia Rating Scale; MMSE = Mini-Mental State Examination; MoCA = Montreal cognitive assessment; MoCA-BC = Montreal cognitive assessment-Basic; STT-A&B = Shape Trail Test-A & B

Appendix 7: Random-effect bivariate model analysis

Disease classification	Test	Sensitivity			Specificity			Generalized		covar	rho	RE vs FE model X2
		Pooled	Tau^2	I^2	Pooled	Tau^2	I^2	Tau^2	I^2			(*p* value)
Dementia	**ACE**	0.92 (0.87-0.95)	0.00	0.01	0.87 (0.77-0.93)	0.04	0.22	0.00	0.00	0.01	1.00	0 (0.99)
	CDT	0.77 (0.71-0.83)	0.00	0.00	0.64 (0.58-0.69)	0.00	0.00					
	DRS	0.87 (0.71-0.95)	0.00	0.00	0.87 (0.79-0.92)	0.00	0.00	0.00	0.00	-0.00	-0.08	0 (1.00)
	MMSE	0.92 (0.88-0.95)	0.29	0.49	0.90 (0.85-0.93)	0.26	0.54	0.01	0.30	0.25	1.00	14 (0.00)
	MoCA	0.93 (0.87-0.96)	0.54	0.60	0.94 (0.88-0.97)	0.72	0.59	0.00	0.00	0.62	1.00	18 (0.00)
	STT-A	0.88 (0.84-0.91)	0.01	0.07	0.71 (0.66-0.76)	0.04	0.68	0.00	0.00 [d]	0.02	1.00	4 (0.24)
	STT-B	0.88 (0.80-0.93)	0.22	0.69	0.71 (0.68-0.74)	0.01	0.30	0.00	0.33 [d]	0.04	0.86	8 (0.06)
MCI	**ACE**	0.82 (0.74-0.87)	0.00	0.00	0.77 (0.67-0.84)	0.00	0.00			0.00		
	MMSE	0.76 (0.70-0.81)	0.19	0.69	0.70 (0.56-0.81)	0.86	0.88	0.05	0.69 [d]	-0.34	-0.84	122 (0.00*)
	MoCA	0.83 (0.74-0.89)	0.56	0.76	0.80 (0.73-0.85)	0.23	0.68	0.12	0.72 [d]	-0.08	-0.21	42 (0.00*)
	MoCA-BC	0.87 (0.81-0.91)	0.20	0.70	0.82 (0.74-0.89)	0.32	0.81	0.05	0.74 [d]	0.11	0.43	36 (0.00*)

Note. The pooled sensitivity and specificity are based on marginal summary measures of test accuracy (absolute measures); [d] The random-effect bivariate model had a heterogeneity lower than 75% while that of the random-effect univariate model was higher than 75%. Addenbrooke's Cognitive Examination-Revised & III; CDT = Clock Drawing Test; DRS/MDRS = Mattis dementia rating scale-Chinese version; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment (including different scoring systems); MoCA-BC = Montreal Cognitive Assessment Basic; STT = Shape Trail Test.

Appendix 8: The HSROC / SROC curves of the eight tests



a. ACE-III & ACE-R (5 studies, 3 articles)

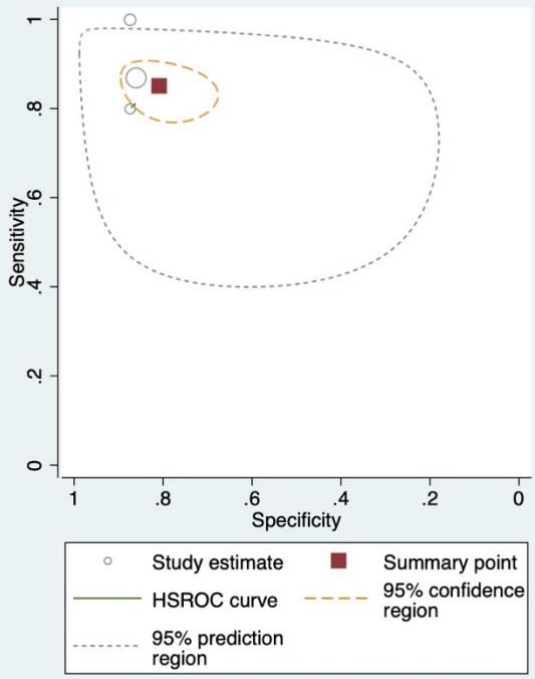
b. MMSE (14 studies, 11 articles)

c. MoCA (19 studies, 10 articles)

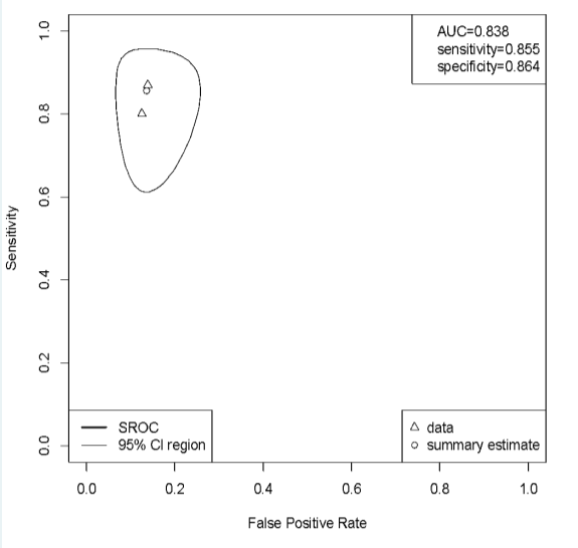
d. MoCA-BC (6 studies, 2 articles)

e. STT-A (4 studies, 1 article)

f. STT-B (4 studies, 1 article)



g. DRS (3 studies, 1 article)



a. CDT (2 studies, 2 articles)

Appendix 9: Likelihood ratio regression based on bivariate models (clinical context)

Test	Where the controls were recruited (N)		Between-study heterogeneity (Tau^2)	Generalised Tau^2	Model comparison: χ^2 (*P*-value)
MMSE (dementia)	**Clinic (5)**	**Community (3)**			
Sensitivity	0.91 (0.85, 0.94)	0.95 (0.89, 0.98)	0.19	0.00 (χ^2 = 8.83, df = 3, *p* = 0.0317)	1.66 (0.20)
Specificity	0.89 (0.82, 0.93)	0.92 (0.84, 0.96)	0.23		0.41 (0.52)
MMSE (MCI)	**Clinic (4)**	**Community (7)**			
Sensitivity	0.76 (0.65, 0.84)	0.77 (0.69, 0.83)	0.19	0.05 (χ^2 = 113.33, df = 3, *p* = 0.00*)	0.01 (0.91)
Specificity	0.72 (0.50, 0.87)	0.69 (0.51, 0.82)	0.85		0.09 (0.76)
MoCA (dementia)	**Clinic (2)**	**Community (5)**			
Sensitivity	0.87 (0.72, 0.95)	0.94 (0.88, 0.97)	0.39	0.00 (χ^2 = 14.52, df = 3, *p* = 0.0023)	1.78 (0.18)
Specificity	0.88 (0.68, 0.96)	0.96 (0.90, 0.98)	0.48		2.24 (0.13)
MoCA (MCI)	**Clinic (3)**	**Community (7)**			
Sensitivity	0.88 (0.76, 0.95)	0.79 (0.68, 0.87)	0.43	0.08 (χ^2 = 25.96, df = 3, *p* = 0.00*).	1.58 (0.21)
Specificity	0.81 (0.69, 0.89)	0.79 (0.71, 0.85)	0.22		1.14 (0.71)

Note. ACE = Addenbrooke's Cognitive Examination-Revised & III; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment (including different scoring systems)

Appendix 10: Likelihood ratio regression based on bivariate models (population)

Test	Population (N)		Between-study heterogeneity (Tau^2)	Generalised Tau^2	Model comparison: χ^2 (*P*-value)
ACE	**China (3)**	**Others (2)**			
Sensitivity	0.90 (0.83, 0.94)	0.85 (0.71, 0.92)	0.16	0.00 (χ^2 = 4.07, df = 3, *p* = 0.25)	0.89 (0.35)
Specificity	0.81 (0.71, 0.88)	0.90 (0.79, 0.96)	0.17		1.98 (0.16)
MMSE	**China (9)**	**Others (10)**			
Sensitivity	0.83 (0.73, 0.89)	0.87 (0.80, 0.92)	0.57	0.41 (χ^2 = 212.64, df = 3, *p* = 0.00*)	0.68 (0.41)
Specificity	0.70 (0.56, 0.81)	0.88 (0.80, 0.93)	0.79		5.74 (0.02*)
MoCA	**China (3)**	**Others (14)**			
Sensitivity	0.86 (0.69, 0.94)	0.88 (0.82, 0.92)	0.69	0.26 (χ^2= 91.10, df = 3, *p* = 0.00*)	0.14 (0.71)
Specificity	0.77 (0.58, 0.89)	0.89 (0.83, 0.93)	0.60		2.65 (0.10)

Note. ACE = Addenbrooke's Cognitive Examination-Revised & III; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment (including different scoring systems)

Appendix 11: Likelihood ratio regression based on bivariate models (subtype)

Test	Subtype (N)		Between-study heterogeneity (Tau^2)	Generalised Tau^2	Model comparison: χ^2 (*P*-value)
MMSE (dementia)	**AD (3)**	**non-AD (5)**			
Sensitivity	0.96 (0.89, 0.98)	0.90 (0.85, 0.94)	0.18	0.02 (χ^2 = 7.04, df = 3, *p* = 0.0708)	2.21 (0.14)
Specificity	0.93 (0.87, 0.96)	0.87 (0.80, 0.92)	0.19		1.99 (0.16)
MMSE (MCI)	**MCI (9)**	**Others (2)**			
Sensitivity	0.80 (0.76, 0.83)	0.60 (0.50, 0.69)	0.03	0.00 (χ^2 = 56.10, df = 3, *p* = 0.00*)	10.23 (0.00*)
Specificity	0.66 (0.52, 0.78)	0.85 (0.63, 0.95)	0.68		2.16 (0.14)
MoCA (dementia)	**AD (3)**	**non-AD (4)**			
Sensitivity	0.93 (0.84, 0.97)	0.92 (0.83, 0.97)	0.48	0.00 (χ^2 = 13.89, df = 3, *p* = 0.0031*)	0.04 (0.84)
Specificity	0.93 (0.82, 0.97)	0.95 (0.88, 0.98)	0.61		0.37 (0.54)

Note. ACE = Addenbrooke's Cognitive Examination-Revised & III; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment (including different scoring systems). AD=Alzheimer's disease; MCI = mild cognitive impairment

Appendix 12: Likelihood ratio regression based on bivariate models (reference standard)

Test	Reference standard (N)		Between-study heterogeneity (Tau^2)	Generalised Tau^2	Model comparison: χ^2 (*P*-value)
MMSE (dementia)	**DSM (6)**	**Others (2)**			
Sensitivity	0.93 (0.88, 0.96)	0.90 (0.78, 0.96)	0.22	0.00 (χ^2 = 9.35, df = 3, *p* = 0.0250*)	0.32 (0.57)
Specificity	0.90 (0.83-0.94)	0.90 (0.79, 0.96)	0.29		0.01 (0.92)
MMSE (MCI)	**Peterson (9)**	**Others (2)**			
Sensitivity	0.76 (0.69, 0.81)	0.80 (0.65, 0.90)	0.19	0.00 (χ^2 = 109.48, df = 3, *p* = 0.0000*)	0.38 (0.54)
Specificity	0.66 (0.51, 0.78)	0.88 (0.62, 0.97)	0.78		2.53 (0.11)
MoCA (MCI)	**Peterson (6)**	**Others (4)**			
Sensitivity	0.85 (0.76, 0.91)	0.78 (0.60, 0.89)	0.48	0.07 (χ^2 = 33.69, df = 3, *p* = 0.00*)	0.88 (0.35)
Specificity	0.76 (0.69, 0.82)	0.86 (0.77, 0.91)	0.14		3.09 (0.08)

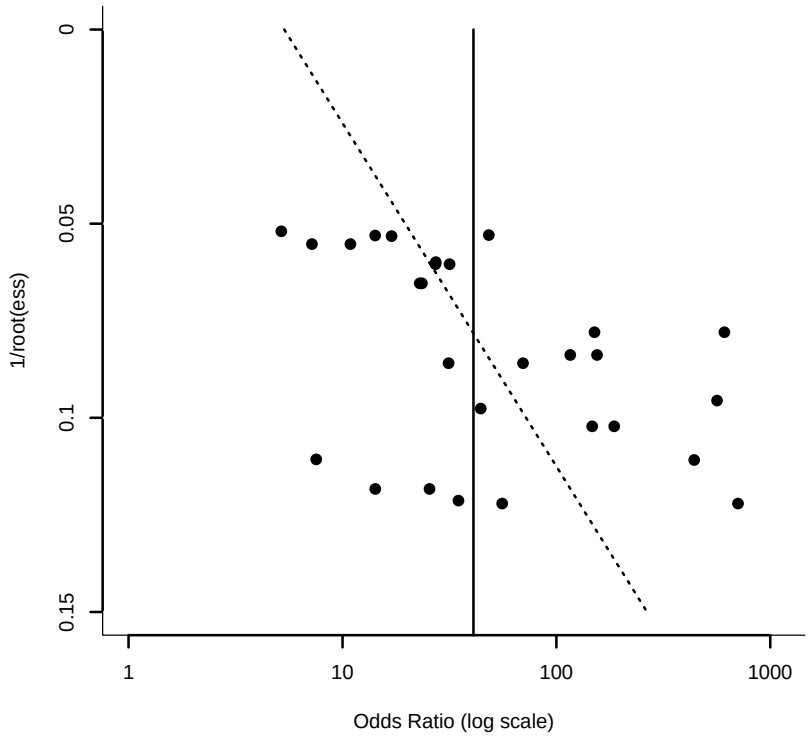
Note. MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment (including different scoring systems); MCI = mild cognitive impairment

Appendix 13: Likelihood ratio regression based on bivariate models (scoring system)

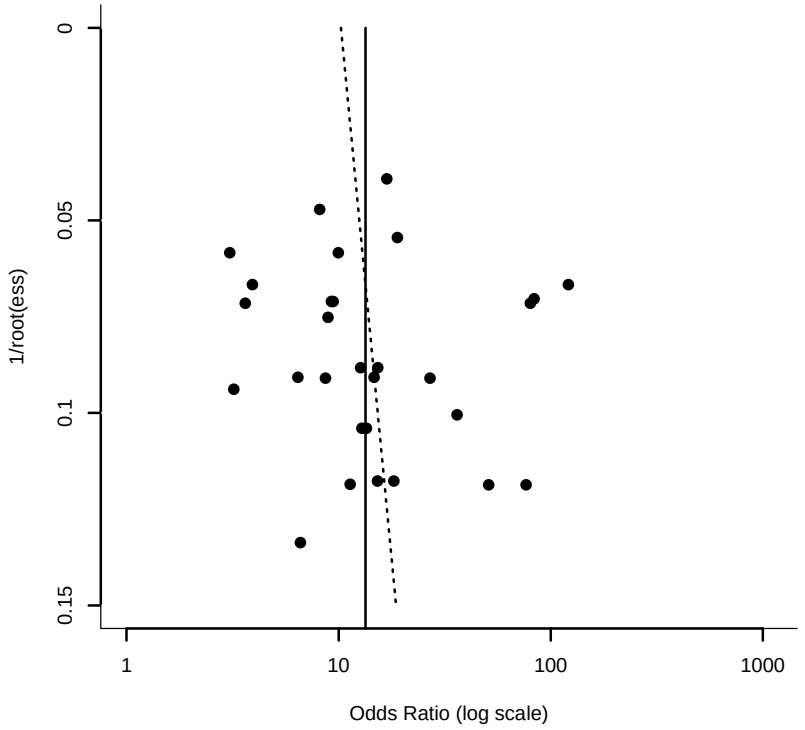
Test	Scoring system (N)		Between-study heterogeneity (Tau^2)	Generalised Tau^2	Model comparison: χ^2 (*P*-value)
MoCA (dementia)	**Original (3)**	**Adjusted (4)**			
Sensitivity	0.96 (0.91, 0.99)	0.88 (0.81, 0.93)	0.16	0.00 (χ^2 = 3.23, df = 3, *p* = 0.3578)	4.38 (0.04*)
Specificity	0.98 (0.95, 1.00)	0.89 (0.84, 0.93)	0.05		9.59 (0.00*)
MoCA (MCI)	**Original (5)**	**Adjusted (5)**			
Sensitivity	0.76 (0.62, 0.86)	0.88 (0.78, 0.93)	0.41	0.08 (χ^2 = 31.83, df = 3, *p* = 0.00*)	2.76 (0.10)
Specificity	0.79 (0.70, 0.86)	0.80 (0.71, 0.87)	0.23		0.03 (0.87)

Note. MoCA = Montreal Cognitive Assessment; MCI = mild cognitive impairment

Appendix 14: Deek's funnel plots



Dementia



MCI

Appendix 15: The MTRQ and MCAR Criteria

Score	MTRQ Definition	MCAR Definition
0	The translation procedure is not mentioned.	The cultural adaptation procedure is not mentioned.
1	The translation procedure is mentioned with no details of the process.	The cultural adaptation procedure is mentioned with no details of the process.
2a	The translation procedure is mentioned in insufficient details for replication.	The cultural adaptation procedure is mentioned in insufficient detail for replication.
2b	The translation procedure is mentioned by referring to another publication that describes the translation process in insufficient detail for replication.	The cultural adaptation procedure is mentioned by referring to another publication that describes the cultural adaptation process in insufficient detail for replication.
3	The translation procedure is only described according to pre-existing guidelines on translating the assessment, with a reference to the guidelines provided.	The cultural adaptation procedure is described only according to pre-existing guidelines on culturally adapting the assessment, with a reference to the guidelines provided.
4a	The translation procedure is described in sufficient detail for replication of the process.	The cultural adaption procedure is described in sufficient details for replication of the process, including reasons for cultural adaption and for the selection and replacement of items in the assessment.
4b	The translation procedure is mentioned by referring to a publication that describes the translation process in sufficient detail for replication, with a reference to that publication.	The cultural adaption procedure is mentioned by referring to a publication that describes the cultural adaption process of that assessment in sufficient detail for replication, including reasons for cultural adaption and for the selection and replacement of items in the assessment, with a reference to that publication.

Note. retrieved from Mirza et al., (2017)

Appendix 16: Published paper from the T-ACE-III study

OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 37 (2022) 692–703

Archives of CLINICAL NEUROPSYCHOLOGY

Validation of the Taiwanese Version of ACE-III (T-ACE-III) to Detect Dementia in a Memory Clinic

Ruan-Ching Yu[1,*,†], Naaheed Mukadam[1], Narinder Kapur[2], Joshua Stott[2], Chaur-Jong Hu[3,4], Chien-Tai Hong[3,4], Cheng-Chang Yang[3,5], Lung Chan[3,4], Li-Kai Huang[3,4], Gill Livingston[1,†]

[1]*Department of Mental Health of Older People, University College London, London, UK*
[2]*Department of Clinical, Educational and Health Psychology, University College London, London, UK*
[3]*Department of Neurology and Dementia Center, Shuang Ho Hospital, Taipei Medical University, New Taipei City, Taiwan*
[4]*Department of Neurology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan*
[5]*Brain and Consciousness Research Center, TMU-Shuang Ho Hospital, New Taipei City, Taiwan*

*Corresponding author at: Department of Mental Health of Older People, University College London, London, UK.
E-mail address: ruan-ching.yu.18@ucl.ac.uk(@naturelovegrace)
[†]Alternate corresponding author: Gill Livingston, g.livingston@ucl.ac.uk

Accepted 6 October 2021

Abstract

Objective: The Addenbrooke's Cognitive Examination III (ACE-III) is a 100-points cognitive test used in detecting dementia in many countries. There has been no validation study of the ACE-III in patients with suspected dementia in a Taiwanese population, where the language is traditional Chinese. We aimed to culturally adapt and validate the ACE-III as a cognitive assessment tool for differentiating between people with and without dementia presenting to healthcare professionals in Taiwan with possible dementia.
Methods: We culturally adapted the ACE-III for Taiwan (T-ACE-III) and tested it with consenting patients with suspected dementia in northern Taiwan who had been through the diagnostic process. We calculated receiver operating characteristic (ROC) curves to test the ability of the T-ACE-III to differentiate between dementia and non-dementia cases using clinician diagnosis as the gold standard. We generated the Youden Index to determine the best cut-off score.
Results: We recruited 90 Taiwanese individuals aged 49–93 years: 24 males and 33 females had dementia and 12 males and 21 females did not. The area under the ROC curve was 0.99 for distinguishing dementia from non-dementia. The T-ACE-III had a sensitivity of 100% and specificity of 78.8% when the cut-off score was 86/87. With a cut-off value of 73/74, the specificity was 100.0%, and sensitivity 89.5%. The highest Youden Index was 0.895, indicating the best overall cut-off point to be 73/74.
Conclusions: The T-ACE-III is an acceptable cognitive test with excellent psychometric properties for discriminating dementia from non-dementia in Taiwanese populations in memory clinic settings.

Keywords: Dementia; Cognitive assessment; Cross-cultural/minority

Introduction

Dementia affects around 5% of people over the age of 60 (Prince et al., 2015) worldwide and 7.19% of those over the age of 65 in Taiwan (Ministry of Health and Welfare in Taiwan, 2020), which has become an aging society particularly in the past 20 years. Studies mostly conducted around two decades ago found that the prevalence of dementia in Taiwan was 3.4% in the age-range of 65–69 years, increasing to 36.9% for the over 90s age range (Sun et al., 2014). A more recent Taiwanese national study between 2011 and 2013, found the data of very mild dementia was less reported and addressed in some early studies

© The Author(s) 2021. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permission@oup.com
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

https://doi.org/10.1093/arclin/acab089 Advance Access publication 3 November 2021

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

(Lin et al., 1998; Liu et al., 1998a; Liu et al., 1996; Liu et al., 1995; Liu et al., 1998b; Liu et al., 1997), indicating possible challenges in diagnosing very mild dementia (Chen, Wang, Ma, & Yang, 2014). There is a pressing need for validating rapid and accurate tests for timely detection of possible dementia among Taiwanese populations.

Early diagnosis of dementia is one of the main aims of the World Health Organization (WHO) for the period from 2017 to 2025 (Alzheimer's disease International, 2019; World Health Organization, 2017). But data collected between 1992 and 2016, found internationally only approximately half of individuals who met dementia diagnostic criteria were formally diagnosed (Amjad, Roth, Samus, Yasar, & Wolff, 2016; Connolly, Gaehl, Martin, Morris, & Purandare, 2011; Lang et al., 2017; Savva & Arthur, 2015). The project of the Ministry of Health and Welfare in Taiwan aimed to increase the diagnosis rate of dementia from 30% to 50% by 2020 and to 70% by 2025 (Ministry of Health and Welfare in Taiwan, 2020). Clinical assessments to precisely identify those individuals can be time-consuming but if only brief cognitive assessments are used, suspected dementia is often misclassified, because scores differ across factors such as age, education, and ethnicity (Ranson et al., 2019). And, screeners often do no adjust for these demographic factors. Therefore, it is important to select an appropriate brief assessment in order to achieve accurate timely detection of dementia.

Taiwan is a mix of ethnic groups, such as Southern Min, Hakka and Taiwanese Aboriginal, and many Chinese dialects are spoken. One of the most commonly spoken Chinese dialects is Taiwanese Hokkien (also called Taiwanese), spoken by approximately 70% of the population. Taiwanese speakers tend to have a low educational level, with a cut-off stage of senior high school (from 16 to 18 years old). A study of 252 patients at a hospital in South Taiwan found that 95.4% of the older group used Taiwanese and all nonliterate people spoke Taiwanese. They also discovered that people who are older are more likely to use Taiwanese Hokkien as their main language (Tsao & Lu, 1999).

