TREATMENT ADHERENCE AND ADVERSE EFFECTS IN MOOD DISORDERS

THE EFFECTS OF ONE-ON-ONE MEDICATION TRAINING ON MEDICATION ADHERENCE IN PATIENTS WITH MOOD DISORDERS AND THE EFFECT OF ELECTROCONVULSIVE THERAPY ON COGNITIVE FUNCTIONING IN PATIENTS WITH DEPRESSION

By CAROLINA OREMUS, M.D.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Doctor of Philosophy

DOCTOR OF PHILOSOPHY (2016)

(Neuroscience)

McMaster University

Hamilton, Ontario

TITLE: The effects of one-on-one medication training on medication adherence in patients with mood disorders and the effect of electroconvulsive therapy on cognitive functioning in patients with depression

AUTHOR: Carolina Oremus, M.D. (Universidad Popular Autónoma del Estado de Puebla [UPAEP])

SUPERVISOR: Margaret McKinnon, Ph.D., C. Psych.

PAGES: xxii - 236

Lay Abstract

Mood disorders (MD), including major depressive disorder (MDD) and bipolar disorder, are among the most common mental disorders worldwide. Treating MD is a challenge because of long treatments, the presence of other illnesses, treatment side effects, problems with memory, attention, and decision-making, a lack of understanding about medications, or incorrect beliefs about medication (BAM). Persons with MD who do not respond to drug treatment are often given electroconvulsive therapy (ECT).

This thesis explored the challenges of treating persons with MD through: (1) a pilot study examining whether a one-on-one personalized medication training program, called PIMM/SAM, would help persons with MD take their medications as prescribed; and (2) a study of the effects of ECT on cognitive functioning in depression. Results: (1) participants randomized to PIMM/SAM group held fewer negative BAM than participants receiving standard care; (2) evidence showed worse cognitive functioning in persons who received more intense forms of ECT.

Word count: 150 (150 words maximum)

Abstract

Mood disorders (MD) are among the most common mental disorders worldwide. Low treatment adherence and treatment resistance are two of the most substantial challenges facing clinicians who treat persons with MD. This thesis examined: (1) a pilot study investigating whether a one-on-one personalized medication training program, called PIMM/SAM, improves medication adherence in persons with MD; and (2) a systematic review and meta-analysis on the effects of electroconvulsive therapy (ECT) on cognitive functioning in persons with depression.

To evaluate the impact of PIMM/SAM on medication adherence, a randomized controlled trial was launched in a mood disorders inpatient unit to compare PIMM/SAM (partnership in medication management/self-administered medication) program to standard prescribing practice (SPP). Over follow-up in the feasibility portion of the trial, participants in the PIMM/SAM group (n = 7) held fewer negative beliefs about medications and had lower depersonalization scores compared to participants in the SPP group (n = 5). Between-group differences on the Medication Adherence Rating Scale favoured the PIMM/SAM group, but were not statistically significant.

To examine the effects of bilateral versus unilateral ECT on cognitive performance in persons with TRD, 18 studies across 10 different cognitive domains were meta-analyzed. In the 8- to 30-day timeframe post-ECT, persons who received bilateral versus unilateral ECT had over double the odds of worse cognitive performance in global cognition, non-

verbal memory delayed recall, verbal memory immediate and delayed recall, subjective memory, and verbal memory immediate recall.

A personalized medication training program in a mood disorders clinic may have positive implications for medication adherence. The trial to evaluate PIMM/SAM versus SPP is ongoing and further evidence about the training program is expected within the next 12 months. The systematic review and meta-analysis showed that cognitive performance was worse in persons who received bilateral versus unilateral ECT in some cognitive domains at 8 to 30 days post-treatment.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Margaret McKinnon, for taking me as a graduate student and for her guidance, support and direction throughout my doctoral studies. I would also like to sincerely thank Dr. Ruth Lanius, my co-supervisor, for her outstanding advice, guidance and support over the last few years. I am also especially grateful to my committee members, Drs. Roberto Sassi and Geoffrey Hall, for providing me with feedback and encouragement.

Thank you to study participants, researchers, and staff at St. Joseph's Healthcare Hamilton and the Mood Disorders Program, especially Sharon Simons, Laura Garrick, and Marija Gojmerac, whose support was paramount for undertaking the work in this thesis. Thanks to Sandra Murphy for all the answers, help, and guidance that you provided me during my grad studies. I would like to extend my sincerest gratitude to the many students and volunteers who helped with my research projects throughout the years, especially to Drs. Heather McNeely and Bruno Losier, the ECT & Cognition Systematic Review Team, Megan McFarlane, Julia Mason, Latisha Rhooms, Joseph Capozza, Diego Lamilla, and Luz Adriana Perdomo.

I want to specially thank Melissa Parlar, Alina Protopopescu, and Yasir Rehman for the endless hours we spent working on and discussing the ECT meta-analysis. I would like to extend a very special appreciation to the former and current graduate students in Dr. McKinnon's lab for all your help, support, and friendship. I will always remember the many anecdotes we shared during our graduate studies.

I would like to sincerely thank Drs. Tunks, Muniz and Trinh, Jareck Krawczyk, Brandy John, Enoch Ho, Bethany Stuart, and Miranda McKinley for going the extra mile to help me to physically function since the car accident.

Many thanks also to my friends and relatives, especially to my cousin Cristy, friends from the Hoffman's cell group, and to Pastors Joanna Adams, Nancy Neuman, Danielle Di Benedito, Karen and Tim Roberts, and Nancy Campbell, who stood by me throughout my PhD. Your love, encouragement, and fellowship kept me focused on the finish line and helped me jump the hurdles along the way.

My sincerest gratitude to my parents, Arturo and Charis, for their love, support, and continuous encouragement for me to never give up and pursue all of my dreams. I specially want to thank my mother for never making education optional. To my father for showing me how to dream and rise above adversity. To my late grandparents, who are always in my heart, for their constant cheerleading, unconditional love, support, and for teaching me to value education, family, and life, caring and respecting others and to fight for what is right. To my in laws, for showing me love and encouragement since the day we met. Last but certainly not least, to my husband Mark, for his love, support, encouragement, for keeping me on track, and for bringing Paws into our home. Thanks Mark for all your constant cheerleading, care, smiles, and laughter. Thank G-d for putting all of these wonderful people and opportunities in my life.

Table of Contents

List of Tables

List of figures

CHAPTER 2

CHAPTER 4

List of abbreviations

ACC	Anterior Cingulate Cortex
AHRQ	Agency for Healthcare Research and Quality
AMI	Autobiographical Memory Interview
ANOVA	Analysis of Variance
BAI	Beck Anxiety Inventory
BD	Bipolar Disorder
BDI	Beck Depression Inventory
BDNF	Brain-derived Neurotrophic Factor
BL	Bilateral
BMC	BioMed Central
BMQ	Beliefs about Medication Questionnaire
BNT	Boston Naming Test
BSRT	Buschke Selective Reminding Test
BVRT	Benton Visual Retention Test
CADTH	Canadian Agency for Drugs and Technologies in Health
CIHR	Canadian Institutes of Health Research
CBT	Cognitive Behavioural Therapy
CDR	Commons Drugs Review
CFQ	Cognitive Failures Questionnaire
CFT-Delay	Cognitive Function Test-Delay

CGI-SS	Clinical Global Impression - Severity Scale
CI	Confidence Intervals
CONSORT	Consolidated Standards of Reporting Trials
COWAT	Controlled Oral Word Association Test
CPT	Continuous Performance Task
CSF	Cerebrospinal Fluid
DBS	Deep Brain Simulation
DENG	Disengagement
DEPR	Depersonalization
DERL	Derealization
D-KEFS	Delis-Kaplan Executive Function System Sorting Test
DPDD	Dysthymia/Persistent Depression Disorder
DS	Dissociative Subtypes
DSM-5	Diagnostic and Statistical Manual of Mental Disorders - 5
EBM	Evidence-based Medicine
ECON	Emotional Constriction
ECT	Electroconvulsive Therapy
GABA	Gamma-aminobutyric Acid
G-BMQ	General-BMQ
GDE-my	Global Self-Evaluation-Memory
GR	Glucocorticoid Receptor

GRADE	Grading of Recommendations Assessment, Development and Evaluation
GSE	General Self-Efficacy
GSE-my	Global Self Evaluation-Memory
HAQ	Helping Alliance Questionnaire
HRQoL	Health-Related Quality-of-life
ICC	Intraclass Correlation Coefficient
IDDIS	Identity Dissociation
IFN	Interferon
IL-6	Interleukin-6
IL-8	Interleukin-8
IPT	Interpersonal Psychotherapy
iSPOT-D	International Study to Predict Optimized Treatment in Depression
κ	Kappa
MARS	Medication Adherence Rating Scale
MCSD	Mean Change Standard Deviation
MD	Mood Disorder
MDD	Major Depressive Disorders
MDE	Major Depressive Episode
MDI	Multi-scale Dissociation Inventory
MEMD	Memory Disturbance
MINDS	McMaster Integrative Neuroscience Discovery & Study

MINI	Mini International Neuropsychiatric Interview
MMSE	Mini Mental State Examination
MOCA	Montreal Cognitive Assessment
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIMH	National Institute of Mental Health
NOS	Newcastle-Ottawa Scale
OR	Odds Ratio
PAL	Paired Associates Learning Test
PIMM/SAM	Partnership in Medication Management/Self-administered Medication
PIMT	Personal and Impersonal Memory Test
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta Analyses
PROSPERO	International Prospective Registry of Systematic Reviews
PTSD	Post-Traumatic Stress Disorder
RAVLT	Rey Auditory Verbal Learning Test
RBANS	Repeatable Battery for the Assessment of Neuropsychological Functioning
RVMT	Rivermead Behavior Memory Test
RCT	Randomized Controlled Trial
RFT	Rey Figure Test
rTMS	Repetitive Transcranial Magnetic Stimulation
sBDNF	Serum Brain-derived Neurotrophic Factor

SC-BMQ	Specific Concern-BMQ
SCG	Subgenual Cingulate Gyrus
SD	Standard Deviation
SSEMF	Subjective Self-Evaluation of Memory Function
SF-36	Short Form-36 Health Survey
SF-36-GH	Short Form-36 Health Survey - General Health
SF-36	Short Form-36 Health Survey – Mental Health
SMD	Standard Mean Difference
SMDs	Standard Mean Differences
SMCQ	Squire Memory Complaint Questionnaire
SPP	Standard Prescribing Practice
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TMS	Transcranial Magnetic Simulation
TRD	Treatment Resistant Depression
UL	Unilateral
VNS	Vagus Nerve Simulation
WAIS	Weschler Adult Intelligence Scale
WMS	Weschler memory Scale
WHO	World Health Organization

Declaration of Academic Achievement

This thesis contains a total of five chapters: Chapters 2 through 4 are empirical articles, Chapter 1 provides background information to the material in Chapters 2 through 4, and Chapter 5 discuss the main conclusions and future directions.

Chapter 2

Carolina Oremus, Sharon Simons, and Margaret McKinnon conceived the study presented in Chapter 2. Carolina Oremus, Mark Oremus, and Margaret McKinnon designed the study and study protocol. Carolina Oremus developed the case report forms, training and advertising materials, trained hospital staff, collected and analyzed the data. Mark Oremus, Sharon Simons, and Margaret McKinnon reviewed and revised the protocol and study forms. Carolina Oremus, Ashok Chaurasia, Mark Oremus, and Margaret McKinnon developed the statistical analysis plan and analyzed the data. Carolina Oremus and Mark Oremus drafted the manuscript. Ashok Chaurasia, Sharon Simons, Jennifer Lowe, and Margaret McKinnon reviewed and revised the manuscript for important intellectual content. All authors read and approved the final manuscript. The manuscript was submitted *BMC Pilot and Feasibility Studies* in July 2016.

Chapter 3

This chapter contains two published manuscripts. One manuscript was published in *BMJ Open* in March 2015 ("Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis"). Carolina Oremus, with help from Mark Oremus, Heather McNeely, Bruno Loiser, Matthew King,

Gary Hasey, Ruth Lanius, and Margaret McKinnon conceived and designed the study. Carolina Oremus conducted the literature search and led the article screening process. Melissa Parlar, Yasir Rehman, Jorge Ortiz, Megan McFarlane, Latisha Rhooms, and Julia Mason helped with the article screening process. Melissa Parlar, Yasir Rehman, Joseph Capozza, Alina Protopopescu, Rachel Wells, Gagan Ferhava, Allyson Graham, Caitlin Gregory, Lindsay Hanford, Anthony Nazarov, Maria Restivo, Erica Tatham, Wanda Truong, and Carson Vanderwel helped Carolina Oremus extract data from the included studies and assess the risk of bias of the included studies. Carolina Oremus drafted the manuscript and the other authors critically revised the manuscript for important intellectual content.

The second manuscript in Chapter 2 was published in *BMJ Open* in July 2012 ("Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales"). Carolina Oremus and Mark Oremus conceived and designed the study. Mark Oremus analysed the data. Carolina Oremus, Mark Oremus, Margaret McKinnon, Geoffrey Hall, and the ECT & Cognition Systematic Review Team interpreted the data. Mark Oremus drafted the manuscript. Carolina Oremus, Margaret McKinnon, Geoffrey Hall, and the ECT & Cognition Systematic Review Team critically revised the manuscript for important intellectual content. All authors approved the final version of the manuscript.

Chapter 4

Chapter 4 contains the results and discussion of the systematic review and meta-analysis described in the protocol that was published in *BMJ Open* in March 2015. Carolina Oremus interpreted the extracted data and the assessments of risk of bias. Alina Protopopescu and Carolina Oremus checked, extracted and entered data into the meta-analysis software. Yasir Rehman, Joseph Capozza helped Carolina Oremus to grade the strength of evidence. Carolina Oremus ran the meta-analyses, and wrote the results and discussion. Drs. David Streiner and Mark Oremus provided Carolina Oremus with statistical and methodological advice. All of these tasks were performed under the supervision of Drs. McKinnon, Lanius, Sassi, and Hall. The results and discussion described in this chapter, plus a summary of the introduction and methods already described in the *BMJ Open* paper from March 2015, will be submitted to *The Lancet* in 2016.

CHAPTER 1:

GENERAL INTRODUCTION

1. General Introduction

This thesis examines two components of treating mood disorders (MD): (1) a one-on-one medication training program to improve medication adherence in persons with MD; and (2) the effects of electroconvulsive therapy (ECT) on cognitive functioning in persons with depression (MDD). The introduction presents background information to contextualize the remainder of the thesis. The introduction begins with an explanation of mood disorders and medication adherence to provide a segue into Chapter 2, which reports on the early results of a randomized controlled trial (RCT) that is part of a larger sequential explanatory mixed methods study to evaluate a personalized medication training versus standard prescribing practice as a means of improving medication adherence in persons with MD. Next, the introduction presents information on treatment-resistant depression (TRD), electroconvulsive therapy (ECT), cognition, and evidence-based medicine (EBM) to place Chapters 3 and 4 in the context of a meta-analysis to investigate the effects of ECT on cognitive functioning. Chapter 5, which is not discussed in the introduction, is a summary of the content and main findings of the thesis. Two important conclusions arising from the thesis, with respect to treating MD, are that personalized medication training programs can improve medication adherence and that ECT is associated with cognitive deficits in persons treated for TRD.

1.1. Mood Disorders

MD, including MDD, dysthymic/persistent depression disorder (DPDD), and bipolar disorder (BD), are among the most common mental disorders in Canada. Statistics

Canada reported an MD prevalence of 6.5% (N = 2,346,244) among all persons in Canada (Government of Canada, 2016). Globally, the World Mental Health Surveys estimate a 12% lifetime prevalence of MD and a 6% 12-month prevalence of MD; MD are the second most prevalent type of mental health disorder, following anxiety disorders (Kessler et al., 2009). Approximately 350 million people worldwide suffer from MDD alone (World Health Organization, 2008). The estimated prevalence of DPDD is 5% (Sadock & Sadock, 2007) and the estimated prevalence of BD ranges from 0.1% in Nigeria to 3.3% in the United States (Merikangas et al., 2011). Persons with MD experience changes in many domains, including behaviour, cognitive/body functions, thoughts, studies/education, family and social interactions (American Psychiatric Association, American Psychiatric Association, & DSM-5 Task Force, 2013). Financially, the burden of MD leads to productivity losses in the workplace, high healthcare costs, and lost familial income. MD adversely affects family income because persons with a mood disorder, and their close family relatives, often require unpaid time away from work. Additionally, the cost of pharmaceutical therapy can be quite high, especially for individuals with limited or without prescription drug insurance (Health Canada, 2002).