In Taiwan, the Clinical Dementia Rating (CDR), is used for both staging and diagnosis (Morris, 1993) as it is internationally. The Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975), is the most commonly used cognitive test which, on average, takes 10–15 min to administer, but does not test memory and executive functioning extensively, has ceiling effects, a major influence of education, and a copyright-protection issue (Carnero-Pardo, 2014). The Cognitive Abilities Screening Instrument (CASI) (Lin et al., 2002), which takes around 15–20 minutes, but up to 30 minutes in patients with dementia (Teng et al., 1994), gives particularly low scores in those who can only speak Taiwanese, after accounting for age and education (Chen, 1996). So far, there has been no validated short (around 20 min or less) assessment that covers a wide range of cognitive domains to detect dementia in the Taiwanese population.

The ACE-III assesses cognitive domains including attention, orientation, fluency, memory, language, visual perceptual function, and visuospatial skills. It takes a short time (<20 min) to administer. There were simplified Chinese, Japanese, and Korean language versions on the official website but not a traditional Chinese version for Taiwan. In China, it has high validity in detecting all types of dementia (sensitivity: 94.0%, 91.1%; specificity: 83.0%, 83.1%; cut-off score: 74/75, 83) (Li, Yang, Yin, Yu, & Ye, 2019; Wang et al., 2017) and is helpful in distinguishing subtypes of dementia in other countries (Elamin, Holloway, Bak, & Pal, 2016; Hsieh, Schubert, Hoon, Mioshi, & Hodges, 2013; Matias-Guiu et al., 2017).

The primary aim of this study was to culturally adapt and validate the Taiwanese version of ACE-III (T-ACE-III) in patients with dementia in clinical settings in Taiwan. The secondary aim was to explore the acceptability of other supplementary tests in this population.

Method

This prospective study was approved by the UCL Research Ethics Committee (REC) in the UK (16845/001) and the Office of Human Research of Taipei Medical University in Taiwan (NO:N202004089), including Data Protection Registration (NO:Z6364106/2020/01/48) in the UK. The data collection process was carried out through convenience sampling in Taiwan from June 2020 to October 2020.

We obtained permission to translate, adapt, and use the cognitive tests from the original authors or the relevant institutions.

Participants

Eligible participants were people, aged 45–100, referred to a neurological clinic in the north of Taiwan in the previous 6 months with memory problems or suspected dementia. Individuals who had in the past been diagnosed with severe psychiatric disorders, or with visual, hearing, or motor impairments which significantly interfered with comprehension or execution of cognitive tasks were excluded.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Dementia Diagnosis

The four neurologists based in the clinic diagnosed the participants as having or not having dementia according to the diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) prior to their recruitment for the study (American Psychiatric Association, 2013) using their medical history, daily functioning, reported cognitive symptoms, laboratory results, and neuroimaging findings. Clinical observation on significant cognitive decline in one or more cognitive domains and the interference with independence in everyday activities were judged by the clinicians at consultation. The study only had access to examine their CDR, MMSE, and CASI scores, but not other tests. They also classified dementia subtypes, based on the International Classification of Diseases, Tenth Revision (ICD-10). Neuropsychological assessments were conducted by three psychologists.

Cultural Adaptation

The term China we use here refers to the People's Republic of China, excluding the Special Administrative Regions of Hong Kong and Macau. Mandarin Chinese is officially spoken in Taiwan and China, whereas the written form of Simplified Chinese is used in China and Singapore and Traditional Chinese is used in Taiwan, Hong Kong, and Macau. The most significant difference between the two written forms of Chinese is the form of the individual characters. Simplified Chinese was implemented by the Chinese government between 1956 and 1964 in order to improve literacy but is often not understood in Taiwan especially by older people. We used the Cross-Cultural Adaptation Process (Beaton, Bombardier, Guillemin, & Ferraz, 2000) to guide the appropriate procedures for cultural adaptation of psychometric tests for the Taiwan population. This is used when assessments in one language are used in other populations with the same language, but differing cultures or language characters. So, for example, the terms used for the leaders of state differ (chairman in China and president in Taiwan). The form of characters (simplified Chinese in China, traditional Chinese in Taiwan) and the culture and history between two populations also differ. In particular, the influence of different Chinese dialects is relevant when culturally adapting neuropsychological assessments (Chen, 1996). Further details are outlined below.

Measures

The ACE-III takes 15–20 min and has a maximum score of 100—a higher score indicates better cognitive function. R-CY, a researcher with a clinical background in dementia who is bilingual and literate in Traditional and Simplified Chinese and English, adapted the Simplified Chinese version of the ACE-III into Traditional Chinese. C-CY who is a clinical neuropsychologist and linguistics expert is also a bilingual researcher, read the draft adapted version and discussed the changes with R-CY. It was adapted for Taiwanese culture after a series of discussions with two neurologists specializing in dementia, C-JH and C-TH, resulting in a consensus. Lastly, after pretesting in a small sample of Taiwanese colleagues and adapting it in line with their feedback, we generated the final version of the Taiwanese ACE-III (T-ACE-III). The main changes include replacing phonemic fluency with semantic fluency (from generating Chinese words starting with the word "car" to "naming vegetables and fruits"), using local Taiwanese name and addresses for memory, using pictures frequently seen in daily life in Taiwan (from pencil, panda, harp, barrel to spoon, kangaroo, drum, bulb) instead of in China for language and understanding, using traditional Chinese words for reading in language, and replacing English letters with numbers for visuospatial abilities. More importantly, alterations were made to language subtest items (asking them to repeat the phrases or proverbs) for Taiwanese language users and Taiwanese language was used throughout the assessments. The validated T-ACE-III is now available online (The University of Sydney, 2021).

In addition to the T-ACE-III, we administered several neuropsychological tests for acceptability. These included: *The Brixton Spatial Anticipation Test from Hayling and Brixton tests* (Burgess & Shallice, 1997) for assessing cognitive flexibility/set shifting, one aspect of executive function—it takes 15–20 min and is a visual test where people follow a rule to position a blue circle from two rows of five circles numbered 1–10 in each page, evaluating the ability to detect a rule, to follow it, and to switch to a new rule. It does not require a verbal or complex motor response by the patient and has no time restriction. The maximum number of errors is 54—higher scores mean worse performance. We directly translated the brief instructions from English to Chinese. R-CY who used this test discussed with N-K (a neuropsychology expert) to ensure the procedure was in line with guidance.

Processing Speed Index from the Wechsler Adult Intelligence Scale IV—(WAIS-IV-PSI) (Wechsler, 2008) measures processing speed through a series of tasks of symbol search, cancelation and coding. It takes less than 10 min. The Taiwanese version of the WAIS-IV was released in 2015 (Wechsler, 2015a; Wechsler, 2015b), and we used it, without it requiring changes for cultural adaptation and translation.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

The Coin-in-the Hand (CIH) test (Kapur, 1994), a two-choice memory test, takes 5 min and measures cognitive effort. It has been used for differentiating malingerers from those with true memory impairment (Schroeder, Peck, Buddin Jr., Heinrichs, & Baade, 2012) including in a pilot study in Taiwan (Yeh et al., 2018), which achieved a high AUC of 0.79 with a cut-off value of 8.27. R-CY translated the administration instructions from English to Chinese for this study.

Apart from the assessments above, we also collected the demographic information including age, sex, language spoken, years of education, the last employment (types and position: managerial/ nonmanagerial), and medical history from the case report form.

Power Calculation

We calculated the numbers of participants needed in each group through "easyROC" (http://www.biosoft.hacettepe.edu.tr/easyROC/). With setting power 90%, type I error 0.05, area under the ROC 0.7, and allocation ratio 0.5, the numbers of participants required in dementia group and non-dementia group were 48 and 24, respectively. For better understanding the influence of languages used on the T-ACE-III scores, we recruited nine more participants in each group.

Statistical Methods

We analyzed data using SPSS (version 26). We first described the demographic characteristics; sex, age, education, language, and raw cognitive test scores in the dementia and non-dementia groups. We also described cognitive tests results, including the mean with standard deviation (SD) for parametric data and median with interquartile range (IQR) for nonparametrically distributed data.

We recorded the time taken for assessments from participant 47 onwards when R-CY experienced in the interview.

The primary analysis was area under the curve (AUC) with receiver operating characteristic (ROC) curve analysis, which we used to calculate the diagnostic accuracy for the T-ACE-III to distinguish between dementia and non-dementia groups. To consider the optimal cut-off point, providing the best measure of sensitivity and specificity combined, we used the Youden Index, which can range between 0 and 1, with 1 indicating no false positives or false negatives. We estimated the reliability using the domain-total internal consistency of the T-ACE-III (Cronbach's α-coefficient). We also examined the convergent validity by calculating correlation using a two-tailed Spearman's test between the scores of the T-ACE-III and the tests previously administered in the neurology clinic (CDR, MMSE, CASI).

To examine the influence of demographic factors on the T-ACE-III scores, we used linear regression analysis, with the T-ACE-III scores as the dependent variable and the independent variables of age, years of education, sex, SES (managerial/ nonmanagerial), and language. We then considered independent predictors of diagnosis of dementia by entering demographic factors (age, years of education, Socioeconomic Status (SES), and sex) and T-ACE-III scores into logistic regression analysis with dementia diagnosis as the outcome. In a post hoc analysis suggested by referees we also explored the relationship of education and age on language (speaking only Taiwanese compared to speaking both Mandarin and Taiwanese) as the outcome in logistic regression analysis. $p < .05$ was considered statistically significant. Lastly, we examined the convergent validity of the supplementary tests with the T-ACE-III using a two-tailed Spearman's correlation analysis.

Results

Demographic and Clinical Profile

We invited 110 participants to take part in this study, of whom 91 (82.7%) consented and 90 (81.8%) completed the neuropsychological assessments. Reasons for refusal included having appointments with other doctors, impaired attention due to bad sleep, and not wanting to take part in this research. There were 57 participants with complete data in the dementia group, including 21 patients with Alzheimer's disease (AD), 15 with vascular dementia (VaD), 6 with Parkinson's disease dementia (PDD), 3 with frontotemporal dementia (FTD), 1 with dementia with Lewy bodies (DLB), 2 patients with unspecified dementia, 9 with mixed-type dementia (3 VaD + PDD, 4AD+VaD, 1AD+FTD, 1AD+Trauma), and 33 in the non-dementia group. All participants completed the T-ACE-III. Overall, 35 (61.4%) participants in the dementia and 32 (97.0%) in the non-dementia groups completed WAIS-IV symbol search, 33 (57.9%) with dementia and 32 (97.0%) without completed WAIS-IV symbol coding, 38 (66.7%) with dementia and 33 (100.0%) without completed Brixton spatial anticipation test, 56 (98.2%) with dementia and 33 (100.0%) without completed the CIH test.

The demographic information of the participant groups, and their scores on neuropsychological assessments and the short cognitive battery are shown in Table 1. Those who had been diagnosed with dementia were older (74.2 vs. 66.4), and with less

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Table 1. Demographic and neuropsychological characteristics of participant groups

Variable (number)	Dementia($N = 57$)	Non-dementia ($N = 33$)
Male	24 (42.1%)	12 (36.4%)
Female	33 (57.9%)	21 (63.6%)
Education, years*	7.70 ± 4.4	11.24 ± 4.0
Illiterate	6 (10.5%)	1 (3.0%)
Elementary school	25 (43.9%)	5 (15.2%)
Junior high school	9 (15.8%)	6 (18.2%)
Senior high school	12 (21.1%)	11 (33.3%)
College	5 (8.8%)	9 (27.3%)
Higher education	0 (0.0%)	1 (3.0%)
Age, years*	74.2 ± 8.4	66.4 ± 7.5
<60	5 (8.8%)	7 (21.2%)
61–70	13 (22.8%)	17 (51.5%)
71–80	23 (40.4%)	7 (21.2%)
81–90	15 (26.3%)	2 (6.1%)
>90	1 (1.8%)	0 (0.0%)
Language used		
Taiwanese only	14 (24.6%)	0 (0.0%)
Chinese only or Chinese and Taiwanese	43 (75.4%)	33 (100.0%)
CDR 0	0 (0.0%)	28 (84.9%)
0.5	36 (63.2%)	5 (15.2%)
1	15 (26.3%)	0 (0.0%)
2	6 (10.5%)	0 (0.0%)
MMSE	18.4 ± 5.2	27.7 ± 1.9
CASI	60.8 ± 16.3	91.2 ± 4.9
Total T-ACE-III scores	55.0 ± 15.8	90.8 ± 6.6
Orientation/attention*	11.3 ± 4.0	17.4 ± 1.1
Memory*	11.7 ± 5.3	23.5 ± 2.3
Verbal fluency*	4.6 ± 3.2	10.5 ± 2.2
Language*	16.6 ± 5.1	24.2 ± 2.3
Visuospatial*	10.8 ± 4.0	15.3 ± 1.3
WAIS-IV-PSI-Symbol search (35/32)*	6.6 ± 2.4	10.7 ± 3.4
WAIS-IV-PSI-Digit symbol coding (33/32)*	6.9 ± 3.0	9.5 ± 2.6
Brixton spatial anticipation test* (38/33)	2.2 ± 1.6	4.3 ± 2.2
The Coin-in-the-hand test* (56/33)	9.7 ± 0.7	9.9 ± 0.3

CDR, Clinical Dementia Rating Scale; MMSE, mini-mental state examination; T-ACE-III, Taiwanese version of Addenbrooke's Cognitive Examination III; WAIS-IV-PSI, Processing Speed subtests of the Wechsler Adult Intelligence Scale-IV.
*Represents the results presented with Mean ± SD.

education than those in the non-dementia group (7.7 and 11.2 years of education). Fourteen (24.6%) participants in the dementia group could speak only Taiwanese language. But all of those in the non-dementia group could speak at least Mandarin Chinese (15.2%) or both (84.8%).

Time to Administer on Average

We measured time taken for testing in 43 participants, of whom 16 did not complete the supplementary tests. The T-ACE-III took a mean (SD) of 24.0 ± 6.1 and 16.3 ± 4.6 min in the dementia group and non-dementia group, respectively. With the supplemental tests, the mean total time taken was 41.1 ± 4.8 and 32.0 ± 5.3 min to administer in the dementia group and non-dementia group, respectively.

The T-ACE-III Scores for Discriminating Dementia from Non-Dementia

Figure 1 shows the ROC analysis. The area under the ROC curve (AUC) was 0.990 for the T-ACE-III ($p < .001$, 95% confidence interval (CI) = 0.977–1.000). As shown in Table 2, when the cut-off score was 86/87, the T-ACE-III had a specificity of 0.788 and a sensitivity of 1, with a value of 0.788 on the Youden index; when the cut-off score was 73/74, the T-ACE-III had a specificity of 1 and a sensitivity of 0.895, with the highest value of 0.895 of Youden index, indicating the best cut-off point.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

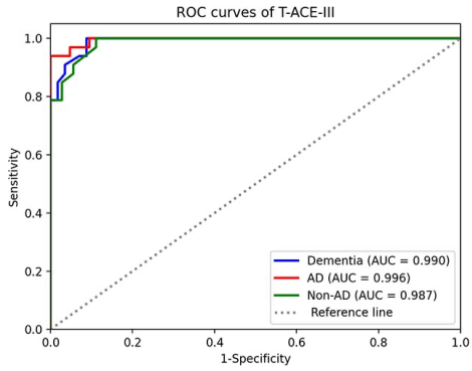


Fig. 1. ROC curves of T-ACE-III.

Table 2. Cut-off scores for the T-ACE-III in differentiating dementia from non-dementia

Cut-off scores	Sensitivity	Specificity	Youden Index
73/74	0.895	1.000	0.895
74/75	0.895	0.970	0.865
75/76	0.912	0.970	0.882
76/77	0.930	0.939	0.869
77/78	0.947	0.939	0.886
80	0.965	0.909	0.874
82/83	0.965	0.879	0.844
83/84	0.982	0.848	0.830
84/85	0.982	0.818	0.800
85/86	0.982	0.788	0.770
86/87	1.000	0.788	0.788

The T-ACE-III Scores for Discriminating AD and Other Dementia Subtypes from Non-Dementia

The AUCs for distinguishing either AD or other dementia subtypes from the non-dementia group were calculated. As shown in Figure 1, the AUCs of T-ACE-III for detecting AD and other dementia subtypes versus no dementia, were 0.996 ($p < .001$, 95% CI = 0.986–1.000) and 0.987 ($p < .001$, 95% CI = 0.969–1.000), respectively, which suggested that the high diagnostic accuracy was not influenced by the type of dementia.

Reliability and Convergent Validity of the T-ACE-III

The internal consistency of the T-ACE-III was 0.895. There was negative correlation between CDR scores and total T-ACE-III scores ($r = -0.8$, $p < .01$) and a correlation between MMSE and CASI scores from psychologists and total T-ACE-III scores (both $r = 0.9$, $p < .01$).

Association of Demographic Factors with the T-ACE-III Scores in Linear Regression

Using linear regression with the T-ACE-III scores as the dependent variable, (see Table 3), we found language used (only Taiwanese compared to both Mandarin and Taiwanese, $p = .005$, 95% CIS for B = −30.868 to −5.856) and every additional year of education ($p = .009$, 95% CIS for B = 0.404–2.760) were significantly associated with the T-ACE-III scores. However, age (for every year older, $p = .213$, 95% CIS for B = −0.810 to −0.183), sex (male compared to female, $p = .292$, 95% CIS for B = −13.406 to −4.085), and employment divided into managerial/ nonmanagerial jobs ($p = .614$, 95% CIS for B = −6.864 to 11.550) were not. The variance inflation factors (VIFs) were from 1.157 to 1.798 (lower than 10), indicating there was no multicollinearity existed among the independent variables (education years, age, sex, language spoken, and SES). The model

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Table 3. Association of demographic factors with T-ACE-III scores and diagnosis of dementia

Dependent variables	Independent variables						
	Coefficients	T-ACE-III (higher)	Education (increasing years)	Age (higher)	Sex (male)	SES (managerial/ nonmanagerial)	Language (Taiwanese)
Linear regression (T-ACE-III scores)	Standardized Coefficients β	-	0.329	−0.127	−0.105	0.052	−0.300
	95% CIS for B	-	0.404–2.760	−0.810-0.183	−13.406-4.085	−6.864-11.550	−30.868 – −5.856
	T	-	2.672	−1.256	−1.060	0.506	−2.921
	p	-	.009	.213	.292	.614	.005
	R^2	0.319					
Logistic regression (outcome diagnosis of dementia)	B	−0.727	0.236	−0.163	−0.945	3.259	-
	Exp(B)	0.483	1.267	0.850	0.389	26.031	-
	95% CIS for Exp(B)	0.260–0.898	0.808–1.985	0.679–1.063	0.005–28.690	0.140–4833.824	-
	p	.021	.302	.155	.667	.221	-
	NagelkerkeR^2	0.927					
Logistic regression (language used—Taiwanese only)	B	-	−0.35	0.061	-	-	
	Exp(B)	-	0.704	1.063	-	-	
	95% CIS for Exp(B)	-	0.586–0.847	0.990–1.141	-	-	
	p	-	<.001	.094	-	-	
	NagelkerkeR^2	-	0.357	0.058	-	-	

T-ACE-III, Taiwanese version of Addenbrooke's Cognitive Examination III. SES = socioeconomic status We could not include language in logistic regression as there were no Taiwanese speakers in the non-dementia groups.

accounted for 31.9% of the variance in the T-ACE-III scores overall model (F = 7.6, $p < .001$). Further analysis on the language used (Taiwanese only) through logistic regression, the results confirmed that every decreasing year of education ($p < .001$, 95% CIS for Exp(B) = 0.586–0.847) was significantly associated. The model accounted for 35.7% of the variance in the overall model ($p < .001$). Although age was not ($p = .935$, 95% CIS for Exp(B) = 0.990–1.141). The model accounted for 5.8% of the variance in the overall model ($p < .001$).

Association of Demographic Factors with Diagnosis of Dementia on Logistic Regression

In the logistic regression model, (see Table 3), lower T-ACE-III ($p = .021$, 95% CIS for Exp(B) = 0.260–0.898) scores were associated with diagnosis of dementia. With T-ACE-III scores in the model, years of education ($p = .302$, 95% CIS for Exp(B) = 0.808–1.985), age ($p = .155$, 95% CIS for Exp(B) = 0.679–1.063), sex ($p = .667$, 95% CIS for Exp(B) = 0.005–28.690) and managerial/nonmanagerial jobs ($p = .221$, 95% CIS for Exp(B) = 0.140–4833.824) were not significantly associated. The model accounted for 92.7% of the variance ($p < .001$). We were unable to add language to the model as no one who spoke only Taiwanese was not diagnosed of dementia.

Correlation between the T-ACE-III and the Supplemental Tests

As shown in Table 4, there was a significant correlation between WAIS-PIS scores (including symbol search and digit symbol coding) and the T-ACE-III scores (including subdomains), and a significant correlation between Brixton spatial anticipation test scores and the T-ACE-III scores (The CIH scores were only slightly correlated with the T-ACE-III scores, including subdomains). The results suggested WAIS-PIS scores and Brixton spatial anticipation test scores increased when the T-ACE-III scores increased, whereas an increase in the CIH scores was unable to reflect the tendency of the T-ACE-III scores and its subdomains.

Discussion

To our knowledge for the first time, we have culturally adapted and validated the T-ACE-III in a Taiwan population, and established cut-off scores in differentiating patients with dementia from healthy participants. The T-ACE-III had excellent

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Table 4. Correlation between T-ACE-III and the supplemental tests

	WAIS-symbol search	WAIS-digit symbol coding	Brixton spatial anticipation test	Coin-in-the-hand test
T-ACE-III total scores	0.610**	0.511**	0.563**	0.265*
Attention	0.549**	0.484**	0.533**	0.247*
Language	0.568**	0.456**	0.539**	0.137
Memory	0.511**	0.349**	0.542**	0.248*
Verbal fluency	0.416**	0.363**	0.351**	0.339**
Visuospatial	0.656**	0.717**	0.488**	0.164

*$p < .05$.
**$p < .01$.
T-ACE-III, Taiwanese version of Addenbrooke's Cognitive Examination III.

accuracy in distinguishing dementia from non-dementia in a memory clinic population with a higher ROC AUC than currently used tests (Chang et al., 2012; Chen et al., 2018; Tsai et al., 2016; Wang et al., 2013). The T-ACE-III was psychometrically robust with the total T-ACE-III scores showing a high correlation with the scores of current assessments used in differentiating dementia from those who did not have dementia in Taiwan (CDR, MMSE, CASI), pointing to good convergent validity and good internal consistency (Cortina, 1993). It had a very high acceptability and completion rate and took a relatively short time to administer compared to current assessments (CASI, CDR, and Neuropsychiatric Inventory) for diagnosing dementia in Taiwan, which usually take around an hour for each patient in a Taiwan clinic (personal communication, C-CY). Linear regression showed that sociodemographic characteristics did not account for most of the variance in the cognitive test scores.

Our secondary aim was to explore the acceptability (completion rate and administration time) of supplementary neuropsychological tests. The completion rates of Symbol Search and Coding subtests of the WAIS-IV and of the Brixton Spatial Anticipation Test were only between 57.9% and 66.7%, and so had low acceptability in the dementia group. We do not know why this was, but some patients may have been tired since the T-ACE-III was always completed first, some struggled with the physical demands of performing the WAIS subtests and others with understanding the Brixton test.