MD are typically diagnosed by examining clinical symptomatology as seen in the Diagnostic and Statistical Manual of Mental Disorders – 5 (DSM-5) (American Psychiatric Association et al., 2013). Additional work has generated evidence that breakdowns in common brain circuits related to cognition and behaviour may be

responsible for the development of the psychopathology and general dysfunction inherent in MD (Iorfino, Hickie, Lee, Lagopoulos, & Hermens, 2016). Therefore, neuroimaging, neurophysiology, and circadian biology, as well as studying related issues such as, may identify biomarkers that indicate a risk of MD, or that can serve as therapeutic targets for treatment. Evolving thinking has posited that neurobehavioural systems, rather than single neurotransmitters, are part of the etiology of MD (Sadock & Sadock, 2007). Components of these overarching neurobehavioural systems include norepinephrine, serotonin, and dopamine (Sadock & Sadock, 2007).

A recent systematic review of functional neurobiological parameters in young people (i.e., 12 to 30 years) with MD or anxiety disorders highlights the interest in researching the neurobiology of MD (Iorfino et al., 2016). The review's authors included 134 studies that were categorized into five functional domains, with many studies featuring more than one domain: social and economic participation: n = 11; physical health: n = 3; suicide and self-harm behaviours: n = 22; alcohol and substance use: n = 10; and clinical syndrome: n = 98. The neurobiological parameters examined in these studies were neuropsychology n = 28, neuroimaging n = 62, sleep-wake and circadian biology n = 23, neurophysiology n = 21, and metabolic measures n = 10.

The authors of the review found a relation between slower reaction times in induced-based decision-making tasks, as well as an increased frequency of self-harm in persons with MDD in remission. In terms of symptomatology, the authors reported a relation between increased startle response and numerous depressive episodes and anxiety

symptoms. A reduction in the startle response was linked to a reduction in anxiety levels following a course of cognitive behavioural therapy (CBT). These findings, combined with the evidence from the neuroimaging studies, confirmed the importance of the amygdala activation in the threat process (Iorfino et al., 2016).

The authors of the review identified a gap in the literature in the area of biomarkers and functional domains (e.g., physical health, social and economic participation, suicide and self-harm, and substance use). Most of the manuscripts in this systematic review describe research related to clinical symptoms. Understanding the relation between biomarkers and functional domains may improve treatment and prognosis in young people with MD (Iorfino et al., 2016).

The search for a biological basis for MD has focused on the monoamine hypothesis. Due to associations between the psychological and cellular actions of psychotropic agents, a view exists that functional deficiencies in catecholamines, especially norepinephrine, can lead to depression. Research has connected other biogenic amines to depression, including serotonin, dopamine, and epinephrine (Barchas & Altemus, 1999).

Another biological basis for MD involves inflammation. Persons with MDD show heightened expression of pro-inflammatory cytokines and receptors, plus high levels of acute-phase reactants, chemokines, and soluble adhesion molecules in peripheral blood and cerebrospinal fluid (CSF). Additionally, pro-inflammatory 'M1' macrophage phenotypes and magnified IL-6, IL-8, and type I IFN-induced signalling pathways are hallmarks of MDD (Miller & Raison, 2016). Early data postulate that depressed moods

may be ameliorated by inhibiting pro-inflammatory cytokines and related signaling pathways (Miller, Maletic, & Raison, 2009).

In recent years, researchers have focused on the relation between MD and trauma. The plight of many returning military personnel from service in Vietnam during the 1970s spurred studies of the etiology, risk factors, and effects of post-traumatic stress disorder (PTSD). Researchers noticed that roughly half of all people with PTSD had MDD (Flory & Yehuda, 2015). Research has also found that a majority of persons with MDD or an anxiety disorder have co-morbid trauma. For example, approximately 91% of persons in a study of 2,000 people with anxiety or depressive disorders reported experiencing a traumatic or bothersome life event (Spinhoven, Penninx, van Hemert, de Rooij, & Elzinga, 2014). Furthermore, the National Institute of Mental Health (NIMH) has reported that 40% of persons with PTSD develop depression within one to four months after experiencing a traumatic event (National Institute of Mental Health, 2016).

Explanations for the PTSD/MDD comorbidity lie in symptom overlap and the possibility of a distinct trauma-related phenotype. Dissociation and altered states of consciousness in time, thought, body, and emotion are thought to have transdiagnostic implications for trauma-related disorders; models based on dissociative states may help identify risk factors for trauma-related disorders (Lanius, 2015). Also, further examination of the glucocorticoid receptor (GR) and FKBP5 genes and related molecular processes can help elucidate the relation between PTSD and MDD (Flory & Yehuda, 2015).

There is an urgent need to identify factors that contribute to treatment efficacy in MD. Following first-line treatment, only 37% percentage of patients will show a positive treatment response, with dwindling positive responses at subsequent phases (i.e., 31% second phase, 14% third phase, and 13% fourth phase) (Rush et al., 2006).Treatments for mood disorders are also associated with significant adverse effects. Accordingly, we sought to examine a means of improving medication adherence in a mood disorders inpatient unit. Where possible, we also explored the relation between MDD and trauma in participants who were part of the study.

1.2 Medication Adherence in Mood Disorders

Adherence is defined as the extent to which an individual behaves in accordance with medical advice (Dunbar-Jacob & Mortimer-Stephens, 2001). The WHO has reported that 50% of persons who are diagnosed with chronic conditions in developed countries do not use their medications as prescribed (World Health Organization, 2008). In MD, the two main adherence problems are medication non-persistence and medication non-compliance (Chong, Aslani, & Chen, 2011). Medication non-persistence is the premature discontinuation of pharmacotherapy. Medication non-compliance is the lack of regularity in taking or using prescribed drugs. Horne has reported that most people's personal understanding and beliefs regarding their MD diagnosis will influence how they evaluate the logic of prescribed medical treatments or medical recommendations (Horne, Weinman, Barber, Elliott, & Morgan, 2006).

In terms of percentages, medication adherence in MD is low, ranging from 30% to 70% in MDD and 18% to 52% in BD (Scott & Pope, 2002). Psychiatric medication non-adherence rates have been constant since the 1950s. The WHO has earmarked MDD as one of the nine chronic conditions that can benefit most from efforts to improve adherence. Nearly a third of persons with MDD discontinue their antidepressant treatment against medical advice during their first month of treatment (Chong et al., 2011). In addition, persons with MDD may not agree with the timing, dose and frequency of medication taking precisely because of their normative beliefs about the disease. For BD, research has shown that the average person lies midway along the continuum between being completely adherent and completely non-adherent. Furthermore, research has shown that 30% to 50% of persons prescribed with a mood stabilizer for prophylaxis have on average one episode of non-adherence per year.

Scott and Pope reported that clinicians were uncertain of the true reasons for non-adherence; clinicians often wrongly believed that persons' self-perceived good health led to non-adherence (Scott & Pope, 2002). Also, previous research regarding medication adherence in psychiatric populations has concluded that patients' attitudes and behaviours encouraged non-adherence. Some other research has blamed medications' side effects as the cause for low medication adherence; however, side effects are actually ranked as the seventh reason for discontinuing medications (Clarworthy, 2009).

Reasons for non-adherence in MD may include adverse effects from medications or the fear of suffering such effects, suicide, hospital re-admission, recurrence of disease

episodes, the effects of disease symptoms (e.g., low motivation, fatigue or loss of energy, lack of concentration, indecisiveness), stigma toward psychiatric medications or belief in propaganda against psychiatric medications, low quality-of-life, negative attitudes toward therapy, job loss due to illness, familial and relationship troubles, lower likelihood of having family members involved in treatment, forgetfulness, belief in being cured and not needing medications, perceived ineffectiveness of therapy, misunderstanding of treatment regimens, self-dosing, or impairment due to alcohol, physical illness, changing patterns of healthcare delivery, increased medication costs, or mania (Coldham, Addington, & Addington, 2002; Dunbar-Jacob & Mortimer-Stephens, 2001). The presence of cognitive deficits, particularly deficits in executive functioning and in verbal memory, has also been associated with poor functional outcomes (e.g., vocational) in patients with mood disorders (Depp et al., 2009; Dickerson et al., 2004; Gildengers et al., 2007; Jaeger & Vieta, 2007). Critically, these cognitive deficits may persist following the resolution of depressive episodes in some patients, and show worsening with subsequent episodes of illness (MacQueen et al., 2003). Moreover, cognitive deficits impact negatively on the outcome of pharmacological and non-pharmacological treatments for mood and related affective disorders, where the ability to engage in and successfully complete treatment relies heavily on higher-order cognitive processes (Dunkin et al., 2000; Polak, Witteveen, Reitsma, & Olff, 2012). From the perspective of adherence, a study of 100 persons with BD found that persons who adhered to treatment had fewer comorbidities, more resources

for coping with stress, and higher contentment with life compared to persons who did not adhere to treatment (Darling, Olmstead, Lund, & Fairclough, 2008).

Numerous attempts to improve medication adherence in MDD have been reported in the literature. A recent systematic review of 26 studies of interventions to improve antidepressant medication adherence found that the most successful interventions were multifactorial in nature (Chong et al., 2011). These successful interventions employed strategies involving mental health professionals, education, telephone monitoring of participants' progress, ongoing support of participants, participants' inclusion in the process of taking medications, and feedback of participant progress to partners in care. Education alone, without monitoring and feedback, was unsuccessful in boosting adherence. The authors of the review concluded that better adherence to antidepressant drug therapy requires behavioural modification through structured programs that reach beyond didactic education sessions or the provision of reading materials. However, the authors could not identify the specific components of multifactorial programs that were the most responsible for improving adherence. No one combination of components appeared to be optimal for promoting medication adherence (Chong et al., 2011).

Similar programs exist in other areas of medicine. A recent systematic review of patient support programs included 64 studies across the spectrum of chronic diseases (Ganguli, Clewell, & Shillington, 2016). Each study described a different program, but the majority of programs shared similar features such as clinic-based face-to-face support mechanisms led by allied healthcare professionals (nurses, pharmacists). Other features

included group teaching, refill reminders, and mailings. Of the 41 programs that measured adherence, 27 (66%) reported positive outcomes on measures such as pill counts, ad hoc and standardized adherence questionnaires, prescription refills, and self-report.

1.3. Partnership in Medication Management/Self-Administered Medication

In this thesis, we evaluated the efficacy of the Partnership in Medication Management/Self-Administered Medication (PIMM/SAM) program versus standard care to improve medication adherence in MD (Chapter 2). PIMM/SAM involved individual, one-on-one sessions between persons with MD and nurses. In PIMM/SAM, study participants were responsible for taking their medications as prescribed, with education, help, and support from nurses in interactive sessions. In the sessions, persons with MD told their nurses about the medications they would administer at home. The nurses gave these persons educational information about their medications and also discussed strategies to help them take their medications as prescribed. Participants were also provided with a choice of notebooks, pens, highlighters, post-its, alarm clocks, checklists, and referrals to online apps to record information about medications and trigger reminders to take medications. On each day participants were in the study, they were responsible for notifying the nurses within an hour of the time when they had to take their medications. After the notification, the participants met with their nurses and were shown all of the medications that they were required to take at that time. Participants then identified each of their medications and described the benefits, purposes, dosage, common adverse

effects of each medication, and the importance of continuing to take the medications and following their treatment plans. The main goal of the PIMM/SAM program was to mimic, as much as possible, in the inpatient setting, the conditions under which participants would take their medications at home. Accordingly, participants were asked to select those reminders most consistent with the procedure they would follow at home to remind themselves to take their medication (i.e., use of an alarm clock). The intent was for participants to develop an understanding of the importance of taking their medications as prescribed, as well as a routine to promote continued medication adherence after discharge into the community.

1.4. Treatment-resistant Depression

A recent analysis of the Global Burden of Disease Study found MDD to have the largest age-standardized rate of disability-adjusted life-years among 23 different mental, neurological, and substance-abuse disorders (Whiteford, Ferrari, Degenhardt, Feigin, & Vos, 2016). While existing antidepressant medications may be effective for many persons with MDD, provided they adhere to treatment, some estimates suggest slightly under half of all persons receiving first-line antidepressants will not experience clinical benefits after treatment (Gartlehner et al., 2008). Additionally, only 33% of persons will fully recover or remit following first-line treatment (Trivedi et al., 2006). Adverse effects from antidepressants are many, including headaches, gastrointestinal upset, insomnia, restlessness, fatigue, anxiety, weight gain, sexual dysfunction, and sedation (Santarsieri & Schwartz, 2015). Traditional approaches to overcome treatment nonresponse have

included supplementing the first-line treatment with additional medications or switching to a new medication. The large Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial is one example of an initiative designed to provide evidence to guide the treatment of TRD (Warden, Rush, Trivedi, Fava, & Wisniewski, 2007). Despite such initiatives, limited evidence is available to guide clinical decision making following initial treatment failures (Gaynes et al., 2012; Souery, Papakostas, & Trivedi, 2006).

Besides a lack of evidence regarding adequate drug therapy, TRD is a vexing problem to treat because of disagreement over the diagnostic criteria used to identify the condition. Consequently, many individuals who are thought to be treatment resistant are actually misdiagnosed. Disagreement even exists at the fundamental level of how many failed treatment trials are required before TRD may be considered to be present (Souery et al., 2006).

Non-drug therapies do exist for TRD. Transcranial magnetic stimulation (TMS) is one alternative that has shown modest benefits in typically younger individuals with lower thresholds of treatment resistance and fewer comorbid anxiety or psychotic symptoms. However, the optimal treatment regimen—i.e., duration, sequential bilateral stimulation, pulses per session—is unknown (Lee, Blumberger, Fitzgerald, Daskalakis, & Levinson, 2012). Deep brain stimulation (DBS) to the subgenual cingulate cortex is a newer potential therapeutic option for TRD. A recent meta-analysis reported high 12-month response (39.9%; 95% confidence interval [CI]: 28.4% to 52.8%) and remission (26.3%; 95% CI: 13.0% to 45.9%) rates, and lower 12-month depressive symptoms

(pooled Hedges *g* effect size: -1.89; 95% CI=-2.64 to -1.15), in persons treated with DBS (Berlim, McGirr, Van den Eynde, Fleck, & Giacobbe, 2014). These meta-analytic findings are preliminary and inconclusive of efficacy because they are based on four observational studies totaling 66 participants. Information about the adverse effects of DBS is also limited. A recent crossover RCT comparing DBS of the ventral anterior limb of the internal capsule versus sham stimulation recruited 25 patients into two-phase study (Bergfeld et al., 2016). The first phase was a 52-week optimization phase to achieve stable responses on DBS, followed by the second phase, which was a 12-week double-blinded crossover of DBS versus sham (only 16 people entered the second phase). Adverse effects included severe nausea (n =1), suicide attempts (n = 4), and suicidal ideation (n = 2). More evidence is required to draw firmer conclusions about the efficacy of DBS to treat TRD.

A recent systematic review (also known as an evidence report) from the Agency for Healthcare Research and Quality (AHRQ) investigated the efficacy of four non-pharmacologic therapies in TRD: ECT, repetitive TMS (rTMS), vagus nerve stimulation (VNS), and cognitive behavioral therapy or interpersonal psychotherapy (CBT or IPT) (Gaynes et al., 2011). The report concluded that research investigating the effects of non-pharmacologic interventions on TRD is in its early stages. In terms of efficacy, the results were limited to a small number of RCTs. Only two RCTs compared nonpharmacological therapies: one examined ECT versus rTMS and the second investigated ECT versus ECT plus rTMS. Neither of these RCTs reported any differences in efficacy between the

treatments. In one RCT involving a group of individuals with MDD or BD, ECT showed a 9-point decrease in the Hamilton Depression Scale (Hamilton, 1960) ($p < 0.05$) versus medications alone. The review identified only one trial of VNS, which was compared to sham treatment. No statistically significant differences in efficacy or adverse effects were found between VNS and sham treatment, although the number of withdrawals due to adverse effects was greater (though not statistically significantly greater) in the VNS group. The review did not find any trials of CBT or IPT.