We found in multiple linear regression analysis that years of education and spoken languages of only Taiwanese were significant predictors of the T-ACE-III scores, but that there was no significant effect of age on the total scores. We found that language spoken (Taiwanese only compared to those who spoke both Mandarin and Taiwanese) was a significant predictor of ACE scores and this may be because those who only spoke Taiwanese had less education, echoing previous research that indicated education levels are negatively correlated with the choice of Taiwanese language (Tsao & Lu, 1999). People who spoke only Taiwanese had less education and this was possibly a risk factor for dementia and may have led to lower scores on the T-ACE-III. However, the clinical diagnosis was made independently of T-ACE-III scores. In terms of the potential predictors on the diagnosis of dementia, logistic regression analysis found that demographic factors had no influence over and above T-ACE-III. In summary, in the cultural context of Taiwan, although demographic factors of language and education years have some influence on the T-ACE-III scores, it is still a powerful diagnostic tool in detecting patients with dementia, with T-ACE-III scores contributing more to diagnostic classification than other factors.

A previous study validated the traditional Chinese version of ACE-III in community settings rather than a clinical setting for detecting MCI in Taiwan, and their translation did not follow cultural adaptation guidelines (Lin, Hsu, Wang, & Ouyang, 2019). However, our own study followed specific guidelines for translation and cultural adaptation (Beaton et al., 2000), incorporating opinions from a bilingual linguist, a bilingual researcher, two neurologists, one psychiatrist, and one clinician.

Most participants completed the CIH test and most participants passed it even if they had dementia. This confirms its validity as a test of cognitive effort and its acceptability (Schroeder et al., 2012; Yeh et al., 2018) in a non-Western population such as that sampled in Taiwan.

There are some limitations to the current study. Firstly, we recruited a consecutive sample of people to reduce selection bias and found that Taiwanese language users in the current study were all in the dementia group. Future research can further confirm the influence of language used on the T-ACE-III performance if necessary, enriching the sample and recruiting Taiwanese users in each group. Also, there were six participants in the dementia and one in the non-dementia groups who were illiterate. They may have led to problems in tasks with reading and writing elements which were not due to dementia, the cut-off values after adjusting the T-ACE-III scores is shown in appendix. We recommend that future research uses a larger sample size to explore different cut-off values of the T-ACE-III for participants with different educational levels and levels of literacy, so as to provide more tailored cut-off scores relating to educational and literacy needs, and thus increase the diagnostic accuracy. Secondly, we calculated internal consistency but no other forms of reliability in the analysis, such as test–retest reliability. Similarly, we did not examine discriminant validity. Thirdly, one question in the attention domain (orientation) has been replaced with the MMSE

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

item due to the instruction of the simplified Chinese version not being clearly translated and tending to misguide the examiners during the assessing procedure. Fourth, the study could not be double-blind as the patients had been given a diagnosis as part of the routine clinical work-up. Both the patient and the researcher were aware of the diagnosis when administering psychometric tests but at that time the cutpoint for the T-ACE-III was not known. Lastly, under the threat of COVID-19, Taiwanese residents were required to wear a mask in all indoor spaces, such as hospitals and public transport. During the assessment procedures, both participants and examiners were wearing a mask, which may potentially have affected the understanding of instructions or conversations, especially for specific items, such as language tasks (asking participants to repeat a sentence, asking them to complete three actions) and attention tasks (asking participants to repeat three words). As a result, this may have contributed to a degree of variability in the T-ACE-III test performance, as well as the average time taken by individual participant to complete the T-ACE-III, which was slightly longer than 20 min in the dementia group. Apart from that, personal and environmental factors during the assessment session may also have contributed to variability in time taken to complete the T-ACE-III. However, since the T-ACE-III showed a high level of diagnostic accuracy for dementia, which was superior to those of MMSE and CASI and took shorter time to administer than CASI, the T-ACE-III would appear to be a promising screening tool for detecting dementia in a Taiwan population.

In conclusion, the T-ACE-III is a valid, reliable, and effective tool to diagnose dementia with a very high AUC in ROC analysis and is relatively quick compared to current test batteries. It has greater validity in differentiating dementia from non-dementia compared to currently used tests so it may prove to be the test of choice in clinical settings in Taiwan in view of its ease of administration and accuracy.

Funding

The Ministry of Education in Taiwan (to R.-C.Y.). Alzheimer's Society Senior Research Fellowship (to N.M.). By NIHR North Thames ARC and as an NIHR senior investigator (to G.L.). UCLH National Institute for Health Research (NIHR) Biomedical Research Centre (to N.M, J.S., and G.L).

Conflict of Interest

None declared.

Authors' Contributions

Ruan-Ching Yu, Naaheed Mukadam, Narinder Kapur, Joshua Stott, Gill Livingston did the conceptualization and design, analysis and interpretation of data, and preparation of manuscript. Chaur-Jong Hu, Chien-Tai Hong, Cheng-Chang Yang, Lung Chan, Li-Kai Huang did the reading and commenting on manuscript, provision of suggestion on literature review and discussion.

Sponsor's Role

University College London (UCL).

References

Alzheimer's disease International. (2019). *World Alzheimer report 2019: Attitudes to dementia*. London: Alzheimer's Disease International. Available at: https://www.alzint.org/resource/world-alzheimer-report-2019/. Accessed March 8, 2021.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, Fifth Edition*. Arlington, VA: American Psychiatric Association.

Amjad, H., Roth, D. L., Samus, Q. M., Yasar, S., & Wolff, J. L. (2016). Potentially unsafe activities and living conditions of older adults with dementia. *Journal of the American Geriatrics Society*, *64*(6), 1223–1232. https://doi.org/10.1111/jgs.14164.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(*24*), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014.

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton tests*. Bury St. Edmunds, England: 1997: Thames Valley test company.

Carnero-Pardo, C. (2014). Should the mini-mental state examination be retired? *Neurología (English Edition)*, *29*(*8*), 473–481. https://doi.org/10.1016/j.nrleng.2013.07.005.

Chang, Y. T., Chang, C. C., Lin, H. S., Huang, C. W., Chang, W. N., Lui, C. C. et al. (2012). Montreal cognitive assessment in assessing clinical severity and white matter hyperintensity in Alzheimer's disease with normal control comparison. *Acta Neurol Taiwan*, *21*(2), 64–73.

Chen, J. Y. (1996). Cognitive abilities screening instrument Chinese version 2.0 (CASI C-2.0): Report of additional psychometric properties and a possible dialectal bias. *Chinese Psychiatry*, *10*(4), 325–333.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Chen, C. H., Wang, L. C., Ma, T. C., & Yang, Y. H. (2014). A walk-in screening of dementia in the general population in Taiwan. *The Scientific World Journal*, *2014*, 1–5. https://doi.org/10.1155/2014/243738.

Chen, S. F., Liu, M. H., Chen, N. C., Horng, H. D., Tsao, W. L., Chang, C. C. et al. (2018). Educational effects on ascertain dementia 8-item informant questionnaire to detect dementia in the Taiwanese population. *International Psychogeriatrics*, *30*(*8*), 1189–1197. https://doi.org/10.1017/S1041610217002733.

Connolly, A., Gaehl, E., Martin, H., Morris, J., & Purandare, N. (2011). Underdiagnosis of dementia in primary care: Variations in the observed prevalence and comparisons to the expected prevalence. *Aging & Mental Health*, *15*(*8*), 978–984. https://doi.org/10.1080/13607863.2011.596805.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(*1*), 98–104. https://doi.org/10.1037/0021-9010.78.1.98.

Elamin, M., Holloway, G., Bak, T. H., & Pal, S. (2016). The utility of the Addenbrooke's cognitive examination version three in early-onset dementia. *Dementia and Geriatric Cognitive Disorders*, *41*(*1-2*), 9–15. https://doi.org/10.1159/000439248.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(*3*), 189–198.

Hsieh, S., Schubert, S., Hoon, C., Mioshi, E., & Hodges, J. R. (2013). Validation of the Addenbrooke's cognitive examination III in frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *36*(*3–4*), 242–250. https://doi.org/10.1159/000351671.

Kapur, N. (1994). The coin-in-the-hand test: A new "bed-side" test for the detection of malingering in patients with suspected memory disorder. *Journal of Neurology, Neurosurgery, and Psychiatry*, *57*(*3*), 385–386. https://doi.org/10.1136/jnnp.57.3.385.

Lang, L., Clifford, A., Wei, L., Zhang, D., Leung, D., Augustine, G. et al. (2017). Prevalence and determinants of undetected dementia in the community: A systematic literature review and a meta-analysis. *BMJ Open*, *7*, e011146. https://doi.org/10.1136/bmjopen-2016-011146.

Li, X., Yang, L., Yin, J., Yu, N., & Ye, F. (2019). Validation study of the Chinese version of Addenbrooke's cognitive examination III for diagnosing mild cognitive impairment and mild dementia. *Journal of Clinical Neurology*, *15*(*3*), 313–320. https://doi.org/10.3988/jcn.2019.15.3.313.

Lin, R. T., Lai, C. L., Tai, C. T., Liu, C. K., Yen, Y. Y., & Howng, S. L. (1998). Prevalence and subtypes of dementia in southern Taiwan: Impact of age, sex, education, and urbanization. *Journal of the Neurological Sciences*, *160*(*1*), 67–75.

Liu, C. K., Lai, C. L., Tai, C. T., Lin, R. T., Yen, Y. Y., & Howng, S. L. (1998a). Incidence and subtypes of dementia in southern Taiwan impact of socio-demographic factors. *Neurology*, *50*(*6*), 1572–1579.

Liu, C. K., Lin, R. T., Chen, Y. F., Tai, C. T., Yen, Y. Y., & Howng, S. L. (1996). Prevalence of dementia in an urban area in Taiwan. *Journal of the Formosan Medical Association Taiwan yi zhi*, *95*(*10*), 762–768.

Liu, H. C., Lin, K. N., Teng, E. L., Wang, S. J., Fuh, J. L., Guo, N. W. et al. (1995). Prevalence and subtypes of dementia in Taiwan: A community survey of 5297 individuals. *Journal of the American Geriatrics Society*, *43*(*2*), 144–149.

Liu, H. C., Fuh, J. L., Wang, S. J., Liu, C. Y., Larson, E. B., Lin, K. N. et al. (1998b). Prevalence and subtypes of dementia in a rural Chinese population. *Alzheimer Disease and Associated Disorders*, *12*(*3*), 127–139.

Liu, H. C., Wang, S. J., Fuh, J. L., Liu, C. Y., Lin, K. P., Lin, C. H. et al. (1997). The kinmen neurological disorders survey (KINDS): A study of a Chinese population. *Neuroepidemiology*, *16*(*2*), 60–68.

Lin, J. Y., Hsu, H. C., Wang, J. Y., & Ouyang, W. C. (2019). Comparison of the Taiwanese versions of the Addenbrooke's cognitive examination-III (ACE-III), mini-mental state examination (MMSE), and Montreal cognitive assessment (MOCA) for screening mild cognitive impairment among older Taiwanese people. *International Journal of Gerontology*, *13*(2), 162–166. https://doi.org/10.6890/IJGE.201906_13(2).0013.

Lin, K. N., Wang, P. N., Liu, C. Y., Chen, W. T., Lee, Y. C., & Liu, H. C. (2002). Cutoff scores of the cognitive abilities screening instrument, Chinese version in screening of dementia. *Dementia and Geriatric Cognitive Disorders*, *14*(*4*), 176–182. https://doi.org/10.1159/000066024.

Matias-Guiu, J. A., Cortés-Martínez, A., Valles-Salgado, M., Rognoni, T., Fernández-Matarrubia, M., Moreno-Ramos, T. et al. (2017). Addenbrooke's cognitive examination III: Diagnostic utility for mild cognitive impairment and dementia and correlation with standardized neuropsychological tests. *International Psychogeriatrics*, *29*(*1*), 105–113. https://doi.org/10.1017/S1041610216001496.

Ministry of Health and Welfare in Taiwan. (2020). *The political framework and action on dementia care and prevention 2.0 (including working items)*. Taipei, Taiwan: Ministry of Health and Welfare. Available at: https://www.health.taichung.gov.tw/media/553446/.pdf. Accessed March 8, 2021.

Morris, J. C. (1993). The clinical dementia rating (CDR): Current version and scoring rules. *Neurology*, *43*(*11*), 2412–2414.

Prince, M., Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., & Prina, M. (2015). The global impact of dementia: An analysis of prevalence, incidence, cost and trends: Alzheimer's disease international. *2015*.

Ranson, J. M., Kuzma, E., Hamilton, W., Muniz-Terrera, G., Langa, K. M., & Llewellyn, D. J. (2019). Predictors of dementia misclassification when using brief cognitive assessments. *Neurology: Clinical Practice*, *9*(2), 109–117. https://doi.org/10.1212/CPJ.0000000000000566.

Savva, G. M., & Arthur, A. (2015). Who has undiagnosed dementia? A cross-sectional analysis of participants of the aging, demographics and memory study. *Age and Ageing*, *44*(*4*), 642–647. https://doi.org/10.1093/ageing/afv020.

Schroeder, R. W., Peck, C. P., Buddin, W. H., Jr., Heinrichs, R. J., & Baade, L. E. (2012). The coin-in-the-hand test and dementia: More evidence for a screening test for neurocognitive symptom exaggeration. *Cognitive and Behavioral Neurology*, *25*(*3*), 139–143. https://doi.org/10.1097/WNN.0b013e31826b71c1.

Sun, Y., Lee, H. J., Yang, S. C., Chen, T. F., Lin, K. N., Lin, C. C. et al. (2014). A nationwide survey of mild cognitive impairment and dementia, including very mild dementia, in Taiwan. *PLoS One*, *9*, e100303. https://doi.org/10.1371/journal.pone.0100303.

Teng, E. L., Hasegawa, K., Homma, A., Imai, Y., Larson, E., Graves, A. et al. (1994). The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. *International psychogeriatrics*, *6*(*1*), 45–58.

The University of Sydney. (2021). *Diagnostic dementia tests-ACE-III*. Australia: Brain and Mind Centre, the University of Sydney (online). Available at: www.ftdrg.org www.ftdrg.org. Accessed March 8, 2021.

Tsai, J. C., Chen, C. W., Chu, H., Yang, H. L., Chung, M. H., Liao, Y. M., & Chou, K. R. (2016). Comparing the sensitivity, specificity, and predictive values of the Montreal Cognitive Assessment and Mini-Mental State Examination when screening people for mild cognitive impairment and dementia in Chinese population. *Archives of psychiatric nursing*, *30*(*4*), 486–491.

Tsao, F. F., & Lu, F. H. (1999). Exploration of the communication between Taiwan physicians and elderly patients from a perspective of discourse analysis. *Ministry of Science and Technology Report*, *1999*.

Wang, B. R., Ou, Z., Gu, X. H., Wei, C. S., Xu, J., & Shi, J. Q. (2017). Validation of the Chinese version of Addenbrooke's cognitive examination III for diagnosing dementia. *International Journal of Geriatric Psychiatry*, *32*(*12*), e173–e179. https://doi.org/10.1002/gps.4680.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Wang, C. S., Pai, M. C., Chen, P. L., Hou, N. T., Chien, P. F., & Huang, Y. C. (2013). Montreal cognitive assessment and mini-mental state examination performance in patients with mild-to-moderate dementia with Lewy bodies, Alzheimer's disease, and normal participants in Taiwan. *International Psychogeriatrics*, *25*(*11*), 1839–1848. https://doi.org/10.1017/S1041610213001245.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV)*. San Antonio, TX: 2008: Pearson Assessment.

Wechsler, D. (2015a). *Wechsler Adult Intelligence Scale–Fourth edition (WAIS–IV): Manual for Taiwan*. Taipei, Taiwan: 2015: Chinese Behavioral Science Corporation.

Wechsler, D. (2015b). *Wechsler Adult Intelligence Scale–Fourth edition (WAIS–IV): Technical and interpretative manual for Taiwan*. Taipei, Taiwan: 2015: Chinese Behavioral Science Corporation.

World Health Organization. (2017). *Global action plan on the public health response to dementia 2017–2025*. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO. Available at: https://apps.who.int/iris/bitstream/handle/10665/259615/9789241513487-eng.pdf?sequence=1. Accessed March 8, 2021.

Yeh, T. C., Chou, Y. C., Weng, J. P., Yeh, H. W., Kao, Y. C., Chiang, W. S. et al. (2018). Detection of malingering in the memory of patients with dementia: A pilot study on coin-in-the-hand test in a northern Taiwan memory clinic. *Journal of Medical Sciences*, *39*(2), 81–89. https://doi.org/10.4103/jmedsci.jmedsci_100_18.

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Appendix 1. Cut-off scores for the T-ACE-III (adjusted) in differentiating dementia from non-dementia

Cut-off scores	Sensitivity	Specificity	Youden Index
73/74	0.895	1.000	0.895
74/75	0.895	0.970	0.864
75/76	0.912	0.970	0.882
76/77	0.930	0.970	0.900
77/78	0.947	0.970	0.917
78/79	0.965	0.939	0.904
80	0.965	0.909	0.874
82/83	0.965	0.879	0.844
83/84	0.982	0.848	0.831
84/85	0.982	0.818	0.801
85/86	0.982	0.788	0.770
86/87	1.000	0.788	0.788

Downloaded from https://academic.oup.com/acn/article/37/3/692/6412401 by guest on 27 February 2023

Appendix 17: T-ACE-III study - UCL ethical approval letter

UCL RESEARCH ETHICS COMMITTEE
OFFICE FOR THE VICE PROVOST RESEARCH

8th April 2020

Professor Gill Livingston
Division of Psychiatry
UCL

Dear Professor Livingston

Notification of Ethics Approval with Provisos
Project ID/Title: 16845/001: Validation of cognitive screening tests in traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia

Further to your satisfactory responses to the Committee's comments, I am pleased to confirm in my capacity as Chair of the UCL Research Ethics Committee (REC) that your study has been ethically approved by the UCL REC until **31st March 2023.** Approval is granted on condition that data collection does not commence until you have secured approval from the TMU-Joint Institutional Review Board in Taiwan with written evidence provided for our records.

Ethical approval is also subject to the following conditions:

Notification of Amendments to the Research
You must seek Chair's approval for proposed amendments (to include extensions to the duration of the project) to the research for which this approval has been given. Each research project is reviewed separately and if there are significant changes to the research protocol you should seek confirmation of continued ethical approval by completing an 'Amendment Approval Request Form'
http://ethics.grad.ucl.ac.uk/responsibilities.php

Adverse Event Reporting – Serious and Non-Serious
It is your responsibility to report to the Committee any unanticipated problems or adverse events involving risks to participants or others. The Ethics Committee should be notified of all serious adverse events via the Ethics Committee Administrator (ethics@ucl.ac.uk) immediately the incident occurs. Where the adverse incident is unexpected and serious, the Joint Chairs will decide whether the study should be terminated pending the opinion of an independent expert. For non-serious adverse events the Joint Chairs of the Ethics Committee should again be notified via the Ethics Committee Administrator within ten days of the incident occurring and provide a full written report that should include any amendments to the participant information sheet and study protocol. The Joint Chairs will confirm that the incident is non-serious and report to the Committee at the next meeting. The final view of the Committee will be communicated to you.

Office of the Vice Provost Research, 2 Taviton Street
University College London
Tel: +44 (0)20 7679 8717
Email: ethics@ucl.ac.uk
http://ethics.grad.ucl.ac.uk/

Final Report

At the end of the data collection element of your research we ask that you submit a very brief report (1-2 paragraphs will suffice) which includes in particular issues relating to the ethical implications of the research i.e. issues obtaining consent, participants withdrawing from the research, confidentiality, protection of participants from physical and mental harm etc.

In addition, please:

- ensure that you follow all relevant guidance as laid out in UCL's Code of Conduct for Research: https://www.ucl.ac.uk/srs/file/579
- note that you are required to adhere to all research data/records management and storage procedures agreed as part of your application. This will be expected even after completion of the study.

With best wishes for the research.

Yours sincerely

Professor Michael Heinrich
Joint Chair, UCL Research Ethics Committee

Appendix 18: T-ACE-III study - Taiwan ethical approval letter (in English & traditional Chinese)

TMU-JIRB Form071/20180201

Taipei Medical University

Certificate of TMU-JIRB Approval

Issue Date: 2020/05/13

TMU-JIRB No.: N202004089
Protocol Title:
Validation of Cognitive Screening tests in Traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia
Principal Investigator: Chaur-Jong Hu
CO- Investigator: Chien Tai Hong, Gill Livingston
Study Member: Ruan Ching Yu
Study Site:TMU-Shuang-Ho Hospital,University College London
Protocol Version/Date:Version 1.3/ 2020.01.20
Interviewee Informed Consent Forms:Version 4.0/ 2020/04/26
Interviewee Informed Consent Forms(Participant and Proxy):Version 4.0/ 2020/04/26
Case Report Forms:Version 3.0/ 2020/04/26
Advertisement:Version 3.0/ 2020/04/26

The above study has been approved by Full-board review process of the TMU-Joint Institutional Review Board in meeting #109-05-2(date:2020/05/12), duration of validity is from 2020/05/13 to 2021/05/12, and must be monitored by TMU-JIRB.

According to Ministry of Health and Welfare and the relevant regulations, follow-up procedures and requirements are as below:

1. Continuing Report: Frequency of the report of this trial/study is every 12 month, and should be submitted to TMU-JIRB for review 2 months prior to expiry date (2021-03-12) or the trial/study must be pending.
2. Final Report: The report should be submitted to TMU-JIRB for review once completed TMU-JIRB may withdraw the approval of the trial/study if the report didn't submitted within 3 months from the date of validity and will suspend PI from new application for 3 months per TMU-JIRB SOPs.
3. SAE: Serious Adverse Event(s) (SAE) Report: SAE report(s) should be submitted to related authorities according to "Regulations for Good Clinical Practice" as well as "Procedures for Reporting Serious Adverse Drug Reaction" by MOHW.

Chairman:

[redacted]

臺北醫學大學暨附屬醫院
聯合人體研究倫理委員會
Taipei Medical University
Joint Institutional Review Board

本會組織與執行皆符合適用法規
The TMU-Joint Institutional Review Board performs its functions according to written operating procedures and complies with GCP and with the applicable regulatory requirements.

421c9510d28df2058d2421228d3e3805

TMU-JIRB Form071/20180201

TMU-JIRB Form071/20180201

臺北醫學大學

臺北醫學大學暨附屬醫院聯合人體研究倫理委員會

TMU-Joint Institutional Review Board

通過證明函 － 一般審查案

開立日期：民國109年05月13日

本會編號：N202004089

計畫名稱：驗證繁體中文版之認知篩檢測驗以偵測台灣人口中失智症個案並提供健康專業人員調查

計畫主持人：胡朝榮

共同主持人：洪千岱、吉爾利文斯頓

研究人員：余阮清

試驗/研究機構：衛生福利部雙和醫院、英國倫敦大學學院

計畫書版本/日期：Version 1.3/ 2020.01.20

受訪者同意書版本/日期：Version 4.0/ 2020/04/26

受訪者同意書(受訪者與代理人)：Version 4.0/ 2020/04/26

個案報告表版本/日期：Version 3.0/ 2020/04/26

招募文宣版本/日期：Version 3.0/ 2020/04/26

上述計畫已通過本會一般審查程序，業經第109-05-2次會期(會議日期：109年05月12日)審查，特此證明。有效期限自民國109年05月13日至民國110年05月12日。試驗/研究期間應接受本會之監督。

依據衛生福利部與相關規定， 後續追蹤程序及要求如下：

1. 期中報告：本計畫期中繳交頻率為**每12個月**，應於有效期限到期**前二個月（民國110年03月12日）**繳交期中報告。有效期限屆滿時若尚未通過期中報告與效期展延審查者，試驗/研究不得繼續執行。
2. 結案報告：試驗/研究完成後，應將執行情形及結果依結案報告表要求送至本會審查。**核准期間到期三個月**仍未繳交者，**本會得撤銷本通過證明函，亦即撤銷本試驗/研究之核准，亦將依本會作業程序暫停主持人(含任何參與形式)申請新試驗/研究案之審查三個月**。
3. 嚴重不良事件(SAE)報告：執行人體試驗或臨床試驗之主持人應根據衛生署「藥品優良臨床試驗準則」和「嚴重藥物不良反應通報辦法」規定，辦理相關事宜。

主任委員：

[redacted]

臺北醫學大學暨附屬醫院
聯合人體研究倫理委員會
Taipei Medical University
Joint Institutional Review Board

本會組織與執行皆符合適用法規
The TMU-Joint Institutional Review Board performs its functions according to written operating procedures and complies with GCP and with the applicable regulatory requirements.

421c9510d28df2058d2421228d3e3805

TMU-JIRB Form071/20180201

Appendix 19: T-ACE-III study - The use of ACE-III approval letter from University of Sydney

Re: About adjusting the ACE-III in traditional Chinese version for using in Taiwan

From: Angela Scharfenberg
Sent: Thursday, 30 January 2020 05:24
To: Yu, Ruan-Ching
Subject: Re: About adjusting the ACE-III in traditional Chinese version for using in Taiwan

Dear Ruan-Ching,

Thank you for your interest in the ACE-III. My apologies for the delayed response.

We would be happy for you to create a traditional Chinese adaptation of the ACE-III for validation. To our knowledge, this does not exist; we typically only allow one translation/adaptation per language, to ensure consistency in application, scoring, and interpretation.

The only requirement of any permission given for adaptations of the ACE-III is that we are provided with a copy of all relevant documentation once completed, which is then uploaded to our website. As with all versions of the ACE-III, please note that copyright is retained by Prof. John Hodges.

If you are happy to proceed on this basis, please let me know.

Kind regards,
Angela

Please note, I work part time: Tuesday and Thursday.

ANGELA SCHARFENBERG | Research Assistant
School of Psychology, Brain and Mind Centre

THE UNIVERSITY OF SYDNEY
94 Mallet Street | The University of Sydney | NSW | 2006
T +61 2 9114 4126 | F +61 2 8322 4000
E angela.scharfenberg@sydney.edu.au | W sydney.edu.au

INSPIRED – the Campaign to support the University of Sydney
sydney.edu.au/inspired

Appendix 20: T-ACE-III study - The use of Brixton Spatial Anticipation test approval letter from the original author(s)

BP

Burgess, Paul
To: Kapur, Narinder
Cc: Yu, Ruan-Ching; Stott, Joshua

Mon 07/10/2019 16:17

Dear Narinder,

I'm glad you are using the test. You are right that it was designed to be, as far as possible, "culture-fair" (although probably no test can be said to be entirely "culture-free").

The Brixton does seem to be particularly good for use with dementia patients. I'm don't have a particularly good account of why this should be, yet: I would love to analyze the individual response patterns of a large sample of dementia cases (perhaps using machine learning) to determine exactly what is going on with them. However, there is enough data now in terms of overall performance on the Brixton with this population to suggest it is useful.

I don't have any spare Brixtons, I'm afraid - I have to buy them from the publishers just like everyone else. (Actually, I have to get someone else to buy it for me since I am not a clinical psychologist so I am not allowed to purchase my own test!) But Pearson, the publishers, will sometimes let you have a copy for research purposes if you contact them. So I suggest you make a request to them.

If they will not give you one, then get back to me and I'll see if I've got any money left over from my student projects last year to buy one for you.

One of my recent PhD students, Hsuan-Chen Wu, is from Taiwan. I will ask him if he has used the Brixton "at home", or in China (where he is now working, in a hospital).

Good luck with your research,

Paul

Professor Paul W. Burgess
Institute of Cognitive Neuroscience,
UCL (University College London)
17 Queen Square,
London WC1N3AR
U.K.
Tel: +44(0)207 679 1139
+44(0)1243 514492

KN

Kapur, Narinder
To: Burgess, Paul
Cc: Yu, Ruan-Ching; Stott, Joshua

Sun 06/10/2019 10:02

Paul

Hope this finds you well.

Along with Josh Stott, and Gill Livingstone and her colleagues in Psychiatry, I am supervising a PhD student from Taiwan who is looking at cognitive screening for dementia in Chinese populations.

The student, Ruan-Ching, is planning to use your excellent Brixton Test, and both I and my colleagues in developing countries have found it very useful as a culture-fair test of executive function. Naturally, any data that are gathered in Taiwan would be good publicity for the test.

Ruan-Ching does not have any funds to purchase tests for her research, and I wonder if you had a spare copy of the test, or know how we can get one?

Kind Regards

Narinder

Narinder Kapur
Visiting Professor of Neuropsychology
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London
WC1E 7HJ

Appendix 21: T-ACE-III study - The use of The-coin-in-the-hand test approval letter from the original author(s)

Suggestions for cognitive testing

7+

KN Kapur, Narinder

To: Yu, Ruan-Ching

Cc: Livingston, Gill; Mukadam, Naaheed; Stott, Joshua

Wed 18/09/2019 10:01

15 KB | 510 KB | 103 KB | 1 MB

Coin-in-Hand Chinese study ... 1,000 KB

coin-in-hand dementia 2012.... 116 KB

6 attachments (3 MB) Save all to OneDrive - University College London Download all

Ruan-Ching

Just to summarise my suggestions for the cognitive testing for the empirical part of your PhD....

1. I think you could do the basic ACE-III, with appropriate minor changes (e.g. Prime Minister question, etc). But you could extend the ACE-III without significantly altering its validity...e.g. after the main test, adding delayed recall and recognition of the pictures used in the naming test, and asking for delayed recall of the figures used in the copying task. You could also test recognition memory for the 4 dot arrays used in the Dot Counting part of the ACE-III. Naturally, there would be some work involved in creating these extra bits.

2. Speed of Information Processing Tests. I suggest the Coding, Symbol Search and Cancellation subtests of the Wechsler Adult Intelligence Scale IV.

3. Brixton Spatial Anticipation Test. This is a largely non-verbal test of executive function. If possible, it would be better if you made up a slightly shorter version of the test.

4. My Coin-in-the-Hand test of cognitive effort. Test procedure and some relevant papers attached.

Some of the above would of course involve costs, and you will need to check how the costs would be met.

Whatever battery your decide on, best if you practise it as much as possible with the rest of us before you go back, and ideally also with a few Chinese-speaking volunteers.

Narinder

Narinder Kapur
Visiting Professor of Neuropsychology
Research Department of Clinical, Educational and Health Psychology
University College London
1-19 Torrington Place
London
WC1E 7HJ

Appendix 22: Discussion on T-ACE-III adaptation

Date	People involved	Cultural adaptation content
15/11/2019	Me	• Translated from Simplified Chinese to Traditional Chinese in all content
16/11/2019	Me	• **MEMORY Anterograde memory (name & address):** Revised the name and address according to Taiwan cultural context
18/12/2019	GL, NM, NK, JS, Me	• **MEMORY Retrograde memory (famous people):** Revised [the name of current Taiwan president, the president experienced gun shooting event, the only female emperor in China, the name of the current China president] according to Taiwan cultural context. • **FLUENCY:** Added "vegetable" as a supplemental test. • **LANGUAGE Object naming:** Replaced some pictures [harp->Guzheng, rhinoceros->elephant, Oak barrel->urn, crown->Emperor's hat] by using painting tool and PowerPoint to draw urn and searching on google pictures • **LANGUAGE Comprehension:** Revised [what is wore by the emperor?]
01/01/2020	GL, NM, NK, JS, Me	• **PERCEPTUAL ABILITIES Identifying letters:** Revised [from English alphabet: K, M, A, T to Arabic numbers: 8, 4, 9, 2] by using PowerPoint
03/01/2020	GL, NM, NK, JS, Me	• **FLUENCY:** Deleted the [generate words starting with car]
13/01/2020	GL, NM, NK, JS, Me	• **LANGUAGE Single word repetition:** (1) Revised the traditional Chinese phase according to the word frequency and rules of English version. (2) Emailed Professor Chan for suggestions on language psychology
20/01/2020		• **LANGUAGE Comprehension:** Revised the [Please point out which animal is reptile?]
29/01/2020	J-CL, Me	• **FLUENCY:** Start the sentences with "car" (simplified Chinese) with "fruit" might be better (women are better in answering vegetable). • **MEMORY Retrograde memory (famous people):** [The current president in China]. Chinese president is no longer restricted with 4-year presidency. so, it is better to remain the same [the current president in America]. • **LANGUAGE Object naming:** (1) Remain [rhinoceros] instead of replacing with [elephants] (2) [The hat of ancient emperor] which is too specific and difficult to elderly people, can be replaced with a “Yuanbao”. [Yuan Dynasty] had a special political meaning in Tang Dynasty. Nowadays, it has been seen as an item for bringing in wealth and blessings and people burn the "paper Yuanbao" to worship god in some Taiwanese traditional religions. https://www.nationsonline.org/oneworld/Chinese_Customs/yuanbao.htm • **LANGUAGE Comprehension:** (1) The first question asking [which object is worn by ancient emperors] can be replaced with [which object can be used for bringing wealth] (2) The second question can be replaced with [which animal eats bamboo as the main food] as asking the concept about reptiles and stringed instrument might be too challenging to elderly people
03/02/2020	C-CY, Me	• **MEMORY Anterograde memory (name & address):** Remember to use current Taiwanese address and name (do not use famous address, such as: Ketagalan Boulevard • **MEMORY Retrograde memory (famous people):** My advice is revising Q2 to ask the name of the Taipei city's leader or President of

		the Executive Yuan, the Q4 asking the Taiwanese president got shot is testing remotely memory, I have no opinion on Q4- the name of the American president ● **FLUENCY:** My suggestion is to reference the Japanese version, generating words with starting certain syllables. The simplified Chinese version's two items are both testing 'category fluency' (starting with "car" and animals) ● **LANGUAGE Object naming:** Guzheng, urn and emperor's hat are not usual things in daily life, the replacements of items from the Eastern culture won't reduce the difficulty. I would rather prefer piano and drum, also, I would recommend that the urn should be replaced with bulb in Japanese version. ● **LANGUAGE Comprehension:** I have no opinion on it [Q1. which item does emperor wear?] [Q2. which animal is reptile] ● **PERCEPTUAL ABILITIES Identifying letters:** I have no opinion on revising [from English alphabet: K, M, A, T to Arabic numbers: 8, 4, 9, 2]. ● **LANGUAGE Reading**: Replacing the 4 clueless simplified Chinese words with 5 traditional Chinese words, I recommend this website: https://language.moe.gov.tw/result.aspx?classify_sn=44&subclassify_sn=454&content_sn=17 which is relatively new. I think the first set of 5 words seems a good one [state, start, service, solution, know]
13/02/2020	NK, Me	● **PERCEPTUAL ABILITIES Identifying letters:** You might make the Arabic numbers look more difficult to guess as they seem easier than English alphabet
19/02/2020	C-CY	● **FLUENCY:** The statement of "you only have 1 minute" will make the participants nervous, I would say "please tell me all the words you know that start with "da" within 1 minute, the more the better". Do it in the same way in animal fluency. The optimal cut-off of 18 should be tested again, you need to pilot test it and see whether they can speak 18 within 1 minute. ● **MEMORY Retrograde memory (famous people):** The first Taiwanese president experienced gun shooting event, should be revised to [the first Taiwanese president experience gun shooting event and got hurt] as the first experienced gun shooting event is not Chen Shui-bian, but Chiang Ching-kuo in America and he did not get hurt. ● **LANGUAGE Comprehension (pick up the paper):** Please touch the paper and [hand in] the pencil to me, it's not easily spoken, change to [give the pencil to me]. ● **LANGUAGE Sentence writing:** [please do not use abbreviations] could change to [please do not use simplified Chinese]. ● **LANGUAGE Proverb repetition:** Pick up some Chinese proverbs that are unusual, hard to speak fluently, but not too long. ● **LANGUAGE Comprehension** [Q4 which item is related to "mooring"] it is too hard to understand, please change to "navigation". ● **LANGUAGE Reading:** I prefer to use words without sound part rather than words with irregular sound parts as it is less complicated.
05/07/2020	Me	● **FLUENCY:** the questions that ask them to start a 'syllable' in traditional Chinese and Hokkien versions are too challenging for participants to understand even for non-dementia group. Therefore, we deleted this a 'syllable' question and used question tests 'category fluency'- vegetables and fruits (its already in the supplemental materials).

Note. C-CY = Cheng-Chang Yang, a postdoc researcher at Taipei Shuang-Ho hospital, Taiwan; GL = Gill Livingston; J-CL = Jen-Chieh Lai, a psychologist Jack at Department of Neurology, Hualien Tzu Chi Medical Centre, Taiwan; JS = Joshua Stott; NK = Narinder Kapur; NM = Naaheed Mukadam.

Appendix 23: T-ACE-III study - poster (in English & traditional Chinese)

Project title: A study to test which memory tests are best at detecting who has dementia and who does not have dementia in the Taiwanese population

Purposes

The project aims to ask questions about memory, fluency, thinking, naming and recognising, in order to find the best assessments for detecting and diagnosing who has and who does not have dementia in Taiwan.

Recruitment of participants

- **Non-dementia group:** people who once went to neurology or memory clinics due to memory issues but was diagnosed as free from dementia.
- **Dementia group**: people who are diagnosed as dementia based on DSM-V
- **All participants** need to have no other illness or disability which significantly interferes with understanding or completing of memory and thinking tests eg hearing loss of lack of sight.
- **No age and gender limits.**

Procedure

Time: 40-60 mins interview in one visit
Location: Interview room, Neurology department, Taipei Medical University - Shuang Ho Hospital
Content: You will be asked about your name, age, education and language spoken, questions from tests designed to test language, memory, word finding, fluency and other brain functions.

Contacts

- **Researcher Ruan-Ching Yu** [redacted]
- Please inform your name, tel number, age, group and available time

Institutions

The department of Mental health of older people, Division of psychiatry, UCL, UK

Neurology department, Taipei Medical University –Shuang Ho Hospital, Taiwan;

Sponsorship: Ministry of Education, Taiwan

計畫題目：偵測台灣人口中失智與非失智族群之最佳測驗

研究目的

該項目旨在詢問有關記憶、流暢度、思考、命名及辨別的問題，以找到最佳的評估方法，以檢測和診斷台灣失智與非失智族群

受試者招募

- **非失智症組**：過去曾因記憶問題至記憶或神經內科門診評估而被排除為失智症
- **失智症組**：符合DSM-V診斷之失智症個案
- **所有受試者**沒有對於完成認知測驗或訪談過程有顯著影響之精神或視覺、聽覺或動作缺損等生理障礙。
- **沒有年齡及性別限制**

研究流程

時間: 一次40-60分之訪問
地點: 衛生福利部雙和醫院神經內科訪談室
內容: 您將被詢問有關姓名、年齡、使用語言及一系列有關語言、記憶力、單字搜尋、流暢度及其他大腦認知功能的相關問題。

聯絡方式

- 余阮清研究員
  [redacted]
- 請告知姓名、電話、年齡、報名組別及可參與時間

相關單位
英國倫敦大學學院
精神醫學所
老年人心理健康學系

台灣衛生福利部
雙和醫院神經內科

贊助單位 台灣教育部公費留學獎勵計畫

Appendix 24: T-ACE-III study - participant information sheet (in English & traditional Chinese)

Participant Information Sheet: A study to test which memory tests are best at detecting who has dementia and who does not have dementia in the Taiwanese population

UCL Research Ethics Committee Approval ID Number: 16845/001

Title of Study:
Validation of Cognitive Screening tests in Traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia
(A study to test which memory tests are best at detecting who has dementia and who does not have dementia in the Taiwanese population)
Department:
Department of Mental Health of Older People, Division of Psychiatry, University College London, UK
Neurology department, Taipei Medical University - Shuang Ho Hospital, Taiwan

Name and Contact Details of the Researcher(s):
Ruan-Ching Yu, ruan-ching.yu.18@ucl.ac.uk (mobile phone number: [redacted]
Name and Contact Details of the Principal Researcher:
Gill Livingston, g.livingston@ucl.ac.uk *(International Principal Researcher)*
Chaur-Jong Hu, chaurjongh@tmu.edu.tw *(Principal Researcher in Taiwan)*

You are being invited to take part in a research project, which is a part of Ruan-Ching Yu PhD student's doctoral dissertation and is being carried out jointly by University College London, UK and Taipei Medical University - Shuang Ho Hospital, Taiwan. The academic supervision is in UCL where Ruan-Ching's PhD is registered and the participant recruitment is in Taiwan. Before you decide whether to take part it is important for you to understand why the research is being done and what participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

1. **What is the project's purpose?**
   The project aims to ask questions about memory, fluency, thinking, naming and recognising, in order to find the best assessments for detecting and diagnosing who has and who does not have dementia in Taiwan.

2. **Why have I been chosen?**
   We are inviting two groups to participate both of whom have attended memory clinics in Taiwan. One group is patients who already have a diagnosis of dementia (such as Alzheimer's disease, Frontotemporal dementia, vascular dementia and Parkinson's disease dementia); another group is older people who do not have dementia. The estimated total number of participants are 72 (48 for dementia group and 24 for the group without dementia).

3. **Do I have to take part?**
   It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep. You can withdraw at any time without giving a reason and without it affecting any benefits that you are entitled to, although once the data has been incorporated in the analysis it will not be possible to withdraw data. All data in any research results and publication will be anonymised so you will not be able to be identified.

4. **What will happen to me if I take part?**
   This study will involve a one-time interview for 40 to 60 minutes. Before starting, you need to sign a consent form. The interview includes two parts - in the first part, you will be asked personal information regarding your name, age, language spoken etc. This is so that we can compare responses from people of different ages and with different backgrounds. In the second part, you will be asked a series of questions from tests designed to test language, memory, word finding, fluency and other brain functions. For some items, it will require to write on the test sheet and for some, the researcher will write down your answers. Before the formal evaluation starts, the researchers will provide further information.

5. **What are the possible disadvantages and risks of taking part?**
   We will be administering a range of cognitive tests. These may be tiring for you or may cause distress in those who find it difficult to complete these tasks. We will ensure that participants are given enough time to complete tasks and are given adequate breaks to mitigate this. Some questions may cause reflection and increase realisation of cognitive deterioration which may be upsetting. The interview will be paused by the researchers according to their observation on whether you are tired or frustrated. Your family or caregivers will be informed and you will have a break till you recover, then decide whether you can continue to finish the interview or come back next time. You can contact the researcher through the contact information on the participant information sheet if you have any problem after taking part in the study. You can also contact the neurology department in Taipei Shuang-Ho hospital if you wish to speak to someone about how you are feeling

6. **What are the possible benefits of taking part?**
   Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will bring some new insights on the memory tests used in Taiwan, which can potentially help doctors in Taiwan to better diagnose dementia and could help people with dementia in Taiwan be diagnosed earlier and more accurately.

7. **What if something goes wrong?**
   There is no treatment or risks in this project, but if you have any concerns, about the research, including the invitation, consent process, interview, and personal data processing, you should contact either the researcher or the principal supervisor of this project let them know what your concerns are. If you are still unsatisfied, you can contact the Chair of the UCL Research Ethics Committee (Email: ethics@ucl.ac.uk). In terms of your rights during the study or suspecting that you have been harmed by participating in the / study, please contact TMU-Joint Institutional Review Board at any time (Tel: +886 (02) 66382736 Ext. 1728 or Email: tmujirb@gmail.com).

8. **Will my taking part in this project be kept confidential?**
   All the identifiable information that we collect about you during the course of the research will be kept strictly confidential. Any material containing your name or details will be kept in secure conditions and will only be accessed by the research team. We will allocate a unique identifying number to you when you enrol on the study and any test results and questionnaires completed by you will only have your number on them and not your name or other details. We will make sure that you will not be able to be identified in any reports or publications.

9. **Limits to confidentiality**
   Confidentiality will be respected subject to legal constraints and professional guidelines. We have a duty to report something if we have concerns about anyone being harmed.

10. What will happen to the results of the research project?

The result of this research will be published after it is concluded and will be written up as a PhD thesis. If you wish, you can request a copy of the published results by emailing to the researcher (ruan-ching.yu.18@ucl.ac.uk) or by telling the researcher. You will not be identified in any report or publication. Test results will be securely archived by the research team and kept for as many years as required by local guidelines. Once your data has been kept for as long as needed, it will be securely disposed of. The original data you provided will be stored and used in the database of Professor Gill Livingston and Doctor Chaur-Jong Hu and will not be shared with others. If but anonymised data without personal details may be shared with other researchers

11. Local Data Protection Privacy Notice

The controllers for this project will be University College London (UCL) and Taipei Medical University - Shuang Ho Hospital. UCL will be responsible for research design, analysis and publication. Taipei Medical University - Shuang Ho Hospital will be responsible for data collection, storage and deletion. The UCL Data Protection Officer and TMU-Joint Institutional Review Board provide oversight of UCL and Taipei Medical University - Shuang Ho Hospital activities involving the processing of personal data, respectively. UCL data protection can be contacted at data-protection@ucl.ac.uk (UCL) or tmujirb@gmail.com, +886 (02) 66382736 Ext. 1728 (TMU-Joint Institutional Review Board).

12. Who is organising and funding the research?

This project is funded by Education Ministry of Taiwan, and is estimated to start from June 2020 to March 2023 and will be jointly carried out through collaboration between Professor Gill Livingston at the division of Psychiatry, UCL and Doctor Chaur-Jong Hu, Doctor Chien-Tai Hong at the department of neurology at Taipei Medical University - Shuang Ho Hospital. It will be supervised by Professor Gill Livingston, Doctor Naaheed Mukadam in the Division of Psychiatry, UCL and Professor Narinder Kapur and Doctor Joshua Stott in the Department of Psychology, UCL.

14. Contact for further information

For further enquiries or information regarding the published details or copy when it is available, Please contact us.

PhD student-
Ruan-Ching Yu,
email: ruan-ching.yu.18@ucl.ac.uk
mobile phone number [redacted]
Principal supervisor-
Gill Livingston,
email: g.livingston@ucl.ac.uk

Thank you for reading this information sheet and for considering to take part in this research study.

研究計畫參與者資料表

研究探討在偵測台灣人口中失智與非失智族群之最佳測驗

UCL Research Ethics Committee Approval ID Number: 16845/001

研究主題:

驗證繁體中文版之認知篩檢測驗以偵測台灣人口中失智症個案並提供健康專業人員調查

(研究探討在偵測台灣人口中失智與非失智族群之最佳測驗)

單位:

英國倫敦大學學院 精神醫學所 老年人心理健康學系

台灣衛生福利部雙和醫院 神經內科

研究人員姓名與聯絡方式:

余阮清, ruan-ching.yu.18@ucl.ac.uk (電話號碼: [redacted]

研究主持人姓名與聯絡方式:

Gill Livingston, g.livingston@ucl.ac.uk (國際研究主持人)

胡朝榮 chaurjongh@tmu.edu.tw (台灣研究主持人)

您已受邀參加一項研究計畫，該研究計畫為博士班學生余阮清的博士學位論文的一部分，由英國倫敦大學學院及台灣台北醫學大學-衛生福利部雙和醫院聯合進行。研究監督機構為英國倫敦大學學院- 余阮清攻讀博士學位之所在地，而受試者之招募則在台灣進行。在決定是否參加之前，對您來說重要的是—*了解為什麼要進行研究以及參與的過程*。請花些時間仔細閱讀以下信息，也可與其他研究人員討論，詢問我們不清楚的地方以及想了解的更多信息。請您花一些時間來決定是否要參加，感謝您閱讀此訊息。

1. 本研究計畫為何?

該項目旨在詢問*有關記憶、流暢度、思考、命名及辨別的問題*，以找到最佳的評估方法，以檢測和診斷台灣失智與非失智族群。

2. 我為何會被選中?