1.5. Electroconvulsive Therapy

ECT is another treatment option for TRD. ECT involves the transmission of electric current through the brain to trigger a brief seizure, change brain chemistry, and reduce symptoms of TRD. The neurotrophic hypothesis of depression has been cited as the mechanism by which ECT works (Tunca et al., 2015). In the 1930s, Lazlo Meduna observed that people who suffered from grand mal seizures did not have schizophrenia. Hence, he began comparing neuroanatomical differences between the brains of people who suffered from epilepsy or schizophrenia. Meduna observed a lack of glial cell growth in the brains of people with schizophrenia; the absence of growth was inversely proportional to the increase in glial cells that he observed in the brains of people with epilepsy. These observations led Meduna to conclude that schizophrenia could be treated with pharmacologically-induced seizures.

Meduna tested his theory by inducing seizures with camphor oil in a catatonic patient who recovered after five treatments. In the late 1930s, two Italian doctors successfully

induced seizures with electricity in a person with schizophrenia, who fully recovered after 11 treatments. In the 1940s, ECT was used in the United States.

Since the initial years of ECT treatment, researchers have studied ways to improve the seizure induction and reduce adverse effects. Today, many different ECT modalities exist, including pulse shape (shift from sine wave to rectangular, pulse width [brief to ultrabrief] and electrode placement (bilateral [BL], unilateral [UL]). Also, ECT is delivered under general anesthesia and patients are provided with muscle relaxants to prevent bone fractures (Dougherty & Rauch, 2007).

Besides the neurotrophic hypothesis, researchers have put forward many additional explanations for ECT's mechanism of action. Neurogenesis is one such explanation. ECT appears to have an impact on neuronal structures, with the shocks from treatment having been observed to increase subgranular zone precursor cell proliferation in the monkey hippocampus. Also, PET studies have shown relations between antidepressants and increased metabolism in the left subgenual anterior cingulate cortex (ACC) and hippocampus (Angela Merkl, Heuser, & Bajbouj, 2009).

Brain-derived neurotrophic factor (BDNF) has been shown to be a component in the pathogenesis of MD. Peripheral BDNF levels are generally lower in MDD and BD. Evidence has shown that BDNF levels improve in tandem with the amelioration of symptoms in persons with MD (Hashimoto, 2010). Therefore, research has sought to explore the possibility of a positive association between ECT and BDNF levels in persons with MD (Bouckaert et al., 2014). However, current results are not promising. For

example, no links have shown between val66val or met versions of the BDNF polymorphism and mood outcomes following ECT (Bennett, Currie, Fernie, Perrin, & Reid, 2016). Also, research has found that hippocampal increases in volume post-ECT were independent of sBDNF and depressive symptomatology (Bouckaert et al., 2016). A recent review concluded that ECT in humans might boost BDNF concentrations, but such increases are not clearly connected to behavioural changes (Polyakova et al., 2015). This conclusion has been echoed by other researchers (Brunoni, Baeken, Machado-Vieira, Gattaz, & Vanderhasselt, 2014).

Further explanations for the mechanism of ECT relate to increased depletion of inhibitory neurotransmission (γ-aminobutyric-acid [GABA]) in the cortical network, which can promote antidepressant and anticonvulsive properties. Also, the electrical discharge from treatment may stimulate monoamine neurotransmitter systems such as dopamine, serotonin, and norepinephrine, or the discharge may promote cell proliferation related to neuroplasticity (the "anticonvulsive hypothesis"). ECT may also restore hypothalamic–pituitary–adrenal axis abnormalities or affect the subgenual cingulate gyrus (SCG), including Brodmann's area 25 and parts of 24 and 32, which demonstrated abnormal metabolic activity in persons with MDD (A. Merkl, Heuser, & Bajbouj, 2009; Angela Merkl et al., 2013).

ECT has evoked diametrically opposed reactions from different quarters of the clinical community, with some clinicians concerned about potential adverse effects, while others believe it is efficacious and safe (The UK ECT Review Group, 2003). In current

clinical practice, ECT prompts a generalized seizure via the delivery of an electrical current to the brain, with leads positioned to the scalp and skull unilaterally or bilaterally. The traditional placements of the leads were bitemporal (or bifrontotemporal or simply bilateral) and right unilateral; more recently, clinicians have been using bifrontal placement. The positioning of the leads affects treatment efficacy and the possibility of experiencing cognitive adverse effects (Kellner, Tobias, & Wiegand, 2010). Common adverse effects from ECT include headache, muscle ache, and nausea. Some ECT recipients experience acute confusion for 30 to 60 minutes post-ECT, largely due to the combined effects of the ECT itself and the anesthesia (Department of Psychiatry, Available at: http://www.psych.med.umich.edu/ect/common-side-effects.asp. Accessed on July19).

Meta-analyses have reported benefits for ECT in the treatment of TRD. A pivotal meta-analysis from the United Kingdom found that real ECT was more effective than simulated ECT (six trials, 256 patients) or drug therapy, with bilateral ECT being more effective than unipolar ECT (22 trials, 1408 participants) (The UK ECT Review Group, 2003). In another meta-analysis, brief right unilateral ECT was statistically significantly more efficacious for depression than ultrabrief right unilateral ECT (standardized mean difference: 0.25; 95% CI: 0.08 to 0.41), although ultrabrief had a lower remission rate than brief (odds ratio: 0.71; 95% CI: 0.51 to 0.99) (Tor et al., 2015).

Meta-analyses have also found that persons who undergo ECT may experience cognitive deficits as adverse effects of ECT. The evidence of cognitive deficits from these

meta-analyses is limited, though, because the number of cognitive domains examined was small (i.e., four domains only [retrograde memory, anterograde memory—learning, anterograde memory—delayed recall, global cognitive function]) (Tor et al., 2015), or studies with substantial clinical heterogeneity were combined under single cognitive domains (Semkovska & McLoughlin, 2010). Alternative explanations suggest that some reports of memory loss following ECT may be manifestations of somatoform disorders (Fink, 2007).

In one of these meta-analyses (Tor et al., 2015), the authors set-out to examine the comparative efficacy of brief pulse versus ultrabrief pulse right unilateral ECT. Cognitive function was a secondary outcome of the review and the authors grouped the neuropsychological instruments used in the included studies into the four domains mentioned in the previous paragraph. The initial literature search yielded 644 references after removal of duplicates and seven studies (5 RCTs, 2 observational) were included in the review. The summary standard mean differences (SMDs) for cognitive function all favoured ultrabrief pulse ECT: retrograde memory (SMD: 0.38; 95% CI: 0.15 to 0.61 [5 studies]); anterograde memory—learning (SMD: 0.45; 95% CI: 0.22 to 0.68 [2 studies]); anterograde memory—delayed recall (SMD: 0.56; 95% CI: 0.40 to 0.73 [3 studies]); global cognitive function (SMD: 0.36; 95% CI: 0.09 to 0.63 [2 studies]).

In the other meta-analysis (Semkovska & McLoughlin, 2010), the authors' primary research question was to examine the evidence for cognitive impairment following ECT. The authors identified 1,525 articles after removing duplicates and included 84 studies in

the review. The articles contained 22 standardized neuropsychological tests spread over eight cognitive domains (global cognitive status, processing speed, attention/working memory, verbal episodic memory, visual episodic memory, spatial problem solving, executive functioning, intellectual ability). The authors created 24 strata based on the type of test and grouped the studies into one or more of these strata. The strata were further sub-divided into three time periods based on the interval between the final ECT session and the last administration of a cognitive test (0 – 3 days, 4 – 15 days, > 15 days). Statistically significant decreases in cognitive performance were observed in the 0 – 3 day period in 72% of the variables, with effect sizes (Cohen's *d*) ranging from -1.10 (95% CI: -1.53 to -0.67) to -0.21 (95% CI: -0.40 to 0.01). In the 4 to 15 period, only one result suggested the presence of cognitive impairment (verbal paired associates delayed recall: Cohen's *d*: -0.36; 95% CI: -0.62 to -0.10 [4 studies]). Beyond 15 days, no results indicated the presence of cognitive impairment.

1.6. Cognition

Cognition concerns the mental processes required to gain knowledge and understanding from thoughts, experiences, and senses. Cognition also involves remembering knowledge and understanding, and being able to reason. Due to the multidimensionality of cognition, the construct has been divided into several domains. To assess cognition, many questionnaire-based and task-oriented instruments have been developed over time to measure cognitive processes in specific domains, as well as globally. The National Institutes of Health has identified 74 different instruments to measure cognition.

Cognitive deficits have been found in many psychiatric disorders, including schizophrenia, unipolar depression, obsessive-compulsive disorder, posttraumatic stress disorder (PTSD), attention-deficit/hyperactivity disorder, and borderline personality disorder {(Hart et al., 2008; Hasselbalch, Knorr, Hasselbalch, Gade, & Kessing, 2012; Iorfino et al., 2016; Murrough, Iacoviello, Neumeister, Charney, & Iosifescu, 2011; Parlar, 2015; Parlar, Frewen, Oremus, Lanius, & McKinnon, 2016a; Polak et al., 2012)}. Over the last 25 years, researchers have become interested in the neurobiology of cognition as a means of understanding the biological and pathophysiological processes linking cognition and psychiatric disorders (Brunoni et al., 2014; Lanius, 2015; Murrough et al., 2011; Parlar, 2015; Williams et al., 2016). Neurobiological evidence is starting to accumulate in the literature, thereby allowing researchers to gain greater insights into the interplay between MD and cognition. This evidence is generating new research as well, with one example being the use of brain circuit functioning to define new dimensions of psychopathology and develop a neural circuit taxonomy for mental disorder (Williams et al., 2016).

Cognitive deficits are manifested in persons with MDD. Cognition-related symptoms such as reduced ability to think, loss of concentration, or difficulty in decision making are hallmarks of MDD. Neuropsychological tests have shown people with MDD, compared to non-depressed controls, perform more poorly on the Trail Making Test, the Symbol Digit Modalities Test, and the Stroop test (Hasselbalch et al., 2012). Research has shown a relation between memory deficits and early depressive symptoms. Deficits in verbal

memory have been associated with the development of depressive symptoms (Iorfino et al., 2016). Additionally, researchers have reported a relation between a decline in executive function and more persistent depressive symptoms (Iorfino et al., 2016). Dysfunctional prefrontal-subcortical circuitry and associated challenges in emotion control are believed to explain cognitive deficits in persons with MDD (Murrough et al., 2011). The recent International Study to Predict Optimized Treatment in Depression (iSPOT-D) trial found that 8-week acute treatment with escitalopram, sertraline, or venlafaxine extended-release did not improve cognitive performance in attention, response inhibition, verbal memory, decision speed, and information processing, even in persons whose depression remitted (Shilyansky et al., 2016).

A meta-analysis by Rock et al., 2014 searched PubMed and Google Scholar between 1980 to December 2012 and included 24 studies comparing currently depressed patients to healthy controls and six studies comparing remitted depressed patients to healthy controls. The authors reported moderate deficits in memory, executive function and attention (Cohen's d effect sizes ranging from -0.34 to -0.65) in currently depressed vs. healthy controls; and memory deficits (Cohen's d ranging from 22 to 0.54) in remitted depressed patients. However, the I^2 for sixteen out of twenty four studies pool results range from 56 to 82 indicating substantial statistical heterogeneity between studies. Another meta-analysis by Lee et al., 2011 searched PubMed and PsychInfo databases from 1990 to February 2011 summarizing 13 studies. The authors reported that patients in their first episode of depression performed worse than healthy controls in attention (SMD:

0.36, 95% CI: 0.13-0.59; I^2=0%), working memory (SMD: 0.16, 95% CI: -0.20 - 0.51; I^2=61%), verbal learning and memory (SMD: 0.13, 95% CI: 0.18-0.45; I^2=81%), visual learning and memory (SMD: 0.53, 95% CI: -0.05 – 1.11; I^2=88%). Unfortunately, the heterogeneity of the pooled results for working memory, verbal and learning memory and visual and learning memory heterogeneity range from 61% to 88% which warns caution in interpreting the findings. In addition, a meta-analysis from Wagner et al., 2012 included 15 studies investigating severity of executive dysfunctions in persons with MDD in comparison to healthy controls and 3 before and after antidepressant treatment studies. The authors found that healthy controls had better cognitive functioning than persons with MDD in semantic and phonemic memory 0.92 SD and 0.71 SD, I^2=52%; Stroop interference test 1.18 SD, Trail Making Test B 1.109 SD. These previous meta-analyses did not strictly adhere to PRISMA guideline, nor did they evaluate the the risk of bias and did not measure the strength of evidence. Interpretation of these findings of should be done with caution due to moderate to high levels of heterogeneity (> 50%).

The primary cognitive deficits in BD are learning and memory, working memory, attention, inhibition, and cognitive control (I. E. Bauer et al., 2016). Functional neuroimaging has traced the potential source of these deficits to changes in neural activity in prefrontal, cingulate, and limbic regions, which occur during response inhibition, cognitive control, and affective processing (J. Bauer et al., 2009). In persons who are genetically predisposed to BD, inflammatory mediators and oxidative stress dysregulate hormonal, metabolic, and circadian homeostasis to increase susceptibility to (and severity

of) the disease (Muneer, 2016). Compared to controls, persons with BD have shown poorer cognitive performance on the DigitSpan Backwards, DigitSpan Forwards, Trail Making Test, and Stroop test (Torrent et al., 2006). A meta-analysis of 45 studies and 18 cognitive variables found that medication use contributed to psychomotor slowing in persons with BD (Bora, Yucel, & Pantelis, 2009).

1.7. Mood Disorders and Post-traumatic Stress Disorder

Persons with MD or post-traumatic stress disorder (PTSD) display impaired performance on the same fronto-temporally mediated cognitive functions, including executive functioning, verbal recollective memory, attention, and processing speed (Parlar, Frewen, Oremus, Lanius, & McKinnon, 2016b). A meta-analysis of 18 studies showed that persons with PTSD displayed poorer executive functioning than controls (Polak et al., 2012). A different meta-analysis of 113 studies comparing persons with MDD to controls also found poorer executive functioning in the diseased group (Snyder, 2013). Indeed, many persons with MDD report a history of trauma. In a sample of 2,000 persons with anxiety or depression, 91.2% claimed to experience a traumatic or troublesome past event (Spinhoven et al., 2014).

Brain alterations in the neural correlates of social cognition—empathy toward others—are evident in persons with MD and PTSD. Specifically, these alterations concentrate in the areas of higher-order cognitive and affective processing, e.g., dorsolateral prefrontal cortex, ventromedial prefrontal cortex anterior cingulate cortex, the amygdala, and the temporoparietal junction (Parlar, 2015).

Dissociation is another common feature of MD and PTSD (Parlar et al., 2016b). Dissociation is characterized by a lack of connectivity between a person's thoughts, memory, and identity. In a recent study of 23 persons with MDD and a trauma history, dissociation was found to be a transdiagnostic risk factor for neuropsychological dysfunction. Derealization was associated with poorer verbal and visuospatial memory recognition, while depersonalization was associated with slower processing speed (Parlar et al., 2016b).

Further work has shown connections between the insula, amygdala, and the pathophysiology of PTSD in non-dissociative and dissociative subtypes (PTSD+DS) during symptom provocation. Relative to controls, persons with the dissociative and non-dissociative PTSD subtypes showed increased insula connectivity, either to basolateral amygdala clusters in both hemispheres (non-dissociative) or the left basolateral amygdala complex (dissociative). Persons with dissociative PTSD demonstrated increased insula subregion connectivity to the left basolateral amygdala compared to persons with non-dissociative PTSD (Nicholson et al., 2016).

In many studies of MD, the presence of trauma is not well studied. In Chapter 2, our analyses of the impact of PIMM/SAM on medication adherence included adjustment for the presence of PTSD.