本研究邀請曾至衛生福利部雙和醫院神經內科就診的兩個小組。一為失智症確診的患者(如阿茲海默症，額顳葉失智症，血管性失智症和帕金森失智症)；一為是沒有失智症的老年人。估計的受試者總數為 72 人*(失智症組為 48 人，無失智症組為 24 人)*。

3. 我必須要參加嗎?

由您決定是否參加。如果您決定參加，此信息表將給予您保留備查。*您也可以在不提出任何原因的前提下，隨時退出本研究，也不會影響您有權獲得的任何利益，唯在資料已被納入分析後將無法撤消，但任何研究結果與出版物中之資料都將被匿名化，您將無法被識別。*

4. 我參加後會發生什麼事?

本研究將進行一次性 40-60 分之訪談。在開始前，您需要完成簽署該份同意書。訪談包含兩部分，第一部分，您將被詢問有關姓名、年齡、使用語言等個人資料。第二部分，您將被詢問一系列有關語言、記憶力、單字搜尋、流暢度及其他大腦認知功能的相關問題。在回答某些題目，需要您寫在測試紙上，而對於某些題目，研究人員則會協助您寫下答案。在正式施測開始前，研究人員將提供更多訊息。

5. 參加研究的可能缺點與風險為何**?**
我們將進行一系列認知測試，可能會使您感到疲倦，或者因感到某些題項較難回答而有挫折感。我們將確保給予您足夠時間來完成，並給予足夠休息以減輕這種情況。 一些問題可能增加對於認知功能退化事實的理解，會讓人感到沮喪。研究人員將根據觀察，看您是否疲倦或沮喪而暫停訪談，將通知您的家屬或照顧者並讓您適當休息，待心情平復後，再依個別狀況繼續施測或預約下一次再回來完成測驗。 如果參加研究後有任何問題，您可以通過「參與者信息表上的聯繫信息」與研究人員聯繫。若您想談談您的感受，也可以聯繫台北衛生福利部雙和醫院神經內科。

6. 參加研究的可能優點為何**?**
即使本研究計*對於參與的受試者並沒有立即的幫助*，但希冀這項研究將對台灣臨床工作運用的記憶測試帶來一些新見解，從而協助醫師更有效率地診斷癡呆症，並幫助台灣之癡呆症個案在早期階段可以更準確地被診斷。

7. 假設研究過程發生錯誤**?**
參與此計畫沒有任何藥物介入或風險，但是若您有任何疑慮，有關研究過程中的細節 (包含邀請、知情同意程序、訪談與個人資料處理方式)，請與該計畫研究人員或主持人聯繫，並讓他們知道您的擔憂，若您對他們的回覆仍不滿意，可進一步聯繫 UCL 研究倫理委員會執行長 (Email: ethics@ucl.ac.uk)；有關您在試驗/研究過程中的權益，或您懷疑因參與試驗/研究而受害，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡（電話：(02)66382736 分機 1728 或 Email：tmujirb@gmail.com）。

8. 我參加研究的過程與資料會被保密嗎**?**
本計畫研究過程中收集與您有關的所有可識別資料將進行嚴格保密，您的姓名及個人資料的任何細節都將被緘密保存在安全的條件下，僅研究團隊人員才能取得。當您參加研究時，我們將為您分配一個獨特的識別號，您填寫的任何測驗結果和訪談表上只會標示該組號碼，不會顯示您的姓名或其他個人資料。我們將確保您在任何研究報告或出版物中都無法被識別。

9. 保密限制
在遵守法律約束和專業準則的前提下，本計畫將尊重並服從機密性。*但若有可能傷害發生之疑慮發生，我們則有責任進行舉報。*

10. 研究結果將會發生什麼事**?**
本計畫研究結果將在完成後進行發表，並將被寫成博士學位論文。若您有興趣進一步了解，可通過電子郵件聯繫研究人員（ruan-ching.yu.18@ucl.ac.uk）或告知研究人員，索取已發表研究之副本。敬請放心，您將不會在任何報告或出版物中被辨別。研究資料將由研究團隊安全保存，並依照單位規定保存數年，當資料保存年限截止，將進行安全銷毀。您提供的原始資料，僅限在 Gill Livingston 教授與胡朝榮醫師資料庫中保管使用，不會與他人共享。但未包含個人詳細資料之匿名化資料則可能與其他研究人員共享。

11. 當地資料保護隱私聲明
該計畫負責單位為倫敦大學學院(University College London, UCL)與衛生福利部雙和醫院，UCL 將負責研究設計、分析與發表，衛生福利部雙和醫院則負責資料搜集、儲存和刪除。UCL 資料保護單位以及臺北醫學大學暨附屬醫院聯合人體研究倫理委員分別負責監督 UCL 與衛生福利部雙和醫院涉及個人資料處理的活動，您可以通過 data-protection@ucl.ac.uk (UCL)或 tmujirb@gmail.com、(02)66382736 分機 1728 (臺北醫學大學暨附屬醫院聯合人體研究倫理委員會)與其聯繫。

12. 誰在組織及資助本項研究**?**

該研究計畫由台灣教育部資助，預計研究期間為 2020 年 6 月至 2023 年 3 月，由英國 UCL 精神醫學學系 Gill Livingston 教授與台灣衛生福利部雙和醫院神經內科胡朝榮醫師、洪千岱醫師合作進行，並由 UCL 精神醫學學系 Gill Livingston 教授、Naaheed Mukadam 助理教授以及心理學系的 Narinder Kapur 教授、Joshua Stott 助理教授共同指導。

13. 更多訊息聯絡方式**:** 若您想了解有關本研究發佈之訊息或任何疑問，請不吝與我們聯繫。

研究生 余阮清 (ruan-ching.yu.18@ucl.ac.uk, [redacted]
計畫主持人 **Gill Livingston** (g.livingston@ucl.ac.uk)

非常感謝您閱讀此訊息並考慮參加本研究計畫

Appendix 25: T-ACE-III study - proxy information sheet (in English & traditional Chinese)

Proxy Participant Information Sheet : A study to test which memory tests are best at detecting who has dementia and who does not have dementia in the Taiwanese population

UCL Research Ethics Committee Approval ID Number: 16845/001

Title of Study:
Validation of Cognitive Screening tests in Traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia
(A study to test which memory tests are best at detecting who has dementia and who does not have dementia in the Taiwanese population)
Department:
Department of Mental Health of Older People, Division of Psychiatry, University College London, UK
Neurology department, Taipei Medical University - Shuang Ho Hospital, Taiwan

Name and Contact Details of the Researcher(s):
Ruan-Ching Yu, ruan-ching.yu.18@ucl.ac.uk (mobile phone number: [redacted]
Name and Contact Details of the Principal Researcher:
Gill Livingston, g.livingston@ucl.ac.uk *(International Principal Researcher)*
Chaur-Jong Hu, chaurjongh@tmu.edu.tw *(Principal Researcher in Taiwan)*

You and your family/ friend are being invited to take part in a research project, which is a part of Ruan-Ching Yu PhD student's doctoral dissertation and is being carried by jointly by University College London, UK and Taipei Medical University - Shuang Ho Hospital, Taiwan. The academic supervision is in UCL where Ruan-Ching's PhD is registered and the participant recruitment is in Taiwan. Before you decide whether to take part it is important for you to understand why the research is being done and what participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

1. **What is the project's purpose?**
   The project aims to ask questions about memory, fluency, thinking, naming and recognising, in order to find the best assessments for detecting and diagnosing who has and who does not have dementia in Taiwan.

2. **Why have we been chosen?**
   We are inviting two groups to participate both of whom have attended memory clinics in Taiwan. One group is patients who already have a diagnosis of dementia (such as Alzheimer's disease, Frontotemporal dementia, vascular dementia and Parkinson's disease dementia); another group is older people who do not have dementia. The estimated total number of participants are 72 (48 for dementia group and 24 for the group without dementia).

3. **Do we have to take part?**
   It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep. You can withdraw at any time without giving a reason and without it affecting any benefits that you are entitled to, although once the data has been incorporated in the analysis it will not be possible to withdraw data. All data in any research results and publication will be anonymised so you and your family/ friend will not be able to be identified.

4. **What will happen to me if we take part?**
   This study will involve a one-time interview for 40 to 60 minutes. Before starting, you need to sign a consent form. The interview includes two parts - in the first part, you will be asked personal information regarding your family's/ friend's name, age, language spoken etc. This is so that we can compare responses from people of different ages and with different backgrounds. In the second part, your family/ friend will be asked a series of questions from tests designed to test language, memory, word finding, fluency and other brain functions. For some items, it will require your family/ friend to write on the test sheet and for some, the researcher will write down the answers. Before the formal evaluation starts, the researchers will provide further information.

5. **What are the possible disadvantages and risks of taking part?**
   We will be administering a range of cognitive tests. These may be tiring for your family/ friend or may cause distress in those who find it difficult to complete these tasks. We will ensure enough time is given to complete tasks and are given adequate breaks to mitigate this. Some questions may cause reflection and increase realisation of cognitive deterioration which may be upsetting. The interview will be paused by the researchers according to their observation on whether your family/ friend are tired or frustrated. You will be informed and they will have a break till they recover, then decide whether they can continue to finish the interview or come back next time. You can contact the researcher through the contact information on the participant information sheet if you have any problem after taking part in the study. You can also contact the neurology department in Taipei Shuang-Ho hospital if you wish to speak to someone about how you are feeling

6. **What are the possible benefits of taking part?**
   Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will bring some new insights on the memory tests used in Taiwan, which can potentially help doctors in Taiwan to better diagnose dementia and could help people with dementia in Taiwan be diagnosed earlier and more accurately.

7. **What if something goes wrong?**
   There is no treatment or risks in this project, but if you have any concerns, about the research, including the invitation, consent process, interview, and personal data processing, you should contact either the researcher or the principal supervisor of this project let them know what your concerns are. If you are still unsatisfied, you can contact the Chair of the UCL Research Ethics Committee (Email: ethics@ucl.ac.uk). In terms of your family's/ friend's rights during the study or suspecting that you and your family/ friend have been harmed by participating in the trial / study, please contact TMU-Joint Institutional Review Board at any time (Tel: +886 (02) 66382736 Ext. 1728 or Email: tmujirb@gmail.com).

8. **Will our taking part in this project be kept confidential?**
   All the identifiable information that we collect about you and your family/ friend during the course of the research will be kept strictly confidential. Any material containing your family's/ friend's name or details will be kept in secure conditions and will only be accessed by the research team. We will allocate a unique identifying number to your family/ friend when enrolling on the study and any test results and questionnaires completed by you and your family/ friend will only have your number on them and not your name or other details. We will make sure that you and your family/ friend will not be able to be identified in any reports or publications.

9. **Limits to confidentiality**
   Confidentiality will be respected subject to legal constraints and professional guidelines. We have a duty to report something if we have concerns about anyone being harmed.

10. What will happen to the results of the research project?

The result of this research will be published after it is concluded and will be written up as a PhD thesis. If you wish, you can request a copy of the published results by emailing to the researcher (ruan-ching.yu.18@ucl.ac.uk) or by telling the researcher. You and your family/ friend will not be identified in any report or publication. Test results will be securely archived by the research team and kept for as many years as required by local guidelines. Once your data has been kept for as long as needed, it will be securely disposed of. The original data you and your family/ friend provided will be stored and used in the database of Professor Gill Livingston and Doctor Chaur-Jong Hu and will not be shared with others. If but anonymised data without personal details may be shared with other researchers.

11. Local Data Protection Privacy Notice

The controllers for this project will be University College London (UCL) and Taipei Medical University - Shuang Ho Hospital. UCL will be responsible for research design, analysis and publication. Taipei university will be responsible for data collection, storage and deletion. The UCL Data Protection Officer and TMU-Joint Institutional Review Board provide oversight of UCL and Taipei Medical University - Shuang Ho Hospital activities involving the processing of personal data, respectively. UCL data protection can be contacted at data-protection@ucl.ac.uk (UCL) or tmujirb@gmail.com, +886 (02) 66382736 Ext. 1728 (TMU-Joint Institutional Review Board).

12. Who is organising and funding the research?

This project is funded by Education Ministry of Taiwan, and is estimated to start from June 2020 to March 2023 and will be jointly carried out through collaboration between Professor Gill Livingston at the division of Psychiatry, UCL and Doctor Chaur-Jong Hu, Doctor Chien-Tai Hong at the department of neurology at Taipei Medical University - Shuang Ho Hospital. Also, it will be supervised by Professor Gill Livingston, Doctor Naaheed Mukadam in the Division of Psychiatry, UCL and Professor Narinder Kapur and Doctor Joshua Stott in the Department of Psychology, UCL.

14. Contact for further information

For further enquiries or information regarding the published details or copy when it is available, Please contact us.

PhD student-
Ruan-Ching Yu,
email: ruan-ching.yu.18@ucl.ac.uk
mobile phone number [redacted]
Principal supervisor-
Gill Livingston,
email: g.livingston@ucl.ac.uk

Thank you for reading this information sheet and for considering to take part in this research study.

研究計畫受試者代理人資料表

研究探討在偵測台灣人口中失智與非失智族群之最佳測驗

UCL Research Ethics Committee Approval ID Number: 16845/001

研究主題:
驗證繁體中文版之認知篩檢測驗以偵測台灣人口中失智症個案並提供健康專業人員調查
(研究探討在偵測台灣人口中失智與非失智族群之最佳測驗)

單位:
英國倫敦大學學院 精神醫學所 老年人心理健康學系
台灣衛生福利部雙和醫院 神經內科

研究人員姓名與聯絡方式:
余阮清, ruan-ching.yu.18@ucl.ac.uk [redacted]

研究主持人姓名與聯絡方式:
Gill Livingston, g.livingston@ucl.ac.uk (國際研究主持人)
胡朝榮 chaurjongh@tmu.edu.tw (台灣研究主持人)

您及您的親人/朋友已受邀參加一項研究計畫，該研究計畫為博士班學生余阮清的博士學位論文的一部分，由英國倫敦大學學院及台灣台北醫學大學-衛生福利部雙和醫院聯合進行。研究監督機構為英國倫敦大學學院- 余阮清攻讀博士學位之所在地，而受試者之招募則在台灣進行。對您來說重要的是—*了解為什麼要進行研究以及參與的過程*。請花些時間仔細閱讀以下信息，也可與其他研究人員討論，詢問我們不清楚的地方以及想了解的更多信息。請您花一些時間來決定是否要參加，感謝您閱讀此訊息。

1. 本研究計畫為何?
該項目旨在詢問*有關記憶、流暢度、思考、命名及辨別的問題*，以找到最佳的評估方法，以檢測和診斷台灣失智與非失智族群。

2. 為何會被選中?
本研究邀請曾至衛生福利部雙和醫院神經內科就診的兩個小組。一為失智症確診的患者(如阿茲海默症，額顳葉失智症，血管性失智症和帕金森失智症)；一為是沒有失智症的老年人。估計的受試者總數為 72 人*(失智症組為 48 人，無失智症組為 24 人)*。

3. 必須要參加嗎?
由您決定是否參加。如果您決定參加，此信息表將給予您保留備查。*您也可以在不提出任何原因的前提下，隨時退出本研究，也不會影響您有權獲得的任何利益，唯在資料已被納入分析後將無法撤消，但任何研究結果與出版物中之資料都將被匿名化，您將無法被識別。*

4. 參加後會發生什麼事?
本研究將進行一次性 40-60 分之訪談。在開始前，您需要完成簽署該份同意書。訪談包含兩部分，第一部分，您將被詢問有關您的親人/朋友之姓名、年齡、使用語言等個人資料。第二部分，您的親人/朋友將被詢問一系列有關語言、記憶力、單字搜尋、流暢度及其他大腦認知功能的相關問題。在回答某些題目，需要您的親人/朋友寫在測試紙上，而對於某些題目，研究人員則會協助寫下答案。在正式施測開始前，研究人員將提供更多訊息。

5. 參加研究的可能缺點與風險為何**?**
我們將進行一系列認知測試，可能會使您的親人/朋友感到疲倦，或者因感到某些題項較難回答而有挫折感。 我們將確保給予足夠時間來完成，並給他們足夠休息以減輕這種情況。 一些問題可能增加對於認知功能退化事實的理解，會讓人感到沮喪。 研究人員將根據對於參與者的觀察，看其是否疲倦或沮喪而暫停訪談，將通知您並讓其適當休息，待心情平復後，再依個別狀況繼續施測或預約下一次再回來完成測驗。 如果參加研究後有任何問題，您可以通過「參與者信息表上的聯繫信息」與研究人員聯繫。若您想談談您的感受，也可以聯繫台北衛生福利部雙和醫院神經內科。

6. 參加研究的可能優點為何**?**
即使本研究計*對於參與的受試者並沒有立即的幫助*，但希冀這項研究將對台灣臨床工作運用的記憶測試帶來一些新見解，從而協助醫師更有效率地診斷癡呆症，並幫助台灣之癡呆症個案在早期階段可以更準確地被診斷。

7. 假設研究過程發生錯誤**?**
參與此計畫沒有任何藥物介入或風險，但是若您有任何疑慮，有關研究過程中的細節 (包含邀請、知情同意程序、訪談與個人資料處理方式)，請與該計畫研究人員或主持人聯繫，並讓他們知道您的擔憂，若您對他們的回覆仍不滿意，可進一步聯繫 UCL 研究倫理委員會執行長 (Email: ethics@ucl.ac.uk)；有關您在試驗/研究過程中的權益，或您懷疑您與您的親人/朋友因參與試驗/研究而受害，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡（電話：(02)66382736 分機 1728 或 Email：tmujirb@gmail.com）。

8. 參加研究的過程與資料會被保密嗎**?**
*本計畫研究過程中收集與您及*您的親人/朋友*有關的所有可識別資料將進行嚴格保密*，您與您的親人/朋友的姓名及個人資料的任何細節都將被緘密保存在安全的條件下，僅研究團隊人員才能取得。當您的親人/朋友參加研究時，我們將為其分配一個獨特的識別號，所填寫的任何測驗結果和訪談表上只會標示該組號碼，不會顯示其姓名或其他個人資料。我們將確保您的親人/朋友在任何研究報告或出版物中都無法被識別。

9. 保密限制
在遵守法律約束和專業準則的前提下，本計畫將尊重並服從機密性。*但若有可能傷害發生之疑慮發生，我們則有責任進行舉報。*

10. 研究結果將會發生什麼事**?**
本計畫研究結果將在完成後進行發表，並將被寫成博士學位論文。若您有興趣進一步了解，可通過電子郵件聯繫研究人員（ruan-ching.yu.18@ucl.ac.uk）或告知研究人員，索取已發表研究之副本。敬請放心，您將不會在任何報告或出版物中被辨別。研究資料將由研究團隊安全保存，並依照單位規定保存數年，當資料保存年限截止，將進行安全銷毀。您與您的親人/朋友提供的原始資料，僅限在 Gill Livingston 教授與胡朝榮醫師資料庫中保管使用，不會與他人共享。但未包含個人詳細資料之匿名化資料則可能與其他研究人員共享。

11. 當地資料保護隱私聲明
該計畫負責單位為倫敦大學學院(University College London, UCL)與衛生福利部雙和醫院，UCL 將負責研究設計、分析與發表，衛生福利部雙和醫院則負責資料搜集、儲存和刪除。UCL 資料保護單位以及臺北醫學大學暨附屬醫院聯合人體研究倫理委員分別負責監督 UCL 與衛生福利部雙和醫院涉及個人資料處理的活動，您可以通過 data-protection@ucl.ac.uk (UCL)或 tmujirb@gmail.com、(02)66382736 分機 1728 (臺北醫學大學暨附屬醫院聯合人體研究倫理委員會)與其聯繫。

12. 誰在組織及資助本項研究**?**
該研究計畫由台灣教育部資助，預計研究期間為 2020 年 6 月至 2023 年 3 月，由英國 UCL 精神醫學學系 Gill Livingston 教授與台灣衛生福利部雙和醫院神經內科胡朝榮醫師、洪千岱醫師合作進行，並由 UCL 精神醫學學系 Gill Livingston 教授、Naaheed Mukadam 助理教授以及心理學系的 Narinder Kapur 教授、Joshua Stott 助理教授共同指導。

13. 更多訊息聯絡方式**:** 若您想了解有關本研究發佈之訊息或任何疑問，請不吝與我們聯繫。

研究生 余阮清 (ruan-ching.yu.18@ucl.ac.uk,[redacted]
計畫主持人 **Gill Livingston** (g.livingston@ucl.ac.uk)

非常感謝您閱讀此訊息並考慮參加本研究計畫

Appendix 26: T-ACE-III study informed consent form for participants (in English & traditional Chinese)

TMU-Joint Institutional Review Board

Informed Consent Form for participants (Interviewees)

We invite you to participate in this study. This consent form provides you with relevant information about this project. The investigators of this study or researcher will explain and answer any related questions for you.

The title of this study	**Validation of cognitive screening tests in Traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia**		
Project members	**name**	**Telephone number**	**E-mail**
Principal investigator	**Chaur-Jong Hu**	[redacted]	
Co-investigator	**Chien-Tai Hong**		
Co-investigator	**Gill Livingston**		
Organizers and contact number	**The department of Mental health of older people, Division of psychiatry, University College London; Neurology department, Taipei Medical University - Shuang Ho Hospital.** **Contact number: +44 (0)2076792000, +886 (0)222490088**		
Conduct period	**1st June 2020 – 31st March 2023, approximately 2 year(s) 10 month(s).**		
Brief introduction of this project	Taiwan is a Chinese-speaking country and is a melting pot of different cultures. Hoklo Taiwanese, who speak the Hokkien language, make up the majority of the population. Hokkien speakers have a comparatively lower education level than people who speak Chinese as their first language. Cognitive tests which are validated in mainland Chinese populations may, consequently, have different psychometric properties in Taiwan. This proposal is therefore to validate a battery of brief cognitive tests for investigating people with possible dementia presenting to health professionals in Taiwan. The tests have been chosen after systematic review of tests available and validated in other Chinese populations. In order to be able to measure each cognitive domain in patients with dementia in Taiwan, my proposed assessment battery will cover a range of cognitive domains.		
Research purposes	The proposed study will aim to validate a commonly used, global cognitive test (ACE-III) and other supplemental assessments via revising to the Traditional Chinese version in Taiwan according to its cultural and historical background, in order to develop an optimal cognitive battery for detecting and diagnosing patients with dementia in Taiwan. The estimated number of participants in this study is 72 in total, including 48 people in dementia group, 24 people in non-dementia group.		
Selection criteria of participants	Inclusion criteria (people who fit those conditions below are suitable to participate in this study. (1) Dementia group: patients with dementia who fulfilled the diagnostic criteria of major neurocognitive disorder (NCD) in DSM-V (American Psychiatric Association, 2013). (2) Non-dementia group: People who come to clinic in Taiwan with memory problems but do not have dementia or Mild Cognitive Impairment. (3) Aged from 45 to 100.		

	Exclusion criteria (with any of those conditions below are unable to participate in this study). (1) Have psychiatric disorder, visual, hearing or motor impairments which significantly interferes with comprehension or execution of cognitive tasks.
Methods and procedure	This study will conduct a one-time interview for 40 to 60 minutes. Before starting, you need to sign this consent form. The interview includes two parts - in the first part, you will be asked personal information regarding your name, age, language spoken etc. This is so that we can compare responses from people of different ages and with different backgrounds. In the second part, you will be asked a series of questions from tests designed to test language, memory, word finding, fluency and other brain functions. For some items, it will require to write on the test sheet and for some, the researcher will write down your answers. Before the formal evaluation starts, the researchers will provide further information.
The possible inconveniences during the interview and the solutions	We will be administering a range of cognitive tests. These may be tiring for you or may cause distress in those who find it difficult to complete these tasks. We will ensure that you are given enough time to complete tasks and are given adequate breaks to mitigate this. Some questions may cause reflection and increase realisation of cognitive deterioration which may be upsetting. The interview will be paused by the researchers according to their observation on whether you are tired or frustrated. Your family or caregivers will be informed and you will have a break till you recover, then decide whether you can continue to finish the interview or come back next time. You can contact the researcher through the contact information on the participant information sheet if you have any problem after taking part in the study. You can also contact the neurology department in Taipei hospital if you wish to speak to someone about how you are feeling.
Estimated benefits	Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will bring some new insights on the memory tests used in Taiwan, which can potentially help doctors in Taiwan to better diagnose dementia and could help people with dementia in Taiwan be diagnosed earlier and more accurately. I wish to receive a copy of the results of this study. ☐ Yes, email address_____ ☐ No.
The things the participants should do during the study	You will need to be stay relaxed during the 40- 60 minutes interview and answer the questions asked by the researchers.
Confidently	The department of mental health of older people, UCL and Neurology department, Taipei Medical University - Shuang Ho Hospital will consider your information as confidential to the extent regulated by law. UCL research ethics committee in the UK and the Ministry of health and welfare and TMU-Joint Institutional Review Board have the right to review your data and will also follow the ethics of confidentiality. Regarding the examination of the sample you provided as well as the collected data and analysis results during and after this study, the **research investigators, Gill Livingston, Chaur-Jong Hu, Chien-Tai Hong and the research assistant Ruan-Ching Yu** will abide by the obligation of confidentiality and are responsible for maintaining your privacy. We will replace your name, initials, national ID number, medical record number, phone number, address, etc.

	that can be used to identify personal information, with a code. This is so that no one can identify the person who has completed the tests. Without your consent, the investigators of this study will not reveal anything that might identify you. If the information obtained is provided to a third party or disclosure, it will be processed in a way that cannot identify your personal data.
Subsidy and compensation	The collection of the data is not part of the routine medical practice. "The department of mental health of older people, UCL" and "Neurology department, Taipei Medical University - Shuang Ho Hospital" will do their best to protect your rights during the study period, and pay additional attention to medical treatment. If injury occurred due to the implementation of this study, the investigators will provide relevant information and consultation, **"The department of mental health of older people, UCL" and "Neurology department, Taipei Medical University - Shuang Ho Hospital"** will be liable for damages according to law. Any adverse effects arising from the study will be covered under a UCL's insurance provider. Except for statutory compensation and medical care, this study does not provide other forms of compensation.
Participants' right	**(1)** If you have any questions or conditions now or during the trial / study, please feel free to contact Researcher Ruan-Ching Yu [redacted] **(2)** The original data you provided will be only remained and used in the database of Professor Gill Livingston and Doctor Chaur-Jong Hu and will not be shared to other departments. If other departments or trials / studies related to public welfare need to use your data, we will first ask for your consent again, otherwise we will never provide to them. Your personal data and privacy will be protected in accordance with relevant laws and regulations of Taiwan. **(3)** If you have any concerns about details in the research procedure, including invitation, informed consent process, interview, and personal data processing, you should contact either the researcher or the principal supervisor of this project let them know what your concerns are. If you are still unsatisfied, you can contact the Chair of the UCL Research Ethics Committee (Email: ethics@ucl.ac.uk). **(4)** If you have doubts about your rights during the trial / study or suspect that you have been harmed by participating in the trial / study, please feel free to contact TMU-Joint Institutional Review Board (Tel: (02)66382736 Ext. 1728 or Email: tmujirb@gmail.com). **(5)** In addition to the above situations, the participant or his/her law representative, guardian, supporter, person entitled to give consent, has no knowledge or is unfamiliar with the current, future or past research, and wants to discuss and ask questions, please feel free to contact TMU-Joint Institutional Review Board. We will provide a person who has no relationship with the study to provide information. If necessary, please contact TMU-Joint Institutional Review Board (Tel: +886 (02) 66382736 ext. 1728 or Email: tmujirb@gmail.com).
Withdrawal and termination	You are free to decide whether to participate in the study; you can also withdraw your consent at any time during the study and withdraw from the study without any reason, and it will not cause any unpleasantness or affect future medical care of the doctor although it will not be possible to withdraw data after it has been incorporated in the research. Your information will be stored confidentially.

	Doctor Chaur-Jong Hu and Professor Gill Livingston may also suspend or terminate the study if necessary.
Statement of person explained this consent form	(In this project: ☐ Principal investigator ☐ Co- / Sub-investigator ☐ Researcher) I have explained the nature and purpose of the above-mentioned trial / study methods in this project in detail, as well as the possible dangers and benefits, and have answered the participants' questions. **Signature of person explained this consent form:** **Date of Signature:** / /
Statement of participant	After the explanation, I have understood the above-mentioned trial / study methods and the possible dangers and benefits. I have also been provided with the detailed explanation on my questions about this trial / study project. I agree and voluntarily participate in this trial / study and will hold a copy of this consent form. **Signature of participant:** **Date of Signature:** / /
Project investigator	**Date of Signature:** / /

臺北醫學大學暨附屬醫院聯合人體研究倫理委員會

TMU-Joint Institutional Review Board

受訪者知情同意書

我們邀請您參加本研究，此份同意書提供您本研究相關資訊，計畫主持人或研究人員將為您詳細說明並回答相關問題。

計畫名稱	驗證繁體中文版之認知篩檢測驗以偵測台灣人口中失智症個案並提供健康專業人員調查		
研究成員	姓名	電話/分機	**E-mail**
主持人	胡朝榮	[redacted]	
共同主持人	洪千岱		
共同主持人	吉爾利文斯頓		
執行機構/單位	英國倫敦大學學院精神醫學所老年人心理健康學系、台灣衛生福利部雙和醫院神經內科 電話**:+44(0)2076792000, 02-22490088**		
計畫執行期限	**2020-06-01~2023-03-31**，約 **2** 年 **10** 月		
計畫簡述	台灣為一中文使用國家，且為不同文化之大熔爐。使用閩南語之河洛台灣人，占據台灣多數人口;然而，閩南語使用者相較於使用中文為第一語言者，擁有相對較低 之教育程度。過去已在中國大陸地區驗證之認知測驗，可能與台灣地區人口具有不 同的心理測量特徵。因此，本研究計畫目標在驗證一系列短版認知測驗以調查台灣可能之失智症族群並供健康專業人員參考。本研究之測驗為經過系統性文獻回顧其他中文使用族群中的有效認知測驗後進行選擇。本計畫之測驗題組將涵蓋一系列之認知領域，以測量台灣失智症個案之不同層面認知能力表現。		
研究目的	本研究計畫旨在驗證一項已被國際普遍使用之認知測驗(ACE-III)及其他補充測驗，並根據台灣特有的文化與歷史背景進行修改為繁體中文版本，致力於發展一最佳化之認知測驗題組，以協助偵測與診斷台灣失智症個案。預計參與研究之受試者共 72 人(失智組 48 人，非失智症組 24 人)。		
受訪者篩選條件	納入條件(符合下列所有條件者，適合參加本試驗/研究) (1) 失智症組:符合 DSM-V (American Psychiatric Association, 2013)之主要神經認知障礙 (major neurocognitive disorder, NCD)診斷標準之失智症個案。 (2) 非失智症組:曾經因記憶問題至診所就診，但被確認排除為失智症或輕度認知 障礙。 (3) 年齡為 45 至 100 歲。 排除條件(若有下列情況之一者，不能參加本試驗/研究) (1) 受試者有明顯影響認知測驗執行或理解之精神疾患，視覺、聽覺或動作缺損。		
方法與程序	本研究將進行一次性 40-60 分之訪談。在開始前，您需要完成簽署該份同意書。訪談包含兩部分，第一部分，您將被詢問有關姓名、年齡、使用語言等個人資料，我們才可用於比較不同年齡與不同背景的受試者。第二部分，您將被詢問一系列有關語言、記憶力、單字搜尋、流暢度及其他大腦認知功能的相關		

	問題。在回答某些題目，需要您寫在測試紙上，而對於某些題目，研究人員則會協助您寫下答案。在正式施測開始前，研究人員將提供更多訊息。
訪程序可能產生之不便及處理	我們將進行一系列認知測試，可能會使您感到疲倦，或者因感到某些題項較難回答而有挫折感。我們將確保給予您足夠時間來完成，並給予足夠休息以減輕這種情況。 一些問題可能增加對於認知功能退化事實的理解，會讓人感到沮喪。研究人員將根據觀察，看您是否疲倦或沮喪而暫停訪談，將通知您的家屬或照顧者並讓您適當休息，待心情平復後，再依個別狀況繼續施測或預約下一次再回來完成測驗。 如果參加研究後有任何問題，您可以通過「參與者信息表上的聯繫信息」與研究人員聯繫。若您想談談您的感受，也可以聯繫台北衛生福利部雙和醫院神經內科。
預期效益	即使本研究計對於參與的受試者並沒有立即的幫助，但希冀這項研究將對台灣臨床工作運用的記憶測試帶來一些新見解，從而協助醫師更有效率地診斷失智症，並幫助台灣之失智症個案在早期階段可以更準確地被診斷。 您是否希望獲得發表研究結果之副本? □是，電子信箱 _____ □否
進行中受訪者應配合事項	您將需要在 40-60 分鐘的訪談中，盡量保持放鬆，並回答研究人員的問題。
機密性	英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」將在法律上所規範的程度內將您的資料視為機密，您亦了解英國倫敦大 學學院倫理委員會、衛生福利部與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會皆有權檢視您的資料，亦會遵守保密之倫理。對於您所提供在研究過程和結束後所收集之資料及分析所得結果，Gill Livingston、胡朝榮、洪千岱研究主持人以及余阮清研究助理將會依法遵守保密義務並負責維護您的隱私。我們將以代碼取代您的姓名、姓名縮寫、國民身份證統一編號、病歷號、電話、地址等可供辨識個人資訊，任何人員皆無法辨認資料的來源。在未經您同意的情況之下，研究主持人不會洩漏任何可能辨認您的訊息。若將資料所得資訊提供予第三人或公開資料時，會以無從識別您個人資料的方式處理 。
補助與損害補償	本研究之訪談收集資料步驟，不是常規醫療作業的一部分。「英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」將盡力維護您在研究期間之權益，並善盡醫療上必要之注意。如因本研究所訂計畫步驟執行而發生傷害時，計畫主持人會提供相關訊息及諮詢。「英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」會依法負損害補償責任。本研究所產生不良後果將由英國倫敦大學學院投保之保險公司負責。提供除法定賠償及醫療照顧外，本研究不提供其他形式之賠償或補償。
受訪者權利	1. 如果您現在或於研究期間有任何問題或狀況，請不必客氣可與研究人員 余阮清 聯絡(電話:[redacted] 2. 您提供的原始資料，僅限在 Gill Livingston 教授與胡朝榮醫師資料庫中保管使用，不會連結到其他單位。如果別的單位或與社會大眾福祉有關的其它試驗/研究計畫需要使用您的資料，我們會再次徵詢您的同意，否則我們絕不會提供給他們。您的個人資料及隱私，會依我國相關法令獲得保障。 3. 如果您對研究過程中的細節有任何疑慮(包含邀請、知情同意程序、訪談與個人資料處理方式)，請透過電子郵件與研究人員或該計畫的主持人聯繫，並讓他們知道您的擔憂，若您對他們的回覆仍不滿意，可進一步聯繫 UCL 研究倫理委員會執行長– ethics@ucl.ac.uk

	4. 如果您在試驗/研究過程中對您的權益有疑義或認為因參與試驗/研究而受害時，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡 (電話:(02) 66382736 分機 1728 或電子郵件信箱:tmujirb@gmail.com)。 5. 除上述情形外，受試者或其法定代理人、監護人、輔助人或有同意權人，對目前、未來或過去的研究有任何不了解、熟悉之處，想要討論及解答疑問，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡，我們將會提供一位與研究無關係的人員來提供資訊，如有需要，請與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡 (電話: (02)66382736 分機 1728 或 Email: tmujirb@gmail.com)。
退出與中止	您可自由決定是否參加本研究; 您也可在研究過程中隨時撤銷同意，退出研究，不需任何理由，且不會引起任何不愉快或影響日後醫師對您的醫療照顧，儘管當資料已被納入研究後將無法撤消，但您的個人資料將會被妥善保存。胡朝榮醫師與 Gill Livingston 教授亦可能於必要時中止或終止本研究之進行。
解釋同意書人聲明	(於本計畫中擔任: □主持人 □共/協同主持人□研究人員) 本人已詳細解釋本計畫中上述試驗/研究方法的性質與目的，及可能產生的危險與利益，並已回答受訪者之疑問。 解釋同意書人簽名:　　簽名日期:　年　月　日
受訪者聲明	經由說明後本人已詳細瞭解上述試驗/研究方法及可能產生的危險與利益，有關本試驗/研究計畫的疑問，亦獲得詳細解釋。本人同意並自願參與本試驗/研究，且將持有同意書副本。 受訪者簽名:　　簽名日期:　年　月　日 出生日期:　年　月　日
計畫主持人	簽名日期:　年　月　日

Appendix 27: T-ACE-III study informed consent form for interviewees and proxy (in English & traditional Chinese)

TMU-Joint Institutional Review Board

Informed Consent Form for participants (Interviewees and Proxy)

We invite you and your family /friend to participate in this study. This consent form provides you with relevant information about this project. The investigators of this study or researcher will explain and answer any related questions for you.

The title of this study	**Validation of cognitive screening tests in Traditional Chinese to detect dementia in the Taiwanese population presenting to health professionals for investigation of possible dementia**		
Project members	**name**	**Telephone number**	**E-mail**
Principal investigator	**Chaur-Jong Hu**	[redacted]	
Co-investigator	**Chien-Tai Hong**		
Co-investigator	**Gill Livingston**		
Organizers and contact number	**The department of Mental health of older people, Division of psychiatry, University College London; Neurology department, Taipei Medical University - Shuang Ho Hospital.** **Contact number: +44 (0)2076792000, +886 (0)222490088**		
Conduct period	**1st June 2020 – 31st March 2023, approximately 2 year(s) 10 month(s).**		
Brief introduction of this project	Taiwan is a Chinese-speaking country and is a melting pot of different cultures. Hoklo Taiwanese, who speak the Hokkien language, make up the majority of the population. Hokkien speakers have a comparatively lower education level than people who speak Chinese as their first language. Cognitive tests which are validated in mainland Chinese populations may, consequently, have different psychometric properties in Taiwan. This proposal is therefore to validate a battery of brief cognitive tests for investigating people with possible dementia presenting to health professionals in Taiwan. The tests have been chosen after systematic review of tests available and validated in other Chinese populations. In order to be able to measure each cognitive domain in patients with dementia in Taiwan, my proposed assessment battery will cover a range of cognitive domains.		
Research purposes	The proposed study will aim to validate a commonly used, global cognitive test (ACE-III) and other supplemental assessments via revising to the Traditional Chinese version in Taiwan according to its cultural and historical background, in order to develop an optimal cognitive battery for detecting and diagnosing patients with dementia in Taiwan. The estimated number of participants in this study is 72 in total, including 48 people in dementia group, 24 people in non-dementia group.		
Selection criteria of participants	Inclusion criteria (people who fit those conditions below are suitable to participate in this study. (1) Dementia group: patients with dementia who fulfilled the diagnostic criteria of major neurocognitive disorder (NCD) in DSM-V (American Psychiatric Association, 2013). (2) Non-dementia group: People who come to clinic in Taiwan with memory problems but do not have dementia or Mild Cognitive Impairment. (3) Aged from 45 to 100.		

	Exclusion criteria (with any of those conditions below are unable to participate in this study). (1) Have psychiatric disorder, visual, hearing or motor impairments which significantly interferes with comprehension or execution of cognitive tasks.
Methods and procedure	This study will conduct a one-time interview for 40 to 60 minutes. Before starting, you and your family /friend need to sign this consent form. The interview includes two parts - in the first part, your family /friend will be asked personal information regarding their name, age, language spoken etc. This is so that we can compare responses from people of different ages and with different backgrounds. In the second part, your family /friend will be asked a series of questions from tests designed to test language, memory, word finding, fluency and other brain functions. For some items, it will require to write on the test sheet and for some, the researcher will write down answers. Before the formal evaluation starts, the researchers will provide further information.
The possible inconveniences during the interview and the solutions	We will be administering a range of cognitive tests. These may be tiring for your family /friend or may cause distress in those who find it difficult to complete these tasks. We will ensure enough time to complete tasks is given and are adequate breaks given to mitigate this. Some questions may cause reflection and increase realisation of cognitive deterioration which may be upsetting. The interview will be paused by the researchers according to their observation on whether your family /friend is tired or frustrated. You will be informed and your family /friend will have a break till they recover, then decide whether your family /friend can continue to finish the interview or come back next time. You can contact the researcher through the contact information on the participant information sheet if you have any problem after taking part in the study. You can also contact the neurology department in Taipei hospital if you wish to speak to someone about how you are feeling.
Estimated benefits	Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will bring some new insights on the memory tests used in Taiwan, which can potentially help doctors in Taiwan to better diagnose dementia and could help people with dementia in Taiwan be diagnosed earlier and more accurately. I wish to receive a copy of the results of this study. ☐ Yes, email address_____ ☐ No.
The things the participants should do during the study	Your family /friend will need to be stay relaxed during the 40- 60 minutes interview and answer the questions asked by the researchers.
Confidently	The department of mental health of older people, UCL and Neurology department, Taipei Medical University - Shuang Ho Hospital will consider your information as confidential to the extent regulated by law. UCL research ethics committee in the UK and the Ministry of health and welfare and TMU-Joint Institutional Review Board have the right to review your and your family's /friend's data and will also follow the ethics of confidentiality. Regarding the examination of the sample you and your family /friend provided as well as the collected data and analysis results during and after this study, the **research investigators, Gill Livingston, Chaur-Jong Hu, Chien-Tai Hong and the research assistant Ruan-Ching Yu** will abide by the obligation of confidentiality and are responsible for maintaining your privacy. We will replace your family's /friend's name, initials, national ID number, medical record number, phone number, address, etc. that can be used to identify personal information, with a code. This is so that no one can identify

	the person who has completed the tests. Without your consent, the investigators of this study will not reveal anything that might identify you and your family /friend. If the information obtained is provided to a third party or disclosure, it will be processed in a way that cannot identify your personal data.
Subsidy and compensation	The collection of the data is not part of the routine medical practice. "The department of mental health of older people, UCL" and "Neurology department, Taipei Medical University - Shuang Ho Hospital" will do their best to protect your rights during the study period, and pay additional attention to medical treatment. If injury occurred due to the implementation of this study, the investigators will provide relevant information and consultation, **"The department of mental health of older people, UCL" and "Neurology department, Taipei Medical University - Shuang Ho Hospital"** will be liable for damages according to law. Any adverse effects arising from the study will be covered under a UCL's insurance provider. Except for statutory compensation and medical care, this study does not provide other forms of compensation.
Participants' right	**(1)** If you have any questions or conditions now or during the trial / study, please feel free to contact Researcher Ruan-Ching Yu [redacted] **(2)** The original data you and your family /friend provided will be only remained and used in the database of Professor Gill Livingston and Doctor Chaur-Jong Hu and will not be shared to other departments. If other departments or trials / studies related to public welfare need to use your and your family's /friend's data, we will first ask for your consent again, otherwise we will never provide to them. Your and your family's /friend's personal data and privacy will be protected in accordance with relevant laws and regulations of Taiwan. **(3)** If you have any concerns about details in the research procedure, including invitation, informed consent process, interview, and personal data processing, you should contact either the researcher or the principal supervisor of this project let them know what your concerns are. If you are still unsatisfied, you can contact the Chair of the UCL Research Ethics Committee (Email: ethics@ucl.ac.uk). **(4)** If you have doubts about your and your family's /friend's rights during the trial / study or think that you have been harmed by participating in the trial / study, please feel free to contact TMU-Joint Institutional Review Board (Tel: (02)66382736 Ext. 1728 or Email: tmujirb@gmail.com). **(5)** In addition to the above situations, the participant or his/her law representative, guardian, supporter, person entitled to give consent, has no knowledge or is unfamiliar with the current, future or past research, and wants to discuss and ask questions, please feel free to contact TMU-Joint Institutional Review Board. We will provide a person who has no relationship with the study to provide information. If necessary, please contact TMU-Joint Institutional Review Board (Tel: +886 (02) 66382736 ext. 1728 or Email: tmujirb@gmail.com).
Withdrawal and termination	You and your family /friend are free to decide whether to participate in the study; you can also withdraw your consent at any time during the study and withdraw from the study without any reason, and it will not cause any unpleasantness or affect future medical care of the doctor although it will not be possible to withdraw data after it has been incorporated in the research. Your and your family's /friend's information will be stored confidentially. Doctor Chaur-Jong Hu and Professor Gill Livingston may also suspend or terminate the study if necessary.