1.8. Systematic Review and Meta-analysis of the Effects of ECT on Cognition

Due to continuing uncertainty over the impact of ECT on cognition in persons with TRD, a systematic review and meta-analysis was conducted on this topic. Chapter 4 reports the

results of this review and meta-analysis. The last review in this area was published in 2015 (Tor et al., 2015); however, this review and meta-analysis was narrowly focused on comparing two specific types of ECT, namely ultrabrief versus brief pulse ECT, and therefore did not cover the full spectrum of literature on cognitive performance post-ECT. In fact, this previous meta-analysis considered cognition only as a secondary outcome, whereas the meta-analysis reported in this thesis examined cognitive performance post-ECT as the primary outcome.

A large-scale meta-analysis investigating ECT and cognition was published in 2010 (Semkovska & McLoughlin, 2010). This review reported that worsening cognitive performance was limited to three days post-ECT; however, the reviewers meta-analyzed studies by cognitive test within specific cognitive domains, thus overweighting studies that included multiple cognitive tests for the same domain. Additionally, the reviewers did not appear to consider inter-study differences in areas such as study samples or ECT modality when deciding which studies to include in the meta-analysis.

The meta-analysis reported in Chapter 4 provides an updated literature search and meta-analysis that considers all of the relevant literature through August 2015. The review focuses on two widely-defined ECT modalities, namely bilateral and unilateral ECT to treat persons with TRD, and seeks to investigate differences in cognitive performance (by cognitive domain) between persons receiving these therapies at several time points: pre-ECT, 1-7 days post-ECT, 8-30 days post-ECT, 31-183 days post-ECT, and 184-365 days post-ECT. The meta-analysis avoids overweighting studies by

including studies only once in each possible analytical stratum (cognitive domain/time point).

Unlike the two other meta-analyses discussed above (Semkovska & McLoughlin, 2010; Tor et al., 2015), the meta-analysis in Chapter 4 assesses the risk of bias of all included studies and grades the strength of evidence for impaired cognitive performance by cognitive domain. Risk of bias and grading the strength of evidence are essential components of assessing the body of knowledge on a topic because conclusions about the evidence must be filtered by the degree to which the studies' results are valid and convincing. The reporting of results is incomplete unless reviewers address the underlying validity of the included studies.

The systematic review and meta-analysis protocol has already been published (C. Oremus et al., 2015), as has some preliminary methods work examining interrater and test-retest reliability in ECT reviews (M. Oremus, Oremus, Hall, & McKinnon, 2012) (see Chapter 3).

1.9. Interrater and Test-retest Reliability in ECT Reviews

One of the studies reported in Chapter 3 (M. Oremus et al., 2012) examined the inter-rater and test-retest reliability of risk of bias assessments conducted by inexperienced student raters. The raters received a training session on risk of bias assessment and independently rated this risk for 13–20 articles that were relevant to the topic of the systematic review and meta-analysis reported in Chapter 4. The findings suggested the need for raters to be trained in the assessment of risk of bias. Therefore, to prepare for the review reported in

this thesis, raters were trained in the use of risk of bias assessment tools and they received instruction on the nuances of study design to enable them to validly extract data from the included studies.

1.10. Evidence-based Medicine

The approach adopted here relies heavily upon the principals of evidence-based medicine (EBM). Specifically, EBM, sometimes referred to as evidence-based practice, involves the use of the best available healthcare evidence to inform clinical practice and health policymaking (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). From a clinical perspective, practitioners of EBM will use healthcare evidence in the form of systematic reviews and meta-analyses, randomized controlled trials, and observational studies to help determine the most appropriate means of treating individual patients. This evidence is not meant to be a substitute for the clinicians' own experience and knowledge, nor is it intended to trump patients' own values or thoughts regarding treatment. Systematic reviews often serve as the basis for developing clinical practice guidelines. For example, a large systematic review of dementia medications (Raina et al., 2008) provided the American College of Physicians and the American Academy of Family Physicians with the evidence to develop guidelines for the pharmacologic treatment of dementia (Qaseem et al., 2008).

In the policy realm, many jurisdictions with publicly-funded healthcare systems have embraced EBM to help guide reimbursement decisions for drug and non-drug health technologies. The National Institute for Health and Care Excellence (NICE) in England

and Wales is a prime example of an organization whose mission includes the development of technology assessments to recommend whether the National Health Service (NHS) should start paying for (list) new health technologies or cease paying for (delist) existing technologies (National Institute for Health and Care Excellence (NICE), 2016). In Canada, the Canadian Agency for Drugs and Technologies in Health (CADTH) (Canadian Agency for Drugs and Technologies in Health (CADTH), 2016) and its Common Drug Review (CDR) (Canadian Agency for Drugs and Technologies in Health (CADTH), 2014) produce similar assessments and recommendations. However, the NHS must list or delist health technologies based on NICE recommendations, while provincial governments in Canada may voluntarily decide whether to follow CADTH or CDR recommendations.

EBM has not escaped criticism. Critics have wondered whether an emphasis on clinical evidence would diminish the value of basic science and physicians' accumulated practical experience. Another concern has been the extent to which evidence from the atypical patients commonly recruited into medical studies applies to the patients regularly seen in the average physician's practice. Health policy analysts sometimes point to EBM as a means of controlling healthcare costs, with the provision of good health care being a secondary aim (Greenhalgh, Howick, & Maskrey, 2014).

Much of the tension in EBM surrounds the 'communal' aspect of the best available evidence versus the 'individual' aspects of the physician's own expertise or the needs of specific patients requiring treatment (M. R. Tonelli, 2011). Attempts to reconcile the

communal and individual components of EBM are ongoing through dialogue and discussion, and the differences between the two camps are less exaggerated than often thought (Dickersin, Straus, & Bero, 2007; McCartney, Treadwell, Maskrey, & Lehman, 2016; M. Tonelli & Guyatt, 2016). In 2007, the British Medical Journal rated EBM as one of the 15 greatest medical breakthroughs in the last century-and-a-half, along with the development of anesthesia, antibiotics, the Pill, and vaccines (British Medical Journal, 2016).

EBM replaced a regime where medical evidence played a relatively minor role in guiding healthcare practice and policy. However, other practice and policy guides, such as conventional wisdom, physicians' experience, and good basic science, were not always enough to establish the efficacy of treatment. For example, over the last two decades, evidence from high-quality randomized controlled trials has overturned entrenched beliefs about the efficacy of encainide and flecainide to treat asymptomatic arrhythmias, and the lack of efficacy of beta-blockers to treat heart failure (M. Tonelli & Guyatt, 2016).

In psychiatry, EBM was initially regarded with skepticism. The subjective nature of most psychiatric symptomatology led to the belief that research evidence could not be easily applied to the average patient. Further, the lack of evidence in favour of many second- or third-generation psychiatric medications produced contradictory conclusions: policy makers argued for continued funding of first-generation medications only, while proponents of newer medications pointed out methodological flaws in research studies as

a rationale for giving second- and third-generation medications further consideration (Emsley & Hawkridge, 2009). Despite the initial skepticism, an emerging consensus has arisen in psychiatry to include the best available research evidence in the clinical decision-making mix. Gray and Pinson described EBM for psychiatric audiences over ten years ago (Gray & Pinson, 2003), a British-based journal called *Evidence-based Mental Health* (http://ebmh.bmj.com/) is in its nineteenth volume of publication, and new methodological advances aimed at research synthesis in psychiatry emerge regularly (e.g., see recent work on the use of Bayesian statistics to analyze antidepressant trials in anxiety disorders (Monden et al., 2016)).

Systematic reviews are a form of literature search employed in EBM. Unlike basic literature searches, systematic reviews are rigorous and transparent literature searches guided by explicit research questions and formalized methodological procedures. Systematic reviews are undertaken to answer clinical research questions by searching for, obtaining, and summarizing all of the available evidence on the topics of interest. Additionally, systematic reviews rate the quality of the evidence to account for the fact that study quality affects the conclusions one can draw from the evidence. The Cochrane Collaboration (Higgins & Green, 2016) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009) outline the specific methods for systematic reviews. Meta-analyses are a type of systematic review that involves the statistical combination of results from individual studies to obtain a single summary estimate of effect across all

studies (Borenstein, Hedges, Higgins, & Rothstein, 2009). The work presented here incorporates each of these approaches, with the aim of informing clinical practice in the area of TRD. The results of the systematic review and meta-analysis may suggest to clinicians that certain types of ECT may entail more cognitive risks than other types. Therefore, clinicians can avoid the riskier types of ECT unless one of these types is specifically required to treat an individual patient.

For PIMM/SAM, the results of the RCT could point to a program that might boost medication adherence in persons admitted to an inpatient mood disorders program. Health professionals and partners in care may wish to emulate PIMM/SAM in their institutions if the final results of the RCT are favourable to this program.

CHAPTER 2:

PARTNERSHIP IN MEDICATION MANAGEMENT (PIMM): THE EFFECTS OF ONE-ON-ONE MEDICATION TRAINING ON MEDICATION ADHERENCE IN PATIENTS WITH MOOD DISORDERS: A PILOT STUDY

Foreword to Chapter 2

This chapter describes a pilot randomized controlled trial that is part of a sequential explanatory mixed-methods study investigating the effect of a novel personalized one-on-one medication training approach for persons with mood disorders. Failing to take medications as prescribed is a common risk factor for treatment failure in mood disorders. Antidepressants and lithium are common and effective treatments for mood disorders (MD). The efficacy of these medications, as reported in clinical trials, however, differs from clinical experience, a finding attributed in large part to non-adherence. From a scholarly perspective, little work has examined the efficacy of medication adherence programs in MDs, despite the fact that research in other disease areas suggests that medication training can improve adherence.

The work in chapter 2 has been submitted in July 24, 2016 to the *BioMed Central (BMC) Pilot and Feasibility Studies*. Chapter 2 contains the final manuscript that has been submitted to the journal.

Partnership in Medication Management (PIMM): The Effects of One-on-one Medication Training on Medication Adherence in Patients with Mood Disorders: A Pilot Study

Carolina Oremus[1,2,3], Mark Oremus[4,5], Ashok Chaurasia[4], Jennifer Lowe[3], Margaret C. McKinnon[1,2,3,7*], Sharon Simons[3]

[1]McMaster Integrative Neuroscience Discovery & Study (MINDS), McMaster University, Hamilton, Ontario, Canada

[2]Department of Psychiatry & Behavioural Neurosciences, McMaster University, McMaster University, Hamilton, Ontario, Canada

[3]Mood Disorders Program, St. Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada

[4]School of Public Health and Health Systems, University of Waterloo, Waterloo, Ontario, Canada

[5]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada

[7]Homewood Research Institute, Guelph, Ontario, Canada

*Corresponding author

CO: coremus@stjoes.ca

MO: moremus@uwaterloo.ca

AC: a4chaurasia@uwaterloo.ca

JL: jlowe@stjoes.ca

MCM: mmckinno@stjoes.ca

SS: ssimons@stjoes.ca

Abstract

Background: Mood disorders (MD), including major depressive disorder, bipolar disorder, and dysthymia/persistent depressive disorder are among the most common mental health conditions in Canada. High levels of medication non-adherence affect negatively treatment efficacy. We investigate the feasibility of conducting a sequential explanatory mixed-methods study, involving a 12-month randomized controlled trial (RCT) examining the effect of a partnership in medication management/self-administered medication (PIMM/SAM) program versus standard prescribing practice (SPP) on medication adherence in persons with MD.

Methods: We recruited English-speaking persons, aged 18 years or older with a primary diagnosis of bipolar disorder I or II, major depressive disorder, or dysthymia from an inpatient mood disorders unit. We excluded persons with cognitive impairment, significant suicidal or homicidal risk, or brain injuries. Participants were randomized to the PIMM/SAM or SPP group and assessed at baseline. A second assessment occurred within two days of discharge. The primary outcome was medication adherence, assessed using the Medication Adherence Rating Scale (MARS) and the Beliefs about Medications Questionnaire (BMQ). Secondary outcomes, assessed using a battery of scales, included anxiety, depression, dissociation, self-efficacy, psychiatrist-patient relationship, and health-related quality-of-life. For each scale, we regressed change scores (pre-discharge –

baseline) onto participants' group assignments to obtain mean change score differences (MCSD) between groups.

Results: Seven participants randomized to PIMM/SAM and five participants randomized to SPP (out of eight initially randomized to each group) completed the baseline and pre-discharge interviews. Between-group differences on the MARS were not statistically significant. However, relative to participants in the SPP group, participants in the PIMM/SAM group held fewer negative beliefs about medications (MCSD: -4.9; 95% confidence interval: -9.0 to -0.8) and had lower depersonalization (MCSD: -3.7; 95% confidence interval: -6.7 to -0.8) scores on a transdiagnostic measure.

Conclusions: This study provides the first evidence demonstrating the feasibility of conducting an RCT to evaluate the impact of a medication education program on treatment adherence in persons with MD. The pilot work described in this paper generated lessons to carry forward to the RCT, including sample size targets, adaptations to staff workflow to enable program implementation, and staff input into training.

Trial registration: ClinicalTrials.gov identifier: NCT02285608; Registration date: October 28, 2014

Keywords: Mood disorders, Medication adherence, Medication education program

Background

Mood disorders (MD), including Major Depressive Disorder (MDD), bipolar disorder (BD), and dysthymia/persistent depressive disorder (American Psychiatric Association, 2013) are among the most common mental health conditions in Canada. Health Canada reported a lifetime prevalence of MD of 12.6% (Health Canada, 2002). Persons with mood disorders experience changes in many areas, including alterations in behaviour, cognitive/body functions, thoughts, studies/education, and family and social interactions (American Psychiatric Association, 2013). Financially, MD is associated with lost workplace productivity, high healthcare costs, and significant loss of income affecting families and the economy. The cost of medication also affects family finances (Health Canada, 2002). According to the World Health Organization (WHO), depression—a major MD—is a leading cause of global disability. Approximately 350 million people suffer from depression worldwide (World Health Organization, 2008). For BD, the lifetime prevalence in11 countries ranges from 0.1% in Nigeria to 3.3% in the United States (Merikangas et al., 2007).

In addition to their core affective components, mood disorders are associated with cognitive deficits that persist following the resolution of major depressive episodes in some patients, and show worsening with subsequent episodes of illness (MacQueen et al., 2003). Critically, cognitive dysfunction impacts negatively on the outcome of pharmacological and non-pharmacological treatments for mood and related affective

disorders. This occurs primarily when the ability to engage in and successfully complete treatment relies heavily on cognitive processes (Dunkin et al., 2000; Polak et al., 2012)

Despite the availability of numerous psychopharmacological treatments, evidence indicates that only 60 to 70% of persons who tolerate anti-depressants will respond to first-line drug therapy for MDD (Souery et al., 2006). Lithium is a common and effective treatment for BD. However, the efficacy of anti-depressants or lithium reported in clinical trials differs from clinical experience (Rosa et al., 2007). Various factors such as non-adherence to treatment, poor tolerability to medications, and medical and psychiatric comorbidities have been related to treatment non-response or treatment failure in depression (Nemeroff et al., 2003).

The WHO defines treatment adherence as "the extent to which a patient follows medical instructions" (World Health Organization, 2008). This definition also refers to various health-related behaviours that extend beyond taking medications, including seeking medical attention, filling prescriptions, proper medication intake, and attending follow-up appointments. The WHO also recognizes the importance of the quality of the relationship between patients and healthcare providers in treatment adherence (World Health Organization, 2008).

Premature discontinuation of treatment for mood disorders is common. Lengthily treatments, patients' beliefs about medications, a lack of knowledge about the purposes of medications/treatments, benefits, dosages, and adverse effects, as well as the relationship between patients and healthcare providers, all affect treatment continuation [3].

Objectives

Here, we investigate the effect of a novel one-on-one nurse-led medication training program called Partnership in Medication Management/Self-Administered Medication (PIMM/SAM) on medication adherence in persons with MD. We developed a sequential explanatory mixed methods study involving a 12-month randomized controlled trial (RCT) to quantitatively investigate the differences between PIMM/SAM versus standard prescribing practice (SPP) on medication adherence in persons with MD. Diagnostic status was confirmed using a semi-structured diagnostic interview that also established the presence of common co-morbidities (e.g., PTSD).

We will conduct the qualitative portion of the study following the completion of the RCT. For the qualitative portion of the study, we will randomly recruit a purposeful sample of participants who scored seven or less on the Medication Adherence Rating Scale (MARS) during the clinical trial. The qualitative portion of this study will explore the reasons for low medication adherence and the factors that might explain the differences and similarities in medication adherence, beliefs and knowledge about medications, patient-psychiatrist or patient-therapist relationship, and satisfaction with quality of care. We will also investigate the participants' own personal experiences and views regarding inpatient and outpatient programs, services, and care, and any medication training or instruction they received during hospitalization or visits to the outpatient clinic. Lastly, we will examine if the medication training or instruction that participants received helped them to feel more confident or empowered to take more responsibility for

their medication, health and well-being after hospitalization. Additional qualitative questions will be generated based on the quantitative findings, current knowledge or research on mental health and medication adherence, and comments or observations from study participants.

Furthermore, we will conduct a health economics assessment from the health system perspective. The health economics assessment will investigate the costs of first re-hospitalization, length of re-hospitalisation, daily use and costs of direct medical resources between each study group.

From the start of RCT, we monitored the recruitment and retention of participants, the study process, and the conduct of the study. This monitoring was important because potential decreases in cognitive abilities (e.g., decreased concentration, difficulty in making decisions, and decreased memory) among persons with MD could affect participation in the study. Further factors that could affect participation in this population include irritability, lack of energy, decreased drive to engage in activities, decreased enjoyment and interest in previous activities, the presence of persecutory delusions, and the presence of other negative symptoms such as a lack of eye contact and unemotional speech (Ahern, McKinnon, Bieling, McNeely, & Langstaff, 2016). In our experience, we have found that some inpatients may also misunderstand the benefits and harms of participating in research. For example, inpatients may agree to participate in studies because they believe doing so will provide them with a better standard of care. Others may believe a refusal to participate could adversely affect their treatment or create

antagonism with hospital staff. A clear explanation of the informed consent process typically clarifies these issues, but may lead to reversals in the decision to participate.

We developed a study protocol, case report forms and training materials for study participants and clinical staff (e.g., guidelines, checklists, instruments and advertisement), and we trained the clinical staff from the MD inpatient unit to assist with identification, training, and monitoring of participants. In addition, we created a computer-generated randomization sequence and recruited participants into the study. The primary outcome is medication adherence and the secondary outcomes include anxiety, depression, negative and positive beliefs about medication use, dissociation, self-efficacy, participant-psychiatrist relationship, and health-related quality-of-life (HRQoL). To the best of our knowledge, no other study has examined the impact of this type of personalized and interactive daily one-on-one medication education program on medication adherence in persons with MD. At the same time, we are investigating the personal, financial, psychological and social reasons behind poor adherence.

We felt a pilot study was necessary given the challenges inherent in recruiting and following a group of persons with MD. Some challenges include potentially high levels of distractibility, illness-related decreases in motivation and/or energy, or itinerancy at follow-up. Recruitment and retention may also be affected by changes in cognitive functioning, which persons in our study population can experience. Some of the lessons learned during this pilot study enabled us to modify the original protocol and promote both recruitment and retention of participants in our study.

We have registered the RCT protocol at ClinicalTrials.gov (identifier: NCT02285608) and we report the pilot results in line with the CONSORT guidelines for reporting RCTs (Additional file 1) (Moher et al., 2010).

Methods

Study Setting and Participants

We recruited participants aged 18 years or older from the inpatient unit of the Mood Disorders Program at St. Joseph's Healthcare Hamilton (Hamilton, Ontario, Canada). The inpatient unit houses 24 beds and provides clinical care and treatment for persons with MD. The unit also prepares inpatients to return to the community. The inpatient program employs treatments such as pharmacotherapy, stress management, individual and group psychological therapy (i.e., cognitive behaviour therapy), recreational, physical, occupational and arts therapy, and provides supports around activities of daily living. Eligible participants had a primary diagnosis of BP I or II, major depressive disorder, or dysthymia/ persistent depressive disorder (American Psychiatric Association, 2013). We used the Mini International Neuropsychiatric Interview (MINI) to confirm diagnosis (Lecrubier et al., 1997; Sheehan et al., 1998). Participants also had to speak, read and understand English. We excluded persons with significant suicidal or homicidal risk, any medical condition known to affect the brain, or acquired brain injury. The attending psychiatrists reviewed the Montreal Cognitive Assessment (MOCA) (Nasreddine et al., 2005) scores and made the final determination of eligibility for each participant.

Recruitment and Study Implementation

Recruitment for this parallel group, 1:1 allocation began in January 2015. Recruitment is ongoing and we are reporting our experiences with the study after having enrolled 16 participants by the end of May 2015. Our early experience with recruitment and follow up required the submission of three ethics amendments and the re-thinking of some study processes. These endeavours convinced us to assess the operationally of the study processes, analyse the existing data, and based on our results to continue to move forward with further recruitment.

Attending physicians assigned to the inpatient unit identified potential participants based on the inclusion/exclusion criteria. Trained staff met with potential participants to conduct the screening MOCA. CO administered written informed consent to eligible participants. Once consent was obtained, CO booked an appointment to meet with the study participant the next day to conduct the baseline interview. CO also contacted an independent hospital administrative assistant to determine group allocation. Only the independent hospital administrative assistant had access to the computer-generated randomization list. The next day, CO met with the study participant to administer the baseline interview and disclose the group allocation. She trained the nurses and provided them with guidance to ensure the implementation of the proper program for each participant. Once attending psychiatrists determined a participant would be discharged, nursing staff informed CO, who would administer the pre-discharge (follow-up) interview within two days of discharge.

A co-investigator (MO) independently developed the 1:1 randomization sequence using SAS v9.4 (The SAS Institute, Cary, NC). The list was emailed directly to the independent hospital administrator assistant and it was not shown to the hospital-based investigators (CO< SS< MCM). The administrative assistant kept the sequence in a locked cabinet to preserve allocation concealment. Due to the nature of the interventions, the study was un-blinded following participants' allocation to the study groups.

Programs

Intervention (PIMM/SAM)

PIMM/SAM involves individual, one-on-one sessions between participants and nurses. At the first education session, nurses and CO met with participants and asked how they administered medications at home (e.g., blister pack). At the same session, the nurses gave participants information about the appearance, dosage, purposes, benefits, common adverse effects, and administration schedules for each medication. CO sat in on each nurse's first education sessions to observe, provide guidance, and ensure proper program delivery. During the first education session, CO and participants discussed the goals and steps of the program, strategies to improve medication adherence and established reminders to take medications as prescribed, both in the hospital and at home. Participants were also provided with a choice of notebooks, pens, highlighters, post-its, alarm clocks, checklists, and referrals to other sources (i.e., online apps) to record information about medications and trigger reminders to take medications. CO also encouraged participants to write down and study all relevant information about their medications. Following the

education session, participants were responsible for notifying the nurses within an hour of the time when they had to take their medications. After the notification, the participants met with their nurses and were shown all of the medications that they were required to take at that time. Participants then identified each of their medications and described the benefits, purposes, dosage, common adverse effects of each medication, and the importance of continuing to take the medications and following their treatment plans.

A primary aim of the PIMM procedure was to mimic, as much as possible, in the inpatient setting the conditions under which participants would take their medications at home. Accordingly, participants were asked to select those reminders most consistent with the procedure they would follow at home to remind themselves to take their medication (i.e., use of an alarm clock). Participants and CO worked together to devise new strategies and reminders that would encourage medication adherence once participants returned home from hospital.

Comparator

SPP is the same standard of care that the typical patient receives in the inpatient unit. In this group, nurses are in charge of administering medications. Participants in the SPP group did not receive personalized, one-on-one medication training program or any other standardized strategy to improve medication adherence.

Diagnostic assessment

The Mini International Neuropsychiatric Interview (MINI) is a short structured interview designed to help researchers make any one of 17 current psychiatric diagnoses including

mood disorders, a range of anxiety disorders, substance dependence or abuse, eating disorders, a range of psychotic disorders and antisocial personality disorder. (Lecrubier et al., 1997). For each disorder, 'no' answers to one or two screening questions with dichotomous 'yes/no' response options rule out the presence of the disorder in question. Diagnosis of MD and comorbidities, including PTSD, were confirmed with the MINI. Severity of depressive and anxiety symptoms was assessed using the BDI and BAI, respectively. Transdiagnostic dissociative symptoms were measured using the MDI.

Study questionnaire

Primary outcome

The Medication Adherence Rating Scale (MARS) (Thompson, Kulkarni, & Sergejew, 2000) is a self-report 10-item scale that assesses adherence to the daily prescribed medication intake. Answers to each question are dichotomized (yes=1, no=0). Scores above 7 are considered good adherence.

The Beliefs about Medicines Questionnaire (BMQ) (Robert Horne, Weinman, & Hankins, 1999) investigates participants' beliefs about their current medications and general attitudes to medications; the scale will be used to quantitatively assess participants' perceptions about medications and their reasons for adherence or non-adherence to medications. BMQ has been validated as a good predictor of medication adherence (Rob Horne et al., 2013; Robert Horne et al., 1999).

Secondary outcomes

The Beck Anxiety Inventory (BAI) (Beck, Epstein, Brown, & Steer, 1988) is a self-report 21-item scale that measures the severity of anxiety symptoms. Each question measures the intensity of the symptom on a 0 (not present) to 4 (too intense) scale. Beck Depression Inventory-II (BDI-II) (Beck, Steer, Ball, & Ranieri, 1996) is a self-report 21-item scale that measures the severity of depressive symptoms. Each question measures the intensity of the symptom on a 0 (not present) to 4 (intense) scale.

The transdiagnostic Multiscale Dissociation Inventory (MDI) (Briere, 2002) is a 30-item self-report test that assesses dissociative symptomatology. The MDI measures five domains of dissociative behaviour: Disengagement (DENG), Depersonalization (DEPR), Derealization (DERL), Emotional Constriction (ECON), Memory Disturbance (MEMD), and Identity Dissociation (IDDIS). The MDI yields six scores, one for each subscale, and one score for total dissociative symptomology.

The General Self-Efficacy Scale (GSE) (Schwarzer & Jerusalem, 1995) is a 10-item self-report scale that assesses individual's perceived self-efficacy. The GSE aims to predict the individual's coping and adaptation after experiencing stressful life events.

The patient version of the Revised Helping Alliance Questionnaire for Treatment with Psychiatrists (HAQ-PC) (Luborsky et al., 1996) is a 19-item self-report questionnaire that assesses different aspects of the psychiatrist-patient relationship such as patients' motivation for treatment and perception of the psychiatrist. Each answer ranges from 1 (strongly disagree) to 6 (strongly agree).

Short Form-36 Health Survey (SF-36), physical and mental health subscales (Ware & Sherbourne, 1992), will measure participants' physical and mental HRQoL. The SF-36 is a self-report questionnaire that measures the individuals' own perceptions of their health status and functioning. The physical health subscale consists of 4 dimensions (physical functioning, role-physical, bodily pain and general health) that measure the burden of disease on day-to-day activities. The mental health subscale is a 14-item self-report survey that measures the burden of disease and the benefits of treatment. The mental health subscale consists of 4 dimensions (vitality, social functioning, role-emotional, and mental health) and is graded on a 5-point scale ranging from 0 (none of the time) to 4 (all of the time).

The Repeatable Battery for the Assessment of Neuropsychological Functioning (RBANS) includes index measures of Immediate Memory, Visuospatial/Construction, Language, Attention and Delayed Memory (Randolph, 1998). The RBANS is a reliable measure of neuropsychological functioning in dementia (Mohr, Walker, Randolph, Sampson, & Mendis, 1996) and schizophrenia (Gold, Queern, Iannone, & Buchanan, 1999).

Statistical analysis

We used the Shapiro-Wilk test to assess whether the scores on each of the study questionnaires came from a normally distributed population. Rejecting the null hypothesis at $\alpha = 0.05$ was evidence that the scores were not from a normally distributed population. In the case of non-normal scores, we reported medians and $25^{th}/75^{th}$ percentiles; when scores were normal, we reported means and standard deviations.

We reported our results according to the group to which participants were randomized. We did not experience any unintended cross-overs. To account for participants' baseline status on each study questionnaire, we computed a change score by subtracting the questionnaire score at baseline from the questionnaire score at pre-discharge. For each questionnaire or subscale thereof, we employed linear regression and regressed participants' change scores onto their group assignment to obtain a mean difference in change score between the PIMM/SAM and SPP groups. Due to our small sample size, the p-values for the mean difference in change score were calculated using ANOVA and linear regression in the case of normally distributed data. If the data were not normally distributed, then we performed a Kruskal-Wallis test and used the resultant p-value. Also, we independently analyzed whether the presence of comorbidities (Y/N), anxiety disorders (Y/N), or post-traumatic stress disorder (Y/N) or number of comorbidities (mean PIMM/SAM: 3.0; SPP; 1.7) has an effect on medication adherence, beliefs about medication, and dissociation. We used R v3.2.4 (R Foundation for Statistical Computing, Vienna, Austria) to conduct all statistical analyses.

Ethics

We obtained ethics approval from the Hamilton Integrated Research Ethics Board (#14-733). All participants provided written informed consent prior to recruitment and randomization.

Results

We randomized 16 participants and achieved a balanced number of eight per group. The mean age was 39.1 years (17.1) in the PIMM/SAM group and 54.8 years (13.7) in the SPP group. The difference in age between the two groups was not significant ($p > 0.05$). Four participants in the PIMM/SAM group and four participants in the SPP group were female. Three PIMM/SAM participants had at least some college education and two had high school or less education; three SPP participants had college education and two had high school or less education (education data were unavailable for the remainder of the sample). Median MOCA scores were 25.0 ($n = 5$) in the PIMM/SAM group and 24.5 ($n = 4$) in the SPP group (MOCA data were unavailable for the remainder of the sample). Data on number of previous depressive episodes were not available for all participants; however, four participants in the PIMM/SAM group reported two or more lifetime episodes and four participants in the SPP group reported two or more lifetime episodes. In the PIMM/SAM group, six participants were diagnosed with BD and two were diagnosed with MDD. In the SPP group, three participants were diagnosed with MDD, four were diagnosed with BD and the diagnosis was unavailable for one of the drop-outs. All of the participants were going through a major depressive episode. The median length of time between the baseline and pre-discharge assessments was 34 days for the PIMM/SAM group and 63 days for the SPP group. This difference was not significant ($p > 0.05$), likely due to sample size.

Participants' baseline sample characteristics are shown in Table 1. There were no statistically significant differences between groups in any of the tests performed at

baseline. Proper randomization helps to ensure that between group differences are eliminated at baseline. Although the means appear to be different the differences are not statistically significant because of the sample size. Seven PIMM/SAM and five SPP participants provided pre-discharge data on the nine outcome measurement instruments employed in the study. Figure 1 shows the flow of participants through the study. One participant in the SPP group withdrew due to a self-reported lack of energy, another was no longer suitable to continue in the study due to cognitive deterioration, and a third was discharged before the inpatient unit could notify CO to make arrangements to administer the pre-discharge interview. One participant in the PIMM/SAM group was also discharged before the inpatient unit informed CO.

Critically, SPP participants held stronger negative beliefs about medications than PIMM/SAM participants, with the mean General BMQ (G-BMQ) change score being 4.9 points higher in the SPP group (95% CI: 0.8 to 9.0; $p = 0.041$). The effect of negative beliefs in SPP participants was exacerbated after we added PTSD to our linear models. The mean G-BMQ change score with PTSD included increased to 5.8 in the SPP group versus PIMM/SAM group (95% CI: 1.3 to 10.2; $p = 0.03$).

Similarly, including PTSD in a model with the Specific Concern BMQ (SC-BMQ) subscale led to a mean change score that was 4.0 points higher in the SPP group (95% CI: 0.57-7.43; $p=0.04$). Prior to the addition of PTSD, the mean SC-BMQ change score was not statistically significant (2.3; 95% CI: -1.7 to 6.2; $p=0.29$).