Statement of person explained this consent form	(In this project: ☐ Principal investigator ☐ Co- / Sub-investigator ☐ Researcher) I have explained the nature and purpose of the above-mentioned trial / study methods in this project in detail, as well as the possible dangers and benefits, and have answered the participants' questions. **Signature of person explained this consent form:** **Date of Signature:** / /
Statement of participant	After the explanation, I have understood the above-mentioned trial / study methods and the possible dangers and benefits. I have also been provided with the detailed explanation on my questions about this trial / study project. I agree and voluntarily participate in this trial / study and will hold a copy of this consent form. **Signature of participant:** **Date of Signature:** / /
Law representative or person entitled to give consent	**Law representative, guardian, supporter, person entitled to give consent please confirm and tick on the following things:** ☐ I am entitled to act as the participant's legal representative according the law in Taiwan. ☐ I have had a full explanation by the researcher regarding the research purpose and procedures. ☐ I fully understand and agree on how our personal data will be used and stored in this research. ☐ I agree to act as the participant's legal representative. ☐ I give consent on behalf on the participant to take part in this research. **Signature:** **Relationship with participant** **(for those with consent rights, please check according to the regulations of this trial / study)**: **Date of Signature:** / / **National ID:** **Contact number:** **Contact address:** *Note: Participants are incapacitated (minors under seven years of age or under guardianship declaration), and are represented by law representative or guardians; persons with restricted capacity (minors over seven years of age or under guardianship declaration), It shall be mutually agreed by the participant and his / her law representative or supporter, and and signed the consent form.*
Project investigator	**Date of Signature:** / /

臺北醫學大學暨附屬醫院聯合人體研究倫理委員會

TMU-Joint Institutional Review Board

受訪者與代理人知情同意書

我們邀請您與您的親人/朋友參加本研究，此份同意書提供您本研究相關資訊，計畫主持人或研究人員將為您詳細說明並回答相關問題。

計畫名稱	驗證繁體中文版之認知篩檢測驗以偵測台灣人口中失智症個案並提供健康專業人員調查		
研究成員	姓名	電話/分機	**E-mail**
主持人	胡朝榮	[redacted]	
共同主持人	洪千岱		
共同主持人	吉爾利文斯頓		
執行機構/單位	英國倫敦大學學院精神醫學所老年人心理健康學系、台灣衛生福利部雙和醫院神經內科 電話**:+44(0)2076792000, 02-22490088**		
計畫執行期限	**2020-06-01~2023-03-31**，約 **2** 年 **10** 月		
計畫簡述	台灣為一中文使用國家，且為不同文化之大熔爐。使用閩南語之河洛台灣人，占據台灣多數人口;然而，閩南語使用者相較於使用中文為第一語言者，擁有相對較低 之教育程度。過去已在中國大陸地區驗證之認知測驗，可能與台灣地區人口具有不 同的心理測量特徵。因此，本研究計畫目標在驗證一系列短版認知測驗以調查台灣可能之失智症族群並供健康專業人員參考。本研究之測驗為經過系統性文獻回顧其他中文使用族群中的有效認知測驗後進行選擇。本計畫之測驗題組將涵蓋一系列之認知領域，以測量台灣失智症個案之不同層面認知能力表現。		
研究目的	本研究計畫旨在驗證一項已被國際普遍使用之認知測驗(ACE-III)及其他補充測驗，並根據台灣特有的文化與歷史背景進行修改為繁體中文版本，致力於發展一最佳化之認知測驗題組，以協助偵測與診斷台灣失智症個案。預計參與研究之受試者共 72 人(失智組 48 人，非失智症組 24 人)。		
受訪者篩選條件	納入條件(符合下列所有條件者，適合參加本試驗/研究) (1) 失智症組:符合 DSM-V (American Psychiatric Association, 2013)之主要神經認 知障礙 (major neurocognitive disorder, NCD)診斷標準之失智症個案。 (2) 非失智症組:曾經因記憶問題至診所就診，但被確認排除為失智症或輕度認知 障礙。 (3) 年齡為 45 至 100 歲。 排除條件(若有下列情況之一者，不能參加本試驗/研究) (1) 受試者有明顯影響認知測驗執行或理解之精神疾患，視覺、聽覺或動作缺損。		
方法與程序	本研究將進行一次性 40-60 分之訪談。在開始前，您與您的親人/朋友需要完成簽署該份同意書。訪談包含兩部分，第一部分，您的親人/朋友將被詢問有關姓名、年齡、使用語言等個人資料，我們才可用於比較不同年齡與不同背景的受試者。第二部分，您的親人/朋友將被詢問一系列有關語言、記憶力、單字搜尋、流暢度及其他大腦認知功能的相關問題。在回答某些題目，需要您的親人/朋友寫在測試紙上，而對於某些題目，研究人員則會協助寫下答案。在正式施測開始前，研究人員將提供更多訊息。		

訪程序可能產生之不便及處理	我們將進行一系列認知測試，可能會使您的親人/朋友感到疲倦，或者因感到某些題項較難回答而有挫折感。我們將確保給予足夠時間來完成，並給予足夠休息以減輕這種情況。 一些問題可能增加對於認知功能退化事實的理解，會讓人感到沮喪。研究人員將根據觀察，看您的親人/朋友是否疲倦或沮喪而暫停訪談，將通知您並讓其適當休息，待心情平復後，再依個別狀況繼續施測或預約下一次再回來完成測驗。如果參加研究後有任何問題，您可以通過「參與者信息表上的聯繫信息」與研究人員聯繫。若您想談談您的感受，也可以聯繫台北衛生福利部雙和醫院神經內科。
預期效益	即使本研究計對於參與的受試者並沒有立即的幫助，但希冀這項研究將對台灣臨床工作運用的記憶測試帶來一些新見解，從而協助醫師更有效率地診斷失智症，並幫助台灣之失智症個案在早期階段可以更準確地被診斷。 您是否希望獲得發表研究結果之副本? □是，電子信箱 _____ □否
進行中受訪者應配合事項	您的親人/朋友將需要在 40-60 分鐘的訪談中，盡量保持放鬆，並回答研究人員的問題。
機密性	英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」將在法律上所規範的程度內將您與您的親人/朋友的資料視為機密，您亦了解英國倫敦大學學院倫理委員會、衛生福利部與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會皆有權檢視您與您的親人/朋友的資料，亦會遵守保密之倫理。對於您所提供在研究過程和結束後所收集之資料及分析所得結果，Gill Livingston、胡朝榮、洪千岱研究主持人以及余阮清研究助理將會依法遵守保密義務並負責維護您們的隱私。我們將以代碼取代您的親人/朋友的姓名、姓名縮寫、國民身份證統一編號、病歷號、電話、地址等可供辨識個人資訊，任何人員皆無法辨認資料的來源。在未經您們同意的情況之下，研究主持人不會洩漏任何可能辨認您們的訊息。若將資料所得資訊提供予第三人或公開資料時，會以無從識別您們個人資料的方式處理 。
補助與損害補償	本研究之訪談收集資料步驟，不是常規醫療作業的一部分。「英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」將盡力維護您們在研究期間之權益，並善盡醫療上必要之注意。如因本研究所訂計畫步驟執行而發生傷害時，計畫主持人會提供相關訊息及諮詢。「英國倫敦大學學院精神醫學所老年人心理健康學系」與「台灣衛生福利部雙和醫院神經內科」會依法負損害補償責任。本研究所產生不良後果將由英國倫敦大學學院投保之保險公司負責。提供除法定賠償及醫療照顧外，本研究不提供其他形式之賠償或補償。
受訪者權利	1. 如果您現在或於研究期間有任何問題或狀況，請不必客氣可與研究人員 余阮清聯絡[redacted] 2. 您與您的親人/朋友提供的原始資料，僅限在 Gill Livingston 教授與胡朝榮醫師資料庫中保管使用，不會連結到其他單位。如果別的單位或與社會大眾福祉有關的其它試驗/研究計畫需要使用您們的資料，我們會再次徵詢您們的同意，否則我們絕不會提供給他們。您與您的親人/朋友的個人資料及隱私，會依我國相關法令獲得保障。 3. 如果您對研究過程中的細節有任何疑慮(包含邀請、知情同意程序、訪談與個人資料處理方式)，請透過電子郵件與研究人員或該計畫的主持人聯繫，並讓他們知道您的擔憂，若您對他們的回覆仍不滿意，可進一步聯繫 UCL 研究倫理委員會執行長– ethics@ucl.ac.uk

	4. 如果您在試驗/研究過程中對您與您的親人/朋友的權益有疑義或認為因參與試驗/研究而受害時，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡 (電話:(02) 66382736 分機 1728 或電子郵件信箱:tmujirb@gmail.com)。 5. 除上述情形外，受試者或其法定代理人、監護人、輔助人或有同意權人，對目前、未來或過去的研究有任何不了解、熟悉之處，想要討論及解答疑問，請隨時與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡，我們將會提供一位與研究無關係的人員來提供資訊，如有需要，請與臺北醫學大學暨附屬醫院聯合人體研究倫理委員會聯絡 (電話: (02)66382736 分機 1728 或 Email: tmujirb@gmail.com)。
退出與中止	您與您的親人/朋友可自由決定是否參加本研究; 您們也可在研究過程中隨時撤銷同意，退出研究，不需任何理由，且不會引起任何不愉快或影響日後醫師對您的醫療照顧，儘管當資料已被納入研究後將無法撤消，但您與您的親人/朋友的個人資料將會被妥善保存。胡朝榮醫師與 Gill Livingston 教授亦可能於必要時中止或終止本研究之進行。
解釋同意書人聲明	(於本計畫中擔任: □主持人 □共/協同主持人□研究人員) 本人已詳細解釋本計畫中上述試驗/研究方法的性質與目的，及可能產生的危險與利益，並已回答受訪者之疑問。 解釋同意書人簽名: 簽名日期: 年 月 日
受訪者聲明	經由說明後本人已詳細瞭解上述試驗/研究方法及可能產生的危險與利益，有關本試驗/研究計畫的疑問，亦獲得詳細解釋。本人同意並自願參與本試驗/研究，且將持有同意書副本。 受訪者簽名: 簽名日期: 年 月 日 出生日期: 年 月 日
法定代理人或有同意權人	請法定代理人、監護人、輔助人或有同意權人依序確認並勾選下列事項: □ 根據台灣法律規定，我有權利作為受訪者之法定代理人。 □ 研究人員已針對研究目的及程序向我進行充分的解釋。 □ 我瞭解並同意我與受訪者的個人資料在該研究之使用與保存方式。 □ 我同意擔任受訪者的法定代理人。 □ 我代表受訪者表示同意參加該項研究。 簽名: 與受訪者之關係(有同意權人請依試驗/研究所屬法規查詢): 簽名日期: 年 月 日 國民身份證統一編號: 聯絡電話: 通訊地址: *註:受訪者為無行為能力(未滿七歲之未成年人或受監護宣告)，由法定代理人或監護人為之;限制行為能力人(滿七歲以上之未成年人或受輔助宣告)，應得其本人及法定代理人或輔助人共同同意並簽署同意書。*
計畫主持人	簽名日期: 年 月 日

Appendix 28: T-ACE-III study- T-ACE-III questionnaire (UK & Taiwanese versions)

ADDENBROOKE'S COGNITIVE EXAMINATION – ACE-III
UK Version A (2012)

Name:	Date of testing: ____/___/___
Date of Birth:	Tester's name: _____
Hospital No. or Address:	Age at leaving full-time education: _____
	Occupation: _____
	Handedness: _____

ATTENTION *(Sum together only the items in BOLD for the M-ACE score)

▯ Ask: What is the	**Day**	**Date**	**Month**	**Year**	Season	**Attention** [Score 0-5] *
▯ Ask: Which	No./Floor	Street/Hospital	Town	County	Country	**Attention** [Score 0-5]

ATTENTION

▯ Tell: "I'm going to give you three words and I'd like you to repeat them after me: lemon, key and ball." After subject repeats, say "Try to remember them because I'm going to ask you later".
▯ Score *only* the first trial (repeat 3 times if necessary).
▯ Register number of trials: _____

Attention [Score 0-3]

ATTENTION

▯ Ask the subject: "Could you take 7 away from 100? I'd like you to keep taking 7 away from each new number until I tell you to stop."
▯ If subject makes a mistake, do not stop them. Let the subject carry on and check subsequent answers (e.g., 93, 84, 77, 70, 63 – score 4).
▯ Stop after five subtractions (93, 86, 79, 72, 65): ____ ____ ____ ____ ____

Attention [Score 0-5]

MEMORY

▯ Ask: 'Which 3 words did I ask you to repeat and remember?' _____ _____ _____

Memory [Score 0-3]

FLUENCY

▯ **Letters**
Say: "I'm going to give you a letter of the alphabet and I'd like you to generate as many words as you can beginning with that letter, but not names of people or places. For example, if I give you the letter "C", you could give me words like "cat, cry, clock" and so on. But, you can't give me words like Catherine or Canada. Do you understand? Are you ready? You have one minute. The letter I want you to use is the letter "P".

Fluency [Score 0 – 7]

≥ 18	7
14-17	6
11-13	5
8-10	4
6-7	3
4-5	2
2-3	1
0-1	0
total	correct

▯ **Animals**
Say: "Now can you name as many animals as possible. It can begin with any letter."

Fluency [Score 0 – 7]

≥ 22	7
17-21	6
14-16	5
11-13	4
9-10	3
7-8	2
5-6	1
<5	0
total	correct

MEMORY

- Tell: "I'm going to give you a name and address and I'd like you to repeat the name and address after me. So you have a chance to learn, we'll be doing that 3 times. I'll ask you the name and address later."

Score only the third trial.

	1st Trial	*2nd Trial*	*3rd Trial*
Harry Barnes 73 Orchard Close Kingsbridge Devon	_____ _____ ___ ____ _____ _____ _____	_____ _____ ___ ____ _____ _____ _____	_____ _____ ___ ____ _____ _____ _____

Memory [Score 0 – 7] ☐☐

MEMORY

- Name of the current Prime Minister.....
- Name of the first female Prime Minister .....
- Name of the USA president.....
- Name of the USA president who was assassinated in the 1960s.....

Memory [Score 0 – 4] ☐

LANGUAGE

- Place a pencil and a piece of paper in front of the subject. As a practice trial, ask the subject to "**Pick up the pencil and then the paper**." If incorrect, score 0 and do not continue further.

- If the subject is correct on the practice trial, continue with the following three commands below.
  - ∀ Ask the subject to "**Place the paper on top of the pencil**"
  - ∀ Ask the subject to "**Pick up the pencil but not the paper**"
  - ∀ Ask the subject to "**Pass me the pencil after touching the paper**"

  Note: Place the pencil and paper in front of the subject before each command.

Language [Score 0-3] ☐

LANGUAGE

- Say: "I want you to write two sentences. It can be about anything that you like. I want you to write in full sentences and avoid abbreviations." If the subject does not know what to write about, you could suggest a few topics. "For instance, you could write about a recent holiday, your hobbies, your family or childhood." If the subject writes only one sentence, then prompt for a second one.

  Sentences must contain a subject and a verb. Spelling and grammar are penalized. Sentences do not need to be about the same topic. See scoring guidelines for more information.

Language [Score 0-2] ☐

LANGUAGE

- Ask the subject to repeat: **'caterpillar'; 'eccentricity; 'unintelligible'; 'statistician'**

Score 2 if all are correct; score 1 if 3 are correct; and score 0 if 2 or less are correct.

Language [Score 0-2] ☐

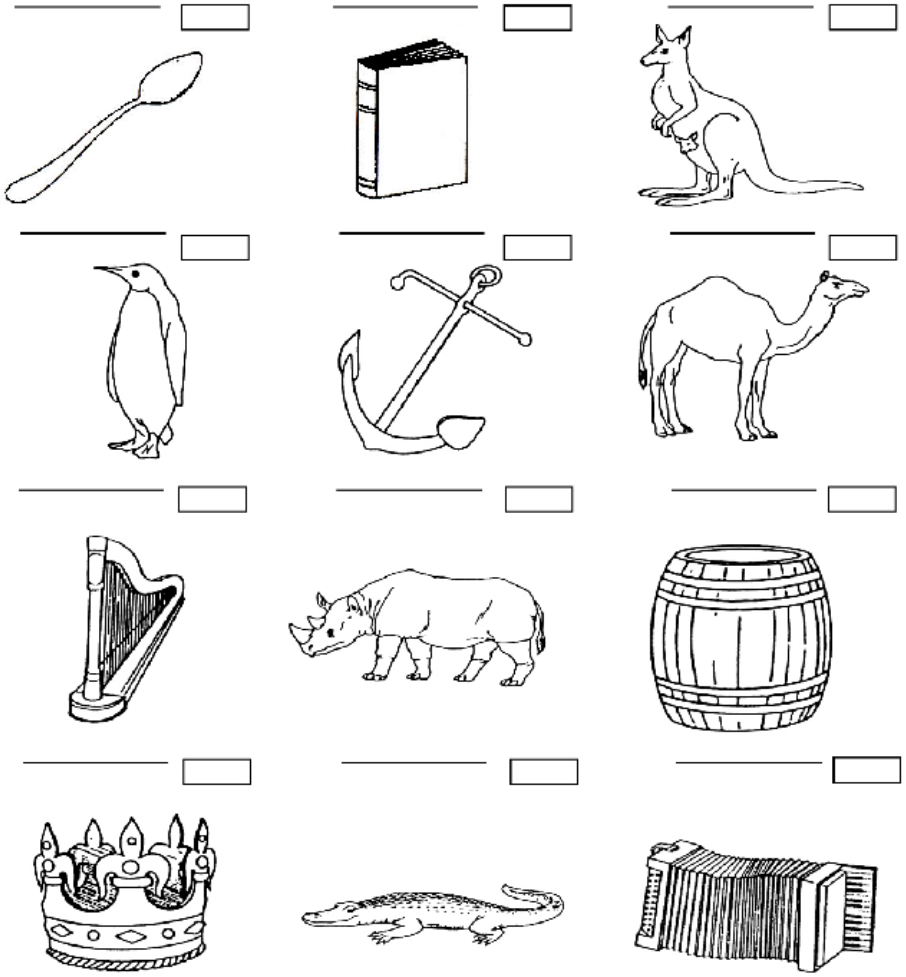
Updated 11/08/2014

LANGUAGE

- Ask the subject to repeat: **'All that glitters is not gold'** — **Language** [Score 0-1] ☐
- Ask the subject to repeat: **'A stitch in time saves nine'** — **Language** [Score 0-1] ☐

LANGUAGE

- Ask the subject to name the following pictures: — **Language** [Score 0-12] ☐



LANGUAGE

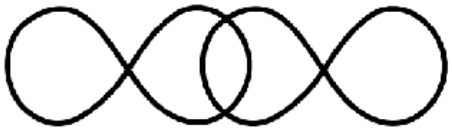
- Using the pictures above, ask the subject to: — **Language** [Score 0-4] ☐
  - ∀ Point to the one which is associated with the monarchy ……
  - ∀ Point to the one which is a marsupial ……
  - ∀ Point to the one which is found in the Antarctic ……
  - ∀ Point to the one which has a nautical connection ……

LANGUAGE

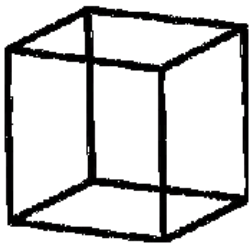
▯ Ask the subject to read the following words: (Score 1 only if all correct) **sew** **pint** **soot** **dough** **height**	**Language** [Score 0-1]

VISUOSPATIAL ABILITIES

▯ Infinity Diagram: Ask the subject to copy this diagram.	**Visuospatial** [Score 0-1]



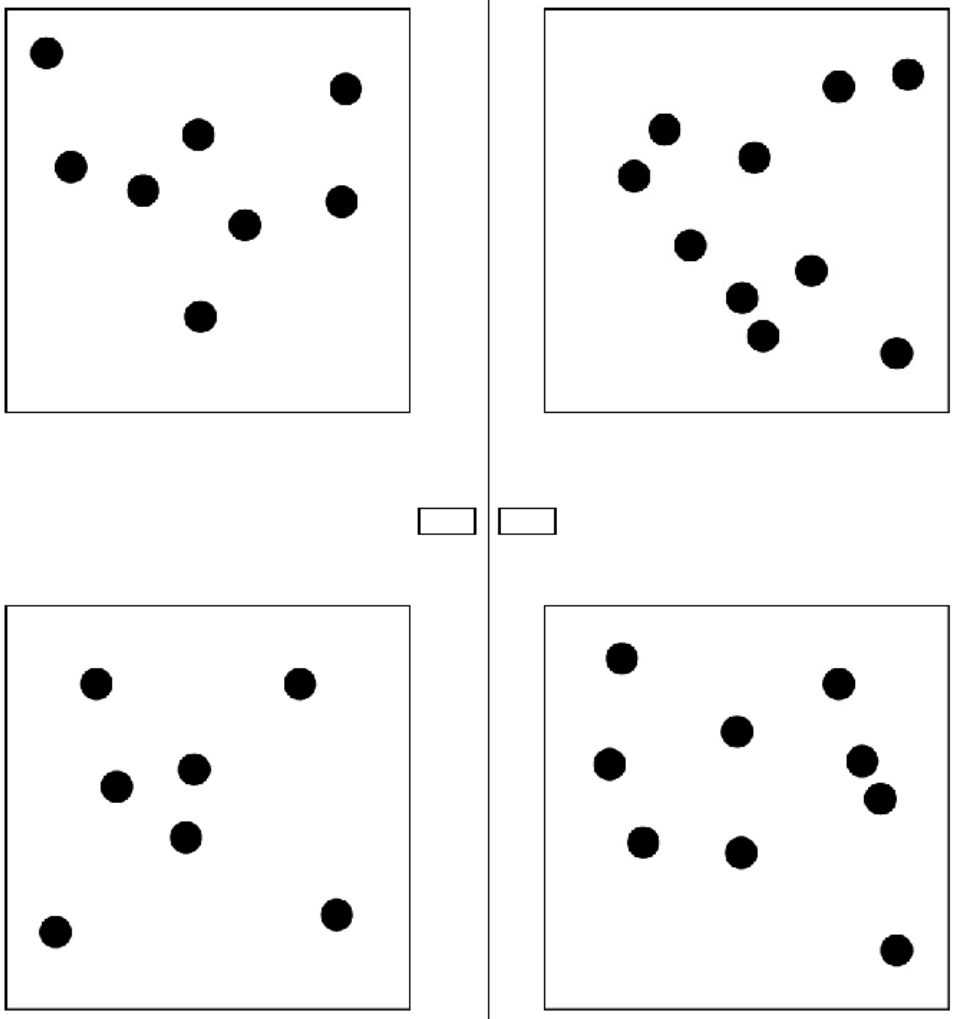
▯ Wire cube: Ask the subject to copy this drawing (for scoring, see instructions guide).	**Visuospatial** [Score 0-2]



▯ Clock: Ask the subject to draw a clock face with numbers. Then, ask the subject to put the hands at ten past five. (For scoring see instruction guide: circle = 1, numbers = 2, hands = 2 if all correct).	**Visuospatial** [Score 0-5]

VISUOSPATIAL ABILITIES

▯ Ask the subject to count the dots without pointing to them	**Visuospatial** [Score 0-4]

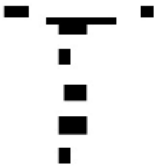


Updated 11/08/2014

VISUOSPATIAL ABILITIES

- Ask the subject to identify the letters

Visuospatial
[Score 0-4]



MEMORY

- Ask "Now tell me what you remember about that name and address we were repeating at the beginning"

Harry Barnes	....	**Memory** [Score 0-7]
73 Orchard Close	....	
Kingsbridge	.....	
Devon	.....	

MEMORY

- This test should be done if the subject failed to recall one or more items above. If all items were recalled, skip the test and score 5. If only part was recalled start by ticking items recalled in the shadowed column on the right hand side; and then test not recalled items by telling the subject "ok, I'll give you some hints: was the name X, Y or Z?" and so on. Each recognised item scores one point, which is added to the point gained by recalling.