Given the presence of dissociative symptoms in depression (Parlar, Frewen, Oremus, Lanius, & McKinnon, 2016) and in bipolar disorder ((Hariri et al., 2015) and their association with a longer illness duration and reduced treatment efficacy (Nuller, 1982), dissociation was assessed transdiagnostically and change measured. Over the course of follow-up, participants in the PIMM/SAM group demonstrated lower depersonalization on the MDI – depersonalization (DEPR) subscale (assessing the extent to which someone experiences the symptoms of depersonalization, i.e., disengagement from self and surroundings) than did their counterparts in the SPP group (Table 2). The mean change score on the MDI-DEPR subscale was 3.7 points lower in PIMM/SAM versus SPP participants (95% confidence interval [CI]: -6.7 to -0.8; p = 0.033). We investigated the effect of PTSD on the mean MDI-DEPR change score. We found that persons in the PIMM/SAM group had a score of 4.40 points lower than those in the SPP group with PTSD included (95% CI: -7.5 to -1.2; p=0.021).

Many of the confidence intervals were bounded very close to the null values, which suggests we may obtain statistically significant results after recruiting more participants. In addition, the p-values associated with these confidence intervals were less than 0.10 (Table 2). The comparisons involving-bounded confidence intervals suggested poorer outcomes for the SPP group versus the PIMM/SAM group: higher anxiety (BAI mean difference in change score: 12.7 [95% CI; 0.8 to 24.5; p =0.074); poorer positive relationships with psychiatrists (HAQ mean difference in positive relationship subscale change score: -13.8 [95% CI: -28.6 to 1.0; p = 0.097); stronger negative relationships with

psychiatrists (HAQ mean difference in negative relationship subscale change score: 6.8 [95% CI: -0.01 to 13.6; p = 0.079); more reported negative side effects and attitudes toward psychiatric medication (MARS factor 3): 1.3 (95% CI: -0.05 to -2.5; p=0.06); and lower health-related quality-of-life in the domain of general health (mean difference in SF-36 general health subscale change score: -22.5 [95% CI: -44.4 to 0.06; p = 0.076]). Number of comorbidities, the presence of anxiety disorder (Y/N) or comorbidities (Y/N), were not found to have an effect on medication adherence.

Discussion

The findings of the between-group comparison are certainly encouraging from an efficacy standpoint. The PIMM/SAM program may confer benefits compared to SPP on certain secondary outcomes, including fewer negative beliefs about medications (MCSD: -4.9; 95% confidence interval: -9.0 to -0.8) and a reduction in symptoms of depersonalization (MCSD: -3.7; 95% confidence interval: -6.7 to -0.8). There were no statistically significant differences between groups at baseline. Although the means appear to be different the differences are not statistically significant because of the sample size. Furthermore, proper randomization helps to ensure that between group differences are eliminated at baseline. Given the small sample size, the absence of findings on other outcomes cannot be taken as evidence for or against the intervention. It is, however, encouraging that many of the confidence intervals were bounded very close to the null values on a range of critical measures, including those assessing general health, quality of relationship with treating psychiatrist, and levels of anxiety.

The main purpose of the pilot study was to explore whether the challenges inherent in recruiting and following a group of persons with MD would affect further the conduct of our study. Our findings show that recruitment and retention will definitely be issues as the RCT progresses. Three out of twenty five persons initially approached to participate in the pilot study could not be enrolled. Further, two participants dropped out after randomization. Two participants were discharged before making arrangements for the pre-discharge interview. To obtain a suitable minimum sample size for the analysis of final results in the later RCT, we will need to randomize 35% more participants over and above the minimum sample size requirement. Further, we will have to increase our recruitment goal by 20% to ensure enough participants are randomized in the first place. A preliminary sample size calculation for the primary outcome of medication adherence suggests we would need 128 participants (64 per group) to detect a medium effect size (Cohen's $d = 0.5$) on the MARS. This calculation is based on a two-tailed t-test with 80% power and a 5% level of significance. To obtain 128 participants in the final analysis at pre-discharge, we would be required to randomize approximately 174 participants. To enable randomization of this number of people, we would have to recruit approximately 209 individuals. The catchment area of the Mood Disorders Program includes a population of 750,000 people. Given a lifetime MD prevalence of 12.6%, the territory should yield enough participants to fulfill our recruitment and randomization targets within a two-year period.

One issue we noticed during the feasibility study was reluctance among some nursing staff to participate in the trial. The nursing staff felt the study added to their daily workload. As well, the nursing staff expressed concerns about the learning curve required to implement the study protocol. For the future progress of the RCT to be successful, staff workflow will have to be modified to more seamlessly integrate the additional requirements of PIMM/SAM into nurses' daily routines. The Mood Disorders Program manager (SS) is aware of this issue and protocols will be developed to facilitate such integration.

Training is another issue that will require further thought. CO developed extensive training materials and held several training sessions with nursing staff to explain the background and objectives of the study, the intervention and comparator programs, and the nurses' study-related duties (e.g., deliver PIMM/SAM, notify one of the study leads of impeding participant discharges). Despite management's strong support of the study, training sessions were not always well attended by on-duty staff and some nurses were unavailable to attend any of the sessions due to shift schedules, workload, or legitimate work absences. CO was required to spend large amounts of time, often fluctuating between 8AM and 11PM, in the inpatient unit to ensure the nursing staff was following the protocol. For the continuation of the RCT, nurses' input into the scheduling of the training sessions will be required to promote successful implementation of the study. To accommodate nurse availability, CO occasionally led training sessions with individual nurses.

The PIMM/SAM program utilized in the present study contains components of other programs that have been designed to improve medication adherence in BD. A recent systematic review of 26 studies of interventions to improve antidepressant medication adherence found that the most successful interventions were multifactorial in nature. These successful interventions employed strategies involving mental health professionals, education, telephone monitoring of participants' progress, ongoing support of participants, participants' inclusion in the process of taking medications, and feedback of participant progress to partners in care. Education alone, without monitoring and feedback, was unsuccessful in boosting adherence. The authors of the review concluded that better adherence to antidepressant drug therapy would require behavioural modification through structured programs that reach beyond didactic education sessions or the provision of reading materials (Chong et al., 2011). Our preliminary findings suggest the PIMM/SAM program can have some benefits with respect to promoting medication adherence in persons admitted to inpatient units for MD. Continuance of the RCT, plus the qualitative investigation and the health economics evaluation, will provide further data to assess the efficacy of PIMM/SAM.

The benefits of programs such as PIMM/SAM may in part accrue from the greater amount of therapeutic contact relative to standard practice settings. Although we found more positive beliefs about medications and a decrease in depersonalization scores in the PIMM/SAM group, we do not believe these results can be explained by therapeutic contact alone. Indeed, our results did not show differences in anxiety, relationships with

psychiatrists, side effects and attitudes toward psychiatric medications, and lower health-related quality-of-life in the domain of general health. The differences for the beliefs about medication scale and depersonalization scale were expected given the domains targeted by PIMM/SAM (i.e., reducing negative views about medications and humanizing participants' inpatient experiences). Still, future studies should control for the level of clinical staff contact between groups.

Conclusions

This study provides the first evidence showing the feasibility of conducting an RCT within a larger program of research to evaluate the impact of a medication education program on treatment adherence in persons with MD. Interestingly, PTSD status impacted the preliminary findings, pointing towards the importance of assessing this frequently co-morbid condition. The results of the RCT will have important implications for medication prescribing practices not only in psychiatry, but also in other areas of medicine. The feasibility work described in this paper provides us with suggestions regarding sample size targets, staff workflow, and training. These lessons will be carried forward to the larger planned sequential explanatory mixed methods study. Most importantly, staff input into workflow and training will be essential to the smooth conduct of the RCT.

List of abbreviations

BAI: Beck Anxiety Inventory; BD: bipolar disorder; BDI-II: Beck Depression Inventory-II; BMQ: Beliefs about Medicines Questionnaire; GSE: General Self-Efficacy Scale;

HAQ-PC: Revised Helping Alliance Questionnaire for Treatment with Psychiatrists-Patient Version; HRQoL: health-related quality-of-life; MARS: Medication Adherence Rating Scale; MD: mood disorder; MDD: major depressive disorder; MDI: Multiscale Dissociation Inventory; PIMM/SAM: partnership in medication management/self-administered medication; RBANS: Repeatable Battery for the Assessment of Neuropsychological Functioning; RCT: randomized controlled trial; SF-36: Short Form-36 Health Survey; SPP: standard prescribing practice; WHO: World Health Organization

Declarations

Ethics approval and consent to participate

We obtained ethics approval from the Hamilton Integrated Research Ethics Board (#14-733). All participants provided written informed consent prior to recruitment and randomization.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are not publicly available because the small number of participants (all recruited from one site) raises the possibility that individual participants could be identified.

Requests to access the data for the purposes of verifying the information presented in this article should be addressed to the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

The present study was funded by a grant from the Canadian Institutes of Health Research (CIHR) to MCM. The funding body played no role in the design of the study and collection, analysis, and interpretation of data, and in writing the manuscript.

Authors' contributions

SS, CO, and MCM conceived the study. CO, MO, and MCM designed the study and study protocol. CO developed the case report forms and training materials, trained hospital staff, and collected and analyzed the data. MO, SS, and MCM reviewed and revised the protocol and study forms. AC, MO, MCM, and CO developed the statistical analysis plan and analyzed the data. CO and MO drafted the manuscript. AC, SS, JL, and MCM reviewed and revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

The authors wish to thank the participants for generously donating their time to participate in this study. The authors also thank the clinical and administrative staff of the inpatient unit of the Mood Disorders Program at St. Josephs Healthcare Hamilton for identifying potential participants, administering the PIMM/SAM and SPP programs, and monitoring participants to pre-discharge.

Tables & Figures

Table 1 Baseline sample characteristics

Characteristic	Group	Mean (SD) or Median (25^{th}/75^{th} percentiles)[a]
BAI	PIMM/SAM	28.0(2.5)
	SPP	28.2(14.7)
BDI	PIMM/SAM	16.7(9.6)
	SPP	32.0(20.5)
BMQ - General	PIMM/SAM	18.7(5.9)
	SPP	19.0(4.3)
BMQ - Specific Necessity	PIMM/SAM	21.5(20 - 21.2)
	SPP	25(21.2 – 25)
BMQ – Specific Concern	PIMM/SAM	14.7(5.2)
	SPP	18.1(1.4)
MDI - DENG	PIMM/SAM	13.8(5.1)
	SPP	13.2(2.5)
MDI - DEPR	PIMM/SAM	10.0(5.4)
	SPP	7.0(0.7)
MDI - DERL	PIMM/SAM	9.7(4.2)
	SPP	9.4(3.2)

MDI - ECON	PIMM/SAM	12.7(5.6)
	SPP	11.4(7.7)
MDI – IDDIS	PIMM/SAM	5 (5/5)
	SPP	5 (5/5)
MDI - MEMD	PIMM/SAM	10.6/(3.6)
	SPP	12.0(5.0)
MDI – Total Score	PIMM/SAM	61.6(19.1)
	SPP	58.0(14.7)
MARS – Factor 1	PIMM/SAM	2.3(1.40)
	SPP	2.0(1.7)
MARS – Factor 2	PIMM/SAM	2.5(2.0 – 3.0)
	SPP	2.0(2.0 – 3.0)
MARS – Factor 3	PIMM/SAM	0.9(0.6)
	SPP	1.3(0.7)
MARS Total Score	PIMM/SAM	5.6(1.5)
	SPP	5.6(1.8)
GSE	PIMM/SAM	26.1(2.7)
	SPP	21(7.9)
HAQ-PC - Pos	PIMM/SAM	60.3(10.6)
	SPP	49.5(33.0)
HAQ-PC - Neg	PIMM/SAM	11.3(4.0)

	SPP	5.9(5.6)
SF-36 – GH	PIMM/SAM	31.7(13.1)
	SPP	49.0(27.0)
SF-36 – MH	PIMM/SAM	26.9(10.8)
	SPP	23.0(13.2)
RBANS[b] - Attention	PIMM/SAM	91.5(16.3)
	SPP	86.9(12.9)
RBANS - Delayed Memory	PIMM/SAM	98.8(12.1)
	SPP	71.3(17.0)
RBANS - Immediate Memory	PIMM/SAM	100.1(14.5)
	SPP	85.1(17.3)
RBANS - Language	PIMM/SAM	105.3(11.1)
	SPP	96.7(6.1)
RBANS - Total Score	PIMM/SAM	97.8(12.9)
	SPP	80.3(9.7)

BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; BMQ: Beliefs in Medications Questionnaire; DENG: disenganement; DEPR: depersonalization; DERL: derealization; ECON: emotional constriction; GSE: General Self-efficacy; HAQ: Helping Alliance Questionnaire; IDDIS: identity dissociation; MARS: Medication Adherence Rating Scale; MDI: Multi-scale Dissociation Inventory; MEMD: memory distburbance; NA: not applicable; Neg: negative relationship with therapist; PIMM/SAM: partnership in

medication management/self-administered medication; Pos: positive relationship with therapist; RBANS: Repeatable Battery for the Assessment of Neuropsychological Functioning; SD: standard deviation; SF-36 – GH: Short-Form 36 General Health Subscale; SF-36 – MH: Short-Form 36 Mental Health; SPP: standard prescribing practice.

[a]Mean (SD) if normally distributed; median (25^{th}/75^{th} percentile) if non-normally distributed.

[b]Administered at baseline only.

Table 2 Between-group differences in mean change scores

Instrument	Difference (95% CI)
BAI	12.7 (0.8 to 24.5)
BDI	2.8 (-7.9 to 13.4)
BMQ – General	4.9 (0.8 to 9.0)
MDI – DENG	-1.3 (-5.6 to 2.9)
MDI – DEPR	3.7 (0.8 to 6.7)
MDI – DERL	1.8 (-0.9 to 4.4)
MDI – ECON	0.7 (-5.4 to 6.8)
MDI – MEMD	-1.1 (-7.2 to 4.9)
MDI – IDDIS	-0.6

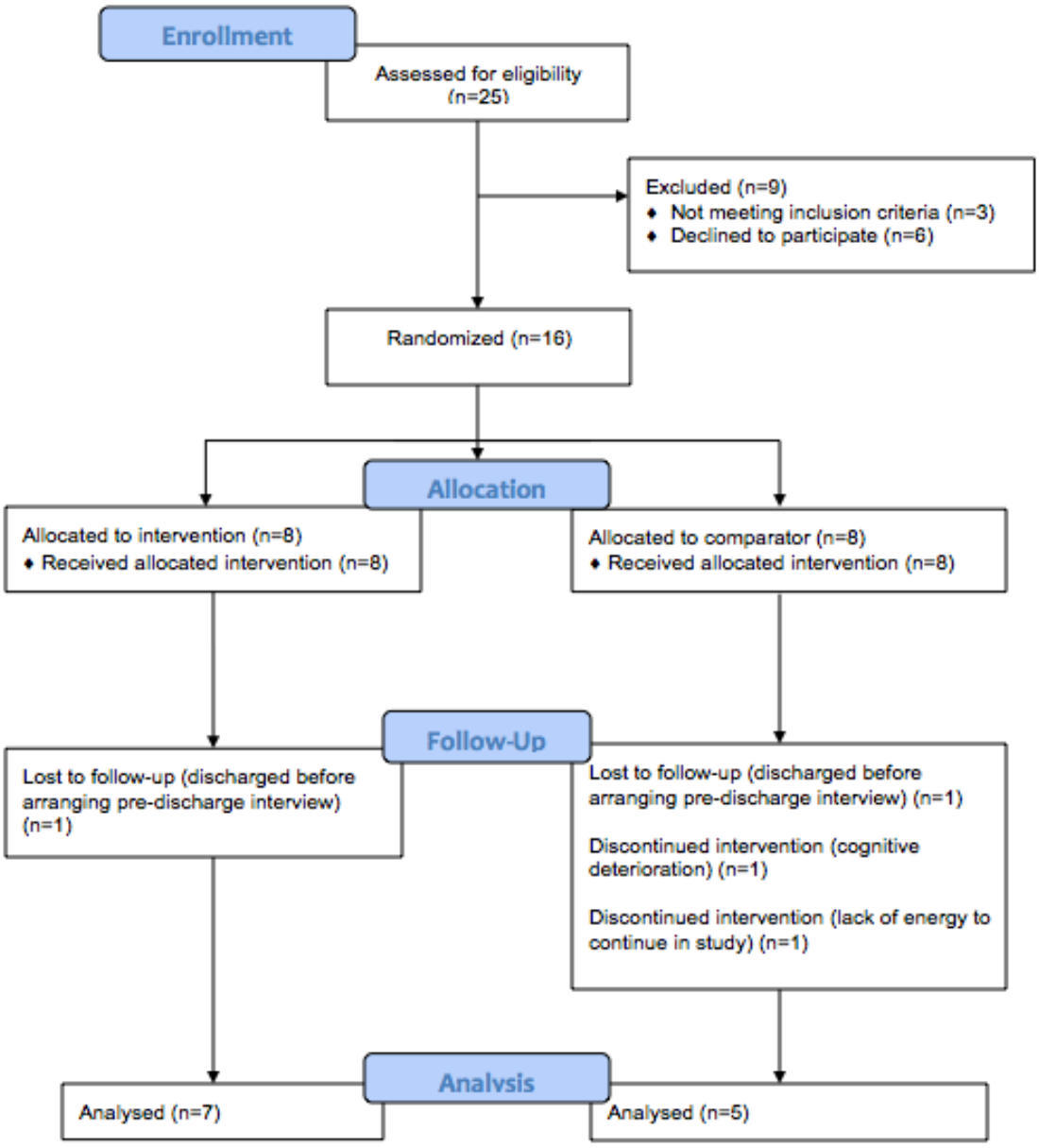
	(-1.6 to 0.4)
MDI – Total Score	3.2
	(-15.8 to 22.1)
MARS – Factor 1	-0.3
	(-1.9 to 1.4)
MARS – Factor 2	0.0
	(0.0 to 0.0)
MARS – Factor 3	-1.1
	(-0.07 to 2,2)
GSE	-1.8
	(-7.3 to 3.7)
HAQ – Positive Relationship with Therapist	-13.8
	(-28.6 to 1.0)
HAQ – Negative Relationship with Therapist	6.8
	(-0.01 to 13.6)
SF-36 – GH	-22.5
	(-44.4 to 0.06)
SF-36 – MH	-2.4
	(-20.6 to 15.8)

Note. For each instrument, mean change scores were calculated by: (1) subtracting every participant's baseline score from her/his pre-discharge score to obtain a change score; (2)

obtaining a mean change score for each study group; and (3) subtracting the mean change score of the SPP group from the mean change score of the PIMM/SAM group.

BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; BMQ: Beliefs in Medications Questionnaire; CI: confidence interval; DENG: disenganement; DEPR: depersonalization; DERL: derealization; ECON: emotional constriction; GSE: General Self-efficacy; HAQ: Helping Alliance Questionnaire; IDDIS: identity dissociation; MARS: Medication Adherence Rating Scale; MDI: Multi-scale Dissociation Inventory; MEMD: memory distburbance; PIMM/SAM: partnership in medication management/self-administered medication; RBANS: Repeatable Battery for the Assessment of Neuropsychological Functioning; SD: standard deviation; SF-36 – GH: Short-Form 36 General Health Subscale; SF-36 – MH: Short-Form 36 Mental Health; SPP: standard prescribing practice.

Figure 1: Study Recruitment Flowchart



Additional file: CONSORT Checklist

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	NA (feasibility study)
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	3-6
	2b	Specific objectives or hypotheses	5-6
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	7
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	NA
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	8
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	7, 9-10
	6b	Any changes to trial outcomes after the trial commenced, with reasons	NA
Sample size	7a	How sample size was determined	NA (feasibility study)
	7b	When applicable, explanation of any interim analyses and stopping guidelines	NA
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	8
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	8
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to	7-8

		interventions		
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	Unblinded	
	11b	If relevant, description of the similarity of interventions	NA	
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	11	
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	NA	
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Figure 1	
	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1	
Recruitment	14a	Dates defining the periods of recruitment and follow-up	7	
	14b	Why the trial ended or was stopped	NA	
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 1	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Table 2	
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Table 2	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	NA	
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	NA	
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	NA	
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	14	
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	NA (feasibility study)	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	13, 15	
Other information				
Registration	23	Registration number and name of trial registry	Title page, p. 6	
Protocol	24	Where the full trial protocol can be accessed, if available	See # 23 above	
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	16	

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

References

Ahern, C., McKinnon, M. C., Bieling, P. J., McNeely, H., & Langstaff, K. (2016). Overcoming the Challenges Inherent in Conducting Design Research in Mental Health Settings Lessons from St. Joseph's Healthcare, Hamilton's Pre and Post-Occupancy Evaluation. *HERD: Health Environments Research & Design Journal*, *9*(2), 119–129. http://doi.org/10.1177/1937586715602219

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5* (5th ed.). Arlington, VA: American Psychiatric Association.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588–597. http://doi.org/10.1207/s15327752jpa6703_13

Briere, J. (2002). *Multiscale Dissociation Inventory*. Odessa, Florida: Psychological Assessment Services.

Chong, W. W., Aslani, P., & Chen, T. F. (2011). Effectiveness of interventions to improve antidepressant medication adherence: a systematic review. *International Journal of Clinical Practice*, *65*(9), 954–975. http://doi.org/10.1111/j.1742-1241.2011.02746.x

Dunkin, J. J., Leuchter, A. F., Cook, I. A., Kasl-Godley, J. E., Abrams, M., & Rosenberg-Thompson, S. (2000). Executive dysfunction predicts nonresponse to fluoxetine in major depression. *Journal of Affective Disorders*, *60*(1), 13–23.

Gold, J. M., Queern, C., Iannone, V. N., & Buchanan, R. W. (1999). Repeatable battery for the assessment of neuropsychological status as a screening test in schizophrenia I: sensitivity, reliability, and validity. *The American Journal of Psychiatry*, *156*(12), 1944–1950. http://doi.org/10.1176/ajp.156.12.1944

Hariri, A. G., Gulec, M. Y., Orengul, F. F. C., Sumbul, E. A., Elbay, R. Y., & Gulec, H. (2015). Dissociation in bipolar disorder: Relationships between clinical variables and childhood trauma. *Journal of Affective Disorders*, *184*, 104–110. http://doi.org/10.1016/j.jad.2015.05.023

Health Canada. (2002). *A Report on Mental Illnesses in Canada*. Ottawa: Health Canada.

Horne, R., Chapman, S. C. E., Parham, R., Freemantle, N., Forbes, A., & Cooper, V. (2013). Understanding Patients' Adherence-Related Beliefs about Medicines Prescribed for Long-Term Conditions: A Meta-Analytic Review of the Necessity-Concerns Framework. *PLoS ONE*, *8*(12). http://doi.org/10.1371/journal.pone.0080633

Horne, R., Weinman, J., & Hankins, M. (1999). The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology & Health*, *14*(1), 1–24. http://doi.org/10.1080/08870449908407311

Lecrubier, Y., Sheehan, D., Weiller, E., Amorim, P., Bonora, I., Harnett Sheehan, K., … Dunbar, G. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry*, *12*(5), 224–231. http://doi.org/10.1016/S0924-9338(97)83296-8

Luborsky, L., Barber, J. P., Siqueland, L., Johnson, S., Najavits, L. M., Frank, A., & Daley, D. (1996). The Revised Helping Alliance Questionnaire (HAq-II) : Psychometric Properties. *The Journal of Psychotherapy Practice and Research*, *5*(3), 260–271.

MacQueen, G. M., Campbell, S., McEwen, B. S., Macdonald, K., Amano, S., Joffe, R. T., … Young, L. T. (2003). Course of illness, hippocampal function, and hippocampal volume in major depression. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(3), 1387–1392. http://doi.org/10.1073/pnas.0337481100

Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M. A., Petukhova, M., & Kessler, R. C. (2007). Lifetime and 12-Month Prevalence of Bipolar Spectrum Disorder in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *64*(5), 543–552. http://doi.org/10.1001/archpsyc.64.5.543

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., … Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated

guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869.

Mohr, E., Walker, D., Randolph, C., Sampson, M., & Mendis, T. (1996). Utility of clinical trial batteries in the measurement of Alzheimer's and Huntington's dementia. *International Psychogeriatrics / IPA*, *8*(3), 397–411.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. http://doi.org/10.1111/j.1532-5415.2005.53221.x

Nemeroff, C. B., Heim, C. M., Thase, M. E., Klein, D. N., Rush, A. J., Schatzberg, A. F., … Keller, M. B. (2003). Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(24), 14293–14296. http://doi.org/10.1073/pnas.2336126100

Nuller, Y. L. (1982). Depersonalisation - symptoms, meaning, therapy. *Acta Psychiatrica Scandinavica*, *66*, 451–458.

Parlar, M., Frewen, P. A., Oremus, C., Lanius, R. A., & McKinnon, M. C. (2016). Dissociative symptoms are associated with reduced neuropsychological performance in patients with recurrent depression and a history of trauma

exposure. *European Journal of Psychotraumatology*, *7*. http://doi.org/10.3402/ejpt.v7.29061

Polak, A. R., Witteveen, A. B., Reitsma, J. B., & Olff, M. (2012). The role of executive function in posttraumatic stress disorder: a systematic review. *Journal of Affective Disorders*, *141*(1), 11–21. http://doi.org/10.1016/j.jad.2012.01.001

Randolph, C. (1998). *The Repeatable Battery for the Assessment of Neuropsychological Status*. San Antonio, TX: The Psychological Corporation.

Rosa, A. R., Marco, M., Fachel, J. M. G., Kapczinski, F., Stein, A. T., & Barros, H. M. T. (2007). Correlation between drug treatment adherence and lithium treatment attitudes and knowledge by bipolar patients. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *31*(1), 217–224. http://doi.org/10.1016/j.pnpbp.2006.08.007

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In *In: Weinman J, Wright S, Johnston M. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs* (pp. 35–37). Windsor, England: NFER-NELSON.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., … Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59 Suppl 20*, 22–33;quiz 34–57.

Souery, D., Papakostas, G. I., & Trivedi, M. H. (2006). Treatment-resistant depression. *The Journal of Clinical Psychiatry*, *67 Suppl 6*, 16–22.

Thompson, K., Kulkarni, J., & Sergejew, A. A. (2000). Reliability and validity of a new Medication Adherence Rating Scale (MARS) for the psychoses. *Schizophrenia Research*, *42*(3), 241–247.

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med.Care*, *30*(0025-7079 (Print)), 473–483.

World Health Organization. (2008). *Global Burden of Disease: 2004 Update*. Geneva: World Health Organization.

CHAPTER 3:

INTER-RATER AND TEST RE-TEST RELIABILITY OF QUALITY ASSESSMENTS BY NOVICE STUDENT RATERS USING THE JADAD AND NEWCASTLE-OTTOWA SCALES & EFFECTS OF ELECTROCONVULSIVE THERAPY ON COGNITIVE FUNCTIONING IN PATIENTS WITH DEPRESSION: PROTOCOL FOR A SYSTEMATIC REVIEW AND META-ANALYSIS

Foreword to Chapter 3

3.1. Reliability Paper

Systematic reviews are a form of literature search employed in EBM. Unlike basic literature searches, systematic reviews are rigorous and transparent literature searches guided by explicit research questions and formalized methodological procedures. Systematic reviews are undertaken to answer clinical research questions by searching for, obtaining, and summarizing all of the available evidence on the topics of interest. Additionally, systematic reviews rate the quality of the evidence to account for the fact that study quality affects the conclusions one can draw from the evidence. The Cochrane Collaboration (J. P. Higgins & Green, 2016) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009a) outline the specific methods for systematic reviews. Meta-analyses are a type of systematic review that involves the statistical combination of results from individual studies to obtain single summary estimate of effect across all studies (Borenstein et al., 2009).

Systematic reviews inform clinical practice and health policy by providing physicians and policy makers with summaries of the current state of knowledge in a particular treatment area. An important component of the practice and policy component of systematic reviews is the GRADE process (GRADE Working Group, 2004). GRADE stands for Grading of Recommendations Assessment, Development and Evaluation and is a formalized means of assessing whether current levels of evidence represent true effects,

or whether new evidence in the future will change the conclusions outlined in a systematic review.

The usefulness of systematic reviews in healthcare decision-making has led to an explosion of work to refine the methods of this type of study. Indeed, checklists now exist to rate the methodological quality of systematic reviews (Shea et al., 2007). Researchers who conduct systematic reviews also undertake methodological work of their own. For example, a major component of systematic reviews is the inclusion of relevant primary research articles. Both the Cochrane and PRISMA guidelines (J. P. Higgins & Green, 2016; Moher et al., 2009a) recommend at least two raters screen the relevance of each citation retrieved in a literature search to ensure relevance. The guidelines also recommend the two-rater system for quality assessment. Methods research suggests rater training (Hartling et al., 2013) is important to ensure the validity of screening and quality assessment.

We undertook a systematic review and meta-analysis of the literature to investigate the effect of electroconvulsive therapy (ECT) on cognitive functioning in depression. In preparation for this review, we examined the scope of the literature and anticipated including a large number of studies. To manage the volume, we invited a pool of student raters to assess study quality. Given the importance of rater training to ensure the validity of systematic review results, we examined the inter-rater and test–retest reliability of student raters with no previous experience assessing study quality (M. Oremus et al.,

2012). This examination helped us identify rater strengths and weaknesses, which informed our approach to developing standardized training for our pool of raters.

3.2. Protocol

ECT involves the transmission of electric current through the brain to trigger a brief seizure, change brain chemistry, and reduce symptoms of depression. Since the initial years of ECT treatment, researchers have studied ways to improve the seizure induction and reduce adverse effects. Today, many different ECT modalities exist, including pulse shape (shift from sine wave to rectangular, pulse width [brief to ultrabrief] and electrode placement (bilateral [BL], unilateral [UL]). Also, ECT is delivered under general anaesthesia and patients are provided with muscle relaxants to prevent bone fractures (Dougherty & Rauch, 2007).

ECT has evoked diametrically opposed reactions from different quarters of the clinical community, with some clinicians concerned about potential adverse effects, while others believe it is efficacious and safe (The UK ECT Review Group, 2003). In current clinical practice, ECT prompts a generalized seizure via the delivery of an electrical current to the brain, with leads positioned to the scalp and skull unilaterally or bilaterally. The traditional placements of the leads were bitemporal (or bifrontotemporal or simply bilateral) and right unilateral; more recently, clinicians have been using bifrontal placement. The positioning of the leads affects treatment efficacy and the possibility of experiencing cognitive adverse effects (Kellner, Tobias, & Wiegand, 2010). Common adverse effects from ECT include headache, muscle ache, and nausea. Some ECT

recipients experience acute confusion for 30 to 60 minutes post-ECT, largely due to the combined effects of the ECT itself and the anaesthesia (Department of Psychiatry, Available at: http://www.psych.med.umich.edu/ect/common-side-effects.asp. Accessed on July19).

Meta-analyses have reported benefits for ECT in the treatment of TRD. A pivotal meta-analysis from the United Kingdom found that real ECT was more effective than simulated ECT (six trials, 256 patients) or drug therapy, with bilateral ECT being more effective than unipolar ECT (22 trials, 1408 participants) (The UK ECT Review Group, 2003). In another meta-analysis, brief right unilateral ECT was statistically significantly more efficacious for depression than ultrabrief right unilateral ECT (standardized mean difference: 0.25; 95% CI: 0.08 to 0.41), although ultrabrief had a lower remission rate than brief (odds ratio: 0.71; 95% CI: 0.51 to 0.99) (Tor et al., 2015).

Meta-analyses have also found that persons who undergo ECT may experience cognitive deficits as adverse effects of ECT. The evidence of cognitive deficits from these meta-analyses is limited, though, because the number of cognitive domains examined was small (i.e., four domains only [retrograde memory, anterograde memory—learning, anterograde memory—delayed recall, global cognitive function]) (Tor et al., 2015), or studies with substantial clinical heterogeneity were combined under single cognitive domains (Semkovska & McLoughlin, 2010). Alternative explanations suggest that some reports of memory loss following ECT may be manifestations of somatoform disorders (Fink, 2007).