Memory
[Score 0-5]

Jerry Barnes	Harry Barnes		Harry Bradford		recalled	
37	73		76		recalled	
Orchard Place	Oak Close		Orchard Close		recalled	
Oakhampton	Kingsbridge		Dartington		recalled	
Devon	Dorset		Somerset		recalled	

SCORES

TOTAL ACE-III SCORE	/100
TOTAL M-ACE SCORE	/30
Attention	/18
Memory	/26
Fluency	/14
Language	/26
Visuospatial	/16

Updated 11/08/2014

ADDENBROOKE'S COGNITIVE EXAMINATION - ACE-III
ADDENBROOKE'S 認知功能測驗（2020）

受試者編號: 施測人員: 受測日期: _____年_____月_____日

注意力

➢ 問：今天是?	幾年 _____	幾月 _____	幾日 _____	星期幾 _____	什麼季節 _____	注意力 [得分0-5] ☐
➢ 問：你現在在?	哪個國家 _____	哪個縣/市 _____	哪個鄉/鎮/區 _____	街道/醫院 _____	樓層/號碼 _____	注意力 [得分0-5] ☐

注意力

➢ 告訴受試者："我說三個字詞，我說完了請你重複一遍：檸檬、鑰匙和皮球"。受試者重複完後，告訴他/她："記住這三個詞，等一下我還會問你。
➢ "只對第一次檢查計分（必要的話可以重複三次）
➢ 重複次數_____

注意力 [得分0-3] ☐

注意力

➢ 問受試者：100 減去 7 是多少？得到的數字再減去 7 是多少？連續做五次。
➢ 如果受試者中間算錯的話，可以繼續進行，將錯就錯，只要下一個答案正確就給分（如 93，84，77，70，63，得分為 4 分，雖然第二次（93-84）答案錯誤，但後 3 次按照 84 依次減 7 都是對的）。
➢ 五次減法算完後停止計算（93，86，79，72，65）____ ____ ____ ____ ____

注意力 [得分0-5] ☐

記憶力

➢ 問受試者：我讓你重複並記住的那三個詞是什麼？ _____ _____ _____

記憶力 [得分0-3] ☐

語言流暢性

➢ 水果和蔬菜
➢ 告訴受試者：請你盡量說出水果和蔬菜的名稱，越多越好。你只有一分鐘的時間，現在開始。

語言流暢性 [得分0-7] ☐

≥22	7
17-21	6
14-16	5
11-13	4
9-10	3
7-8	2
5-6	1
<5	0
總數	正確的

➢ 動物
告訴受試者：請你盡量說出動物的名稱，越多越好。你只有一分鐘的時間，現在開始。

語言流暢性 [得分0-7] ☐

≥22	7
17-21	6
14-16	5
11-13	4
9-10	3
7-8	2
5-6	1
<5	0
總數	正確的

© John R Hodges, translated by Ruan-Ching Yu,& Cheng-Chang Yang, 2020.6

記憶力	
➢ 告訴受試者：我會說一個名字和地址，我說完之後你要重複一遍。我們會重複三次，所以你要試著把它記起來，我之後會再問你一次（只對第三次檢查計分）	記憶力 [得分0-7] ☐

	第一次檢查	第二次檢查	第三次檢查
王志偉	___ ___ ___	___ ___ ___	___ ___ ___
高雄市	___	___	___
楠梓區	___	___	___
創新路 16 號	___ ___	___ ___	___ ___

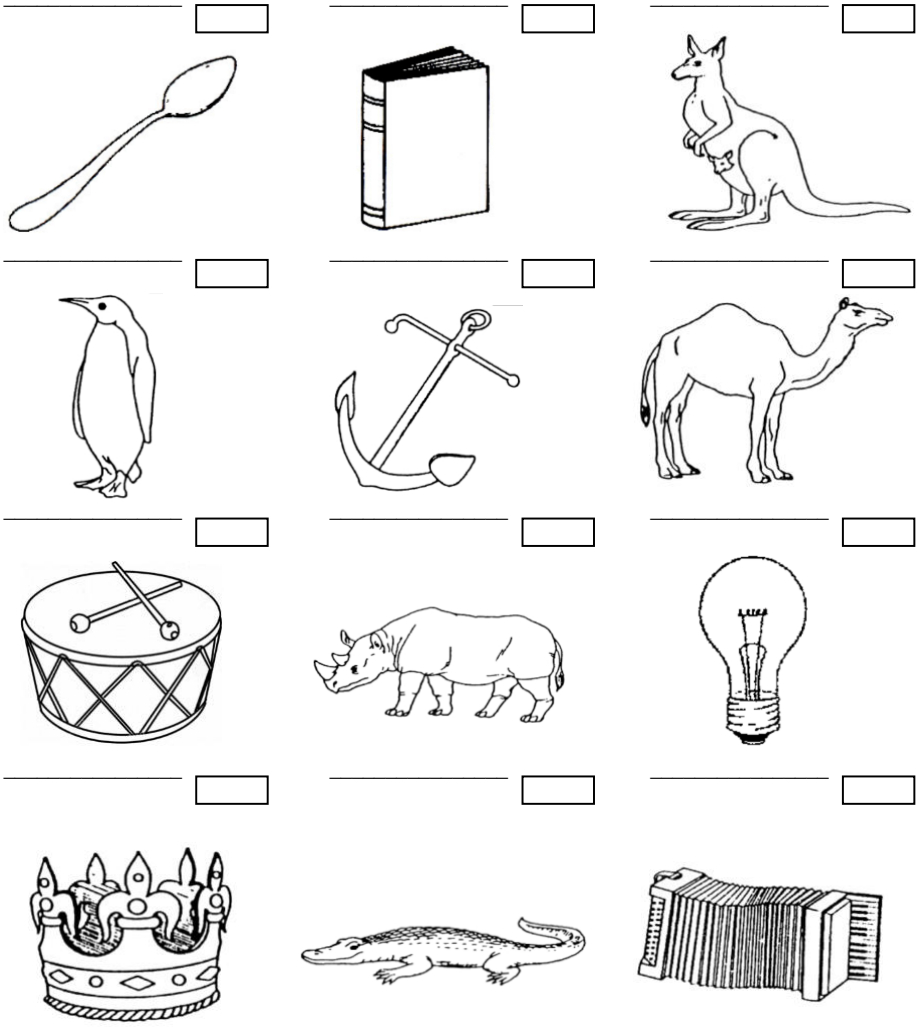
記憶力	
➢ 現任台灣總統的名字 _____ ➢ 第一位民選總統的名字 _____ ➢ 現任美國總統的名字 _____ ➢ 台灣歷史上在國內遭到槍擊受傷的總統的名字 _____	記憶力 [得分0-4] ☐
語言	
➢ 在受試者前放一隻鉛筆和一張紙，然後，跟受試者說“請您拿起鉛筆，然後再拿起紙”。如果受試者不能正確完成這個指令，就計 0 分，不繼續檢查。 ➢ 如果受試者第一個問題回答正確，請繼續讓受試者完成以下三個動作： • 請把紙放在鉛筆的上面 • 請拿起鉛筆，但不要拿紙 • 請摸一下紙後，再把鉛筆拿給我	語言 [得分0-3] ☐
語言	
➢ 要求受試者至少寫出兩個關於他/她的近期活動的句子。要求受試者寫出完整的句子，不要用簡體字。如果受試者圍繞上述某個主題至少寫出兩個句子，就計 1 分。如果句子的語法和拼寫都正確，再計 1 分。	語言 [得分0-2] ☐
語言	
➢ 讓受試者重複（國語）：向聲背實；白雲蒼狗；去偽存真；囊螢照書 ➢ 讓受試者重複（台語）：鴨子聽雷；食飽換枵；剖雞教猴；軟土深掘 4 成語都正確的話給 2 分，3 個正確給 1 分，2 個(或少於)正確給 0 分。	語言 [得分0-2] ☐
➢ 讓受試者重複（國語）：春江水暖鴨先知 ➢ 讓受試者重複（台語）：一樣米飼百樣人	語言 [得分0-1] ☐
➢ 讓受試者重複（國語）：業精於勤荒於嬉 ➢ 讓受試者重複（台語）：七坐八爬九發牙	語言 [得分0-1] ☐

© John R Hodges, translated by Ruan-Ching Yu,& Cheng-Chang Yang, 2020.6

語言

讓受試者說出以下圖片的名稱：

語言
[得分0-12]



語言 — 理解

➢ 用上面的圖片，讓受試者說出：

語言
[得分0-4]

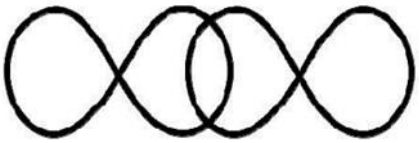
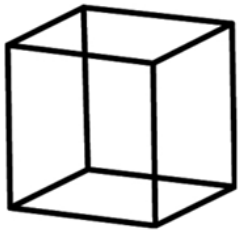
指出哪個動物為有袋動物? _____

指出哪個動物為爬蟲類？ _____

指出哪個動物生活在沙漠？ _____

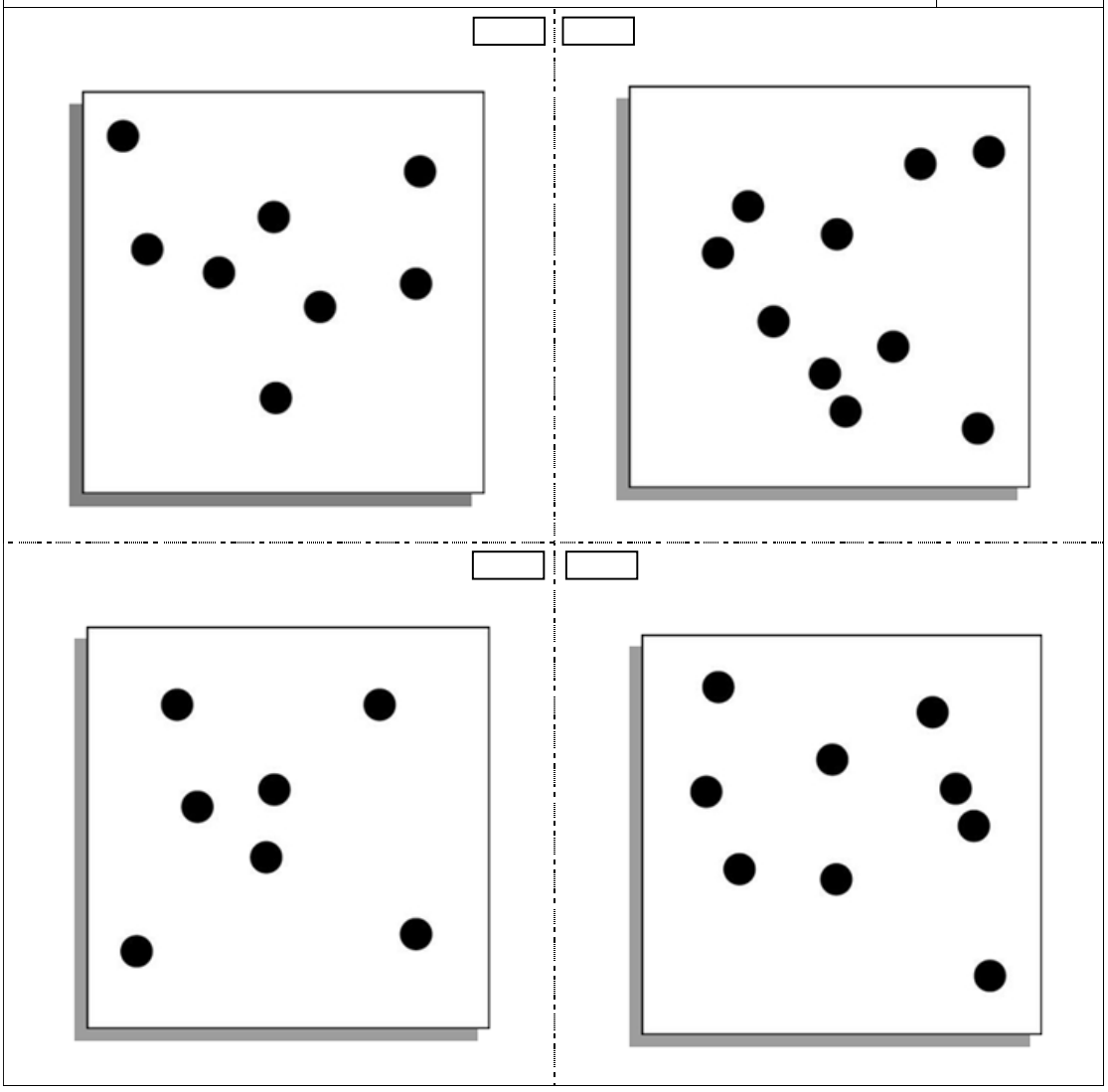
指出哪個與航海有關？ _____

© Jo0n R Ho-g.6, tran62at.- ,y Ruan-&0ing Yu & &0.ng-&0ang Yang 2020.6

語言	
➢ 讓受試者念出以下單詞：[所有的都正確給 1 分] 用 心 為 止 公 開 龍 舟 品 質	語言 [得分0-1] ☐
視覺空間	
➢ 相互重疊的無窮大數學符號：讓受試者遵循以下的圖形畫出相同的圖形	視覺空間 [得分0-1] ☐
	
➢ 立方體：讓受試者畫出相同的立方體圖案	視覺空間 [得分0-2] ☐
	
➢ 時鐘：讓受試者畫出帶有數字的鐘面，指針指向 5 點 10 分	視覺空間 [得分0-5] ☐

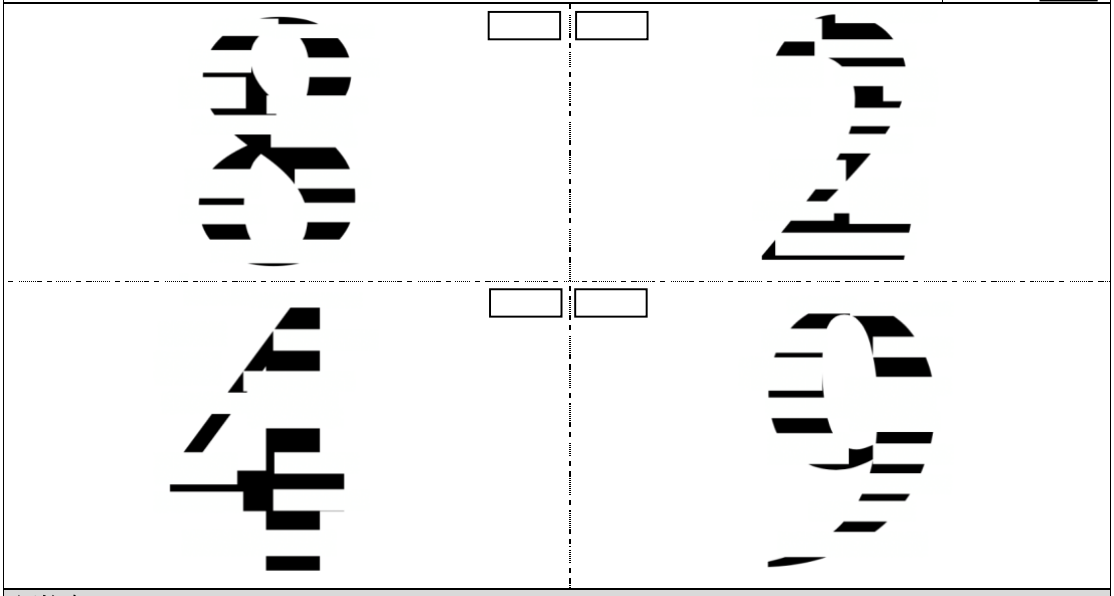
© Jo0n R Ho-g.6, tran62at.- ,y Ruan-&0ing Yu & &0.ng-&0ang Yang 2020.6

視覺空間能力	
➢ 讓受試者說出黑點的個數，但是不能用手指去數	視覺空間能力 [得分0-4] ☐



© Jo0n R Ho-g.6, tran62at.- ,y Ruan-&0ing Yu & &0.ng-&0ang Yang 2020.6

視覺空間能力	
➢ 讓受試者說出下面的數字	視覺空間能力 [得分0-4] ☐



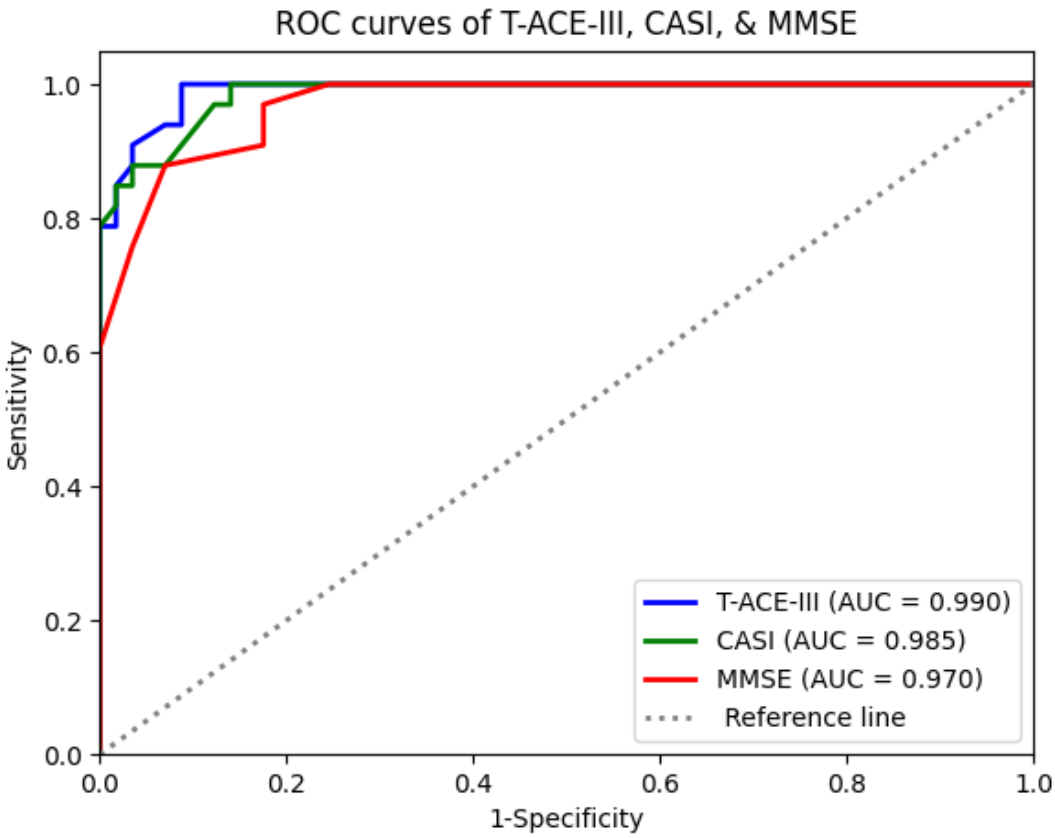
記憶力		
➢ 問受試者：“告訴我剛開始的時候我們重複練習過的名字和地址。”		
王志偉 高雄市 楠梓區 創新路 16 號	___ ___ ___ ___ ___ ___ ___	記憶力 [得分0-7] ☐
➢ 如果受試者有一個或以上項目想不起來，則進行下面的測試。如果都能回憶起來的話可以跳過下面這個測試，並且直接計 5 分。讓受試者選擇：“下面我給你一些提示，名字（或者其他需要選擇的項目）是“X”,“Y”還是“Z”？”如果受試者能找出正確答案可以得 1 分，並可以把這個分數加到回憶一欄的得分裡。		記憶力 [得分0-5] ☐

王志明		王志偉		張文雄		回憶	
高雄市		桃園市		新北市		回憶	
鹽埕區		左營區		楠梓區		回憶	
建國路		創新路		延平南路		回憶	
61 號		17 號		16 號		回憶	

ACE-III 總分	/100
注意力	/18
記憶力	/26
語言流暢性	/14
語言	/26
視覺空間	/16

© Jo0n R Ho-g.6, tran62at.- ,y Ruan-&0ing Yu & &0.ng-&0ang Yang 2020.6

Appendix 29: ROC curves of T-ACE-III, CASI, and MMSE



Appendix 30: T-ACE-III study- Brixton test questionnaire (UK & Taiwanese versions)

The Brixton Spatial Anticipation Test

- *'There are many pages here which all have the same basic design on them. There are always ten positions, and one of them is always coloured blue'* [point to filled circle on page one]. *'However the coloured one moves around according to various patterns that come and go without warning. These numbers* [point to numbers underneath the circles] *are just here to refer to the position – there is nothing complicated or mathematical about this test'.*
- *'Now, as I turn the pages over, your job is to pick up on the pattern as best you can, and point to where you think the blue one is going to be on the next page. It's not guess-work – you can work it out. For instance, imagine the blue one was here* [point to position 6], *and then when I turn the page it goes to 7, and then to 8, then to 9 – you might reasonably expect it next to go to 10'.*
- *'From time to time the pattern changes without warning, and then it is your job to pick up on the new pattern as best you can. Do you understand?'*
- Give further assistance if necessary
- *'Obviously the first time you have nothing to go on, so your first answer will have to be a guess – have a guess as to where the blue one will be next'*

Item/ page	Correct answer	Subject's response	Correct/ incorrect
1	any		
2	3		☐
3	4		☐
4	5		☐
5	6		☐
6 *	7		☐
7	4		☐
8	3		☐
9	2		☐
10	1		☐
11	10		☐
12 *	9		☐
13	10		☐
14	5		☐
15	10		☐
16	5		☐
17	10		☐
18	5		☐
19 *	10		☐
20	7		☐
21	8		☐
22	9		☐
23	10		☐
24	1		☐
25	2		☐
26 *	3		☐
27	10		☐
28	9		☐
29 *	8		☐
30	1		☐
31	2		☐
32	3		☐
33	4		☐
34 *	5		☐
35	4		☐
36	10		☐
37	4		☐
38	10		☐
39	4		☐
40	10		☐
41 *	4		☐
42	9		☐
43	9		☐
44	9		☐
45	9		☐
46	9		☐
47	9		☐
48 *	9		☐
49	9		☐
50	8		☐
51	9		☐
52	8		☐
53	9		☐
54	8		☐
55	9		☐

Total number of errors (raw score)

* Scaled score

Table D

Raw score	Scaled score	Classification
0–7	10	Very superior
8	9	Superior
9–10	8	Good
11–13	7	High average
14–17	6	Average
18–20	5	Moderate ave.
21–23	4	Low average
24–25	3	Poor
26–31	2	Abnormal
> 31	1	Impaired

3

The Brixton Spatial Anticipation Test 布里克斯頓空間預測測驗

- 接下來要進行的測驗有很多頁，每一頁都有相同的設計：兩排圓圈，一排有五個，共有十個圓圈。其中只會有一個是藍色的圓圈[指向第一頁上的藍色圓圈]。 這個藍色圓圈會在沒有提示的狀況下，根據不同模式來回移動。 這些數字[指向圓圈下方的數字]只是標示圓圈的參考位置，這個測驗沒有複雜的數學原理。
- 現在，當我翻頁時，你要進行的工作是預測圓圈，指向你認為下一頁上的藍色圓圈會出現的位置。 這不是用猜的，你可以盡量自己去思考並完成它。 例如，假設藍色圓圈的位置在這邊[指向位置6]，然後當我翻頁時，它的位置依序移動到 7、8、9，你可能會預測下一個是位置是 10。
- 在測驗中，藍色圓圈的模式會在沒有提示下改變，您需要盡可能地去預測到新的模式。 這樣你有了解嗎？
- 必要時給予進一步協助。
- 當然，測驗的第一題你不能根據之前的規則去預測，所以你的第一個答案必須用猜的- 猜下一個藍色的圓圈會在哪一個位置。

題目/頁面	正確答案	受試者反應	正確/不正確
1	任何		
2	3		
3	4		
4	5		
5	6		
6*	7		
7	4		
8	3		
9	2		
10	1		
11	10		
12*	9		
13	10		
14	5		
15	10		
16	5		
17	10		
18	5		
19*	10		
20	7		
21	8		
22	9		
23	10		
24	1		
25	2		
26*	3		
27	10		
28	9		

題目/頁面	正確答案	受試者反應	正確/不正確
29*	8		
30	1		
31	2		
32	3		
33	4		
34*	5		
35	4		
36	10		
37	4		
38	10		
39	4		
40	10		
41*	4		
42	9		
43	9		
44	9		
45	9		
46	9		
47	9		
48*	9		
49	9		
50	8		
51	9		
52	8		
53	9		
54	8		
55	9		

總錯誤次數（原始分數）

標準化分數

表D

原始分數	標準化分數	分類
0-7	10	極優秀
8	9	優秀
9-10	8	良好
11-13	7	尚可
14-17	6	中等
18-20	5	中低
21-23	4	低
24-25	3	差
26-31	2	異常
>31	1	受損

Appendix 31: T-ACE-III study- The-coin-in-the-hand test (UK & Taiwanese versions)

Version: 3.0 Date: 2020/04/26

COIN IN THE HAND TEST

I would like you first to close your eyes and count back from 10. I am now going to test your memory for which hand I have hidden a coin. I will put a coin in one of my two hands [SHOW COIN IN ONE HAND…USE LARGE COIN], then I will ask you to close your eyes and count back to 10 [DO SAMPLE COUNT BACK 10-1], and then I will ask you to open your eyes and tell me which hand the coin was in. Do you understand?

(Examiner's)

1. **LEFT**
2. **RIGHT**
3. **RIGHT**
4. **LEFT**
5. **RIGHT**
6. **LEFT**
7. **LEFT**
8. **RIGHT**
9. **LEFT**
10. **RIGHT**

Score …. / 10

手中的硬幣中文版測驗

受試者編號　　施測人員　　受測日期　　年　　月　　日

請您先閉上眼睛，並從　　開始倒數到　。
我現在要測試您的記憶，讓您看硬幣在我哪一隻手中。 首先，我會將這個硬幣放在我其中一隻手中 展示一隻手中的新台幣　　元硬幣…… 使用大一點的硬幣　，然後，請您閉上眼睛，然後倒數　　從　　數到　，然後我會請你睜開眼睛，再告訴我硬幣在哪隻手上。

你了解了嗎？

（施測人員）

1. 左
2. 右
3. 右
4. 左
5. 右
6. 左
7. 左
8. 右
9. 左
10. 右

得分　　/ 10