In one of these meta-analyses (Tor et al., 2015), the authors set-out to examine the comparative efficacy of brief pulse versus ultrabrief pulse right unilateral ECT. Cognitive function was a secondary outcome of the review and the authors grouped the neuropsychological instruments used in the included studies into the four domains mentioned in the previous paragraph. The initial literature search yielded 644 references after removal of duplicates and seven studies (5 RCTs, 2 observational) were included in the review. The summary standard mean differences (SMDs) for cognitive function all favoured ultrabrief pulse ECT: retrograde memory (SMD: 0.38; 95% CI: 0.15 to 0.61 [5 studies]); anterograde memory—learning (SMD: 0.45; 95% CI: 0.22 to 0.68 [2 studies]); anterograde memory—delayed recall (SMD: 0.56; 95% CI: 0.40 to 0.73 [3 studies]); global cognitive function (SMD: 0.36; 95% CI: 0.09 to 0.63 [2 studies]). The clustering of these findings in the domain of memory is consistent with the preponderance of individual studies that have revealed changes in memory following ECT. Here, a total of 7 studies reveal worse memory performance following the administration of brief versus ultrabrief ECT. Our own review (see below) highlights the 14 studies that revealed worse memory performance following bilateral versus unilateral ECT. Finally, 3 pre- and post-ECT studies reveal worse memory performance following administration of ECT.

In the other meta-analysis (Semkovska & McLoughlin, 2010), the authors' primary research question was to examine the evidence for cognitive impairment following ECT. The authors identified 1,525 articles after removing duplicates and included 84 studies in the review. The articles contained 22 standardized neuropsychological tests spread over

eight cognitive domains (global cognitive status, processing speed, attention/working memory, verbal episodic memory, visual episodic memory, spatial problem solving, executive functioning, intellectual ability). The authors created 24 strata based on the type of test and grouped the studies into one or more of these strata. The strata were further sub-divided into three time periods based on the interval between the final ECT session and the last administration of a cognitive test (0 – 3 days, 4 – 15 days, > 15 days). Statistically significant decreases in cognitive performance were observed in the 0 – 3 day period in 72% of the variables, with effect sizes (Cohen's *d*) ranging from -1.10 (95% CI: -1.53 to -0.67) to -0.21 (95% CI: -0.40 to 0.01). In the 4 to 15 period, only one result suggested the presence of cognitive impairment (verbal paired associates delayed recall: Cohen's *d*: -0.36; 95% CI: -0.62 to -0.10 [4 studies]). Beyond 15 days, no results indicated the presence of cognitive impairment.

Cognitive deficits are manifested in persons with MDD. Cognition-related symptoms like loss of concentration, or difficulty in decision making are hallmarks of MDD. Neuropsychological tests have shown people with MDD, compared to non-depressed controls, perform more poorly on the Trail Making Test, the Symbol Digit Modalities Test, and the Stroop test (Hasselbalch et al., 2012). Research has shown a relation between memory deficits and early depressive symptoms. Deficits in verbal memory have been associated with the development of depressive symptoms (Iorfino et al., 2016). Additionally, researchers have reported a relation between a decline in executive function and more persistent depressive symptoms (Iorfino et al., 2016). Dysfunctional prefrontal-

subcortical circuitry and associated challenges in emotion control are believed to explain cognitive deficits in persons with MDD (Murrough et al., 2011).

A meta-analysis by Rock et al., 2014 searched PubMed and Google Scholar between 1980 to December 2012 and included 24 studies comparing currently depressed patients to healthy controls and six studies comparing remitted depressed patients to healthy controls. The authors reported moderate deficits in memory, executive function and attention (Cohen's d effect sizes ranging from -0.34 to -0.65) in currently depressed vs. healthy controls; and memory deficits (Cohen's d ranging from 22 to 0.54) in remitted depressed patients. However, the I^2 for sixteen out of twenty four studies pool results range from 56 to 82 indicating substantial statistical heterogeneity between studies.

Another meta-analysis by Lee et al., 2011 searched PubMed and PsychInfo databases from 1990 to February 2011 summarizing 13 studies. The authors reported that patients in their first episode of depression performed worse than healthy controls in attention (SMD: 0.36, 95% CI: 0.13-0.59; I^2=0%), working memory (SMD: 0.16, 95% CI: -0.20 - 0.51; I^2=61%), verbal learning and memory (SMD: 0.13, 95% CI: 0.18-0.45; I^2=81%), visual learning and memory (SMD: 0.53, 95% CI: -0.05 – 1.11; I^2=88%). Unfortunately, the heterogeneity of the pooled results for working memory, verbal and learning memory and visual and learning memory heterogeneity range from 61% to 88% which warns caution in interpreting the findings.

In addition, a meta-analysis from Wagner et al., 2012 included 15 studies investigating severity of executive dysfunctions in persons with MDD in comparison to

healthy controls and 3 before and after antidepressant treatment studies. The authors found that healthy controls had better cognitive functioning than persons with MDD in semantic and phonemic memory 0.92 SD and 0.71 SD, I^2=52%; Stroop interference test 1.18 SD, Trail Making Test B 1.109 SD. These previous meta-analyses did not strictly adhere to PRISMA guidelines, nor did they evaluate the the risk of bias and did not measure the strength of evidence. Interpretation of these findings of should be done with caution due to moderate to high levels of heterogeneity (> 50%).

We registered the protocol for our systematic review and meta-analysis in the PROSPERO database of systematic reviews (Booth et al., 2012) (Protocol #: CRD42014009100). We also published the protocol in *BMJ Open* (C. Oremus et al., 2015). The purpose of registering and publishing the protocol was to permit a peer review of the methods and to identify any duplication of effort among research teams around the world. Methodologists encourage registering and publishing systematic review protocols to reduce selective reporting of outcomes (i.e., only reporting the subset of outcomes with positive results) and publication bias (i.e., failing to publish an entire review because the results suggest no differences between the exposures and outcomes) (Straus & Moher, 2010).

The manuscripts presented in this chapter have been cited 39 times (Google Scholar). The Reliability paper was cited in a new epidemiology textbook for Canadian students (Patton, 2015).

Citations:

Oremus, C., Oremus, M., McNeely, H., Losier, B., Parlar, M., King, M., … McKinnon, M. (2015). Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis. *BMJ Open*, *5*(3), e006966. http://doi.org/10.1136/bmjopen-2014-006966

Oremus, M., Oremus, C., Hall, G. B. C., & McKinnon, M. C. (2012). Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales. *BMJ Open*, *2*(4). http://doi.org/10.1136/bmjopen-2012-001368

Open Access Research

BMJ open accessible medical research

Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales

Mark Oremus,[1,2] Carolina Oremus,[3,4] Geoffrey B C Hall,[3,4] Margaret C McKinnon,[3,4] ECT & Cognition Systematic Review Team[3,4,*]

To cite: Oremus M, Oremus C, Hall GBC, *et al*. Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales. *BMJ Open* 2012;**2**:e001368. doi:10.1136/bmjopen-2012-001368

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-001368).

*The ECT & Cognition Systematic Review Team includes Allyson Graham, Caitlin Gregory, Gagan Fervaha, Lindsay Hanford, Anthony Nazarov, Melissa Parlar, Maria Restivo, Erica Tatham and Wanda Truong.

Received 23 April 2012
Accepted 29 June 2012

For numbered affiliations see end of article.

Correspondence to
Dr Mark Oremus;
oremusm@mcmaster.ca

ABSTRACT

Introduction: Quality assessment of included studies is an important component of systematic reviews.

Objective: The authors investigated inter-rater and test–retest reliability for quality assessments conducted by inexperienced student raters.

Design: Student raters received a training session on quality assessment using the Jadad Scale for randomised controlled trials and the Newcastle–Ottawa Scale (NOS) for observational studies. Raters were randomly assigned into five pairs and they each independently rated the quality of 13–20 articles. These articles were drawn from a pool of 78 papers examining cognitive impairment following electroconvulsive therapy to treat major depressive disorder. The articles were randomly distributed to the raters. Two months later, each rater re-assessed the quality of half of their assigned articles.

Setting: McMaster Integrative Neuroscience Discovery and Study Program.

Participants: 10 students taking McMaster Integrative Neuroscience Discovery and Study Program courses.

Main outcome measures: The authors measured inter-rater reliability using κ and the intraclass correlation coefficient type 2,1 or ICC(2,1). The authors measured test–retest reliability using ICC (2,1).

Results: Inter-rater reliability varied by scale question. For the six-item Jadad Scale, question-specific κs ranged from 0.13 (95% CI −0.11 to 0.37) to 0.56 (95% CI 0.29 to 0.83). The ranges were −0.14 (95% CI −0.28 to 0.00) to 0.39 (95% CI −0.02 to 0.81) for the NOS cohort and −0.20 (95% CI −0.49 to 0.09) to 1.00 (95% CI 1.00 to 1.00) for the NOS case–control. For overall scores on the six-item Jadad Scale, ICC(2,1)s for inter-rater and test–retest reliability (accounting for systematic differences between raters) were 0.32 (95% CI 0.08 to 0.52) and 0.55 (95% CI 0.41 to 0.67), respectively. Corresponding ICC(2,1)s for the NOS cohort were −0.19 (95% CI −0.67 to 0.35) and 0.62 (95% CI 0.25 to 0.83), and for the NOS case–control, the ICC(2,1)s were 0.46 (95% CI −0.13 to 0.92) and 0.83 (95% CI 0.48 to 0.95).

Conclusions: Inter-rater reliability was generally poor to fair and test–retest reliability was fair to excellent. A pilot rating phase following rater training may be one way to improve agreement.

ARTICLE SUMMARY

Article focus

- To examine the inter-rater and test–retest reliability of inexperienced raters' quality assessments of articles included in a systematic review.

Key messages

- Among inexperienced raters, inter-rater reliability using the Jadad Scale and Newcastle–Ottawa Scale was generally poor to fair; test–retest reliability was fair to excellent.
- Systematic reviewers must pay special attention to training inexperienced quality raters; a pilot rating phase might be a helpful means of improving reliability among inexperienced raters, especially when rating observational study quality.

Strengths and limitations of this study

- No other study has examined the reliability of quality assessments in a group of inexperienced raters.
- Results may differ depending on rater background and experience, rater training, quality assessment instruments and topic under study.

INTRODUCTION

Systematic reviews summarise healthcare research evidence, and they are useful for assessing whether treatment benefits outweigh risks.[1 2] Accordingly, conclusions drawn from systematic reviews may impact clinical care and patient outcomes, thereby necessitating high standards of methodological rigour.

One critical component of conducting systematic reviews involves evaluation of the

methodological quality of included studies. Study quality may influence treatment effect estimates and the validity of conclusions drawn from such estimates.[3] Through quality assessment, researchers identify strengths and weaknesses of existing evidence[4] and suggest ways to improve future research.

Careful work has identified key quality assessment domains.[1 5] For randomised controlled trials (RCTs), these domains include appropriate generation of random allocation sequences, concealment of allocation sequences, blinding (of participants, healthcare providers, data collectors and outcome assessors) and reporting of proportions of patients lost to follow-up.[1] For observational studies, key domains include the adequacy of case definition, exposure ascertainment and outcome assessment,[5] as well as selection and attrition biases.

Numerous scales exist to help raters assess study quality.[5–11] The majority of these scales list quality assessment domains and require raters to indicate whether each domain is present or absent from the studies under consideration. Some scales (eg, Jadad,[6] Newcastle–Ottawa Scale (NOS)[5]) assign points when quality domains are present, thus permitting the calculation of overall 'quality scores'. Other scales (eg, risk of bias[8]) ask raters to rank the degree of bias (high, low, unclear) associated with each quality domain.

Generally, quality scales demonstrate good inter-rater and test–retest reliability. Reliability coefficients such as κ are typically >0.60,[9–17] although recent work reports κs of <0.50 for eight of the nine questions on the NOS.[18]

Although quality assessment is now regarded as a standard component of systematic reviews, one issue that has received little attention in the literature is the effect of rater experience on the reliability of quality assessments. This issue is important because raters may be drawn from vast pools of persons with varying degrees of methods expertise, from experienced faculty to inexperienced students.

We investigated inter-rater and test–retest reliability for student raters with no previous experience in the quality assessment of RCTs and observational studies. To the best of our knowledge, no other study has examined this topic.

METHODS

Study design

In an ongoing systematic review of cognitive impairment following electroconvulsive therapy (ECT) to treat major depressive disorder, 78 published articles passed title and abstract and full-text screening. These articles formed the basis of this study. Fifty-five of the articles reported the results of RCTs, with one article containing results of five separate studies and two other articles each containing results of two separate studies, for a total of 61 RCTs. Fifteen articles reported on cohort studies and eight reported on case–control studies. Eleven articles were published prior to 1980, 17 between 1980 and 1989, 15 between 1990 and 1999, and 35 since 2000.

We invited all 10 students (three undergraduate and seven graduate) taking a 'special topics' course in the McMaster Integrative Neuroscience Discovery and Study Program to participate in this study. All 10 students accepted the invitation. One author (MO) with systematic review experience trained the students to rate the methodological quality of published study reports using the six-item Jadad Scale for RCTs[6 19] and the NOS for observational studies.[5] Training consisted of a 90 min didactic session divided into two parts: part one highlighted the importance of quality assessment in systematic reviews and part two contained a question-by-question description of the Jadad and NOS instruments. We provided a standardised tabular spreadsheet for student raters to use during quality assessment.

We used a random number table to assign the student raters into five pairs and we randomly distributed between 13 and 20 articles to each pair. None of the 78 articles was assigned to more than one pair; pairs received a mix of RCTs and observational studies. The number of articles assigned to the pairs depended on the amount of time each rater could devote to this study.

Raters determined the type of study design (ie, RCT or observational) for each of their assigned articles and one author (CO) verified their choices. Raters then independently rated their assigned articles to permit us to examine inter-rater reliability.

Statistical analysis

We used κ (kappa)[20 21] to measure inter-rater reliability for individual Jadad and NOS questions. We interpreted κ values as follows: >0.80 was very good, 0.61–0.80 was good, 0.41–0.60 was moderate, 0.21–0.40 was fair and <0.21 was poor.[22]

For test–retest reliability, each rater re-assessed half of the articles to which they had been assigned during the inter-rater reliability phase. The re-assessments took place 2 months after the inter-rater reliability phase[13] to minimise the possibility that recall of the first assessments would influence the second assessments.

We employed the intraclass correlation coefficient-model 2,1 or ICC(2,1)[23] to measure inter-rater and test–retest reliability for the Jadad and NOS total scores. We computed separate ICC(2,1) values for consistency (systematic differences between raters are considered irrelevant) and absolute agreement (systematic differences between raters are considered relevant).[24] ICC (2,1) values were interpreted as follows: >0.75 was excellent, 0.40–0.75 was fair to good and <0.40 was poor.[25]

We calculated two sets of ICC(2,1)s for the Jadad Scale. The first set pertained to the six-item Jadad Scale,[19] and the second set pertained to the original three-item Jadad Scale.[6]

SAS V.9.2 (The SAS Institute) was used to calculate κ; SPSS V.20 (IBM Corp.) was used to calculate ICC(2,1). The level of significance was α=0.05.

RESULTS

Inter-rater reliability

For inter-rater reliability, agreement between raters on individual questions was generally poor (table 1). Half of the questions on the Jadad Scale had moderate κs and the other half had poor κs. On the NOS, all κs were poor for the cohort study questions (NOS cohort) and six of the eight κs were poor for the case–control study questions (NOS case–control).

Examining total scale scores within rater pairs (table 2), agreement was poor for the Jadad Scale (six- and three-item versions) and NOS cohort and fair for the NOS case–control. However, point estimate ICC(2,1)s for the NOS cohort and case–control were not statistically significantly different from zero. Point estimate ICC (2,1)s and 95% CIs did not appreciably differ according to calculation based on consistency or absolute agreement.

Test–retest reliability

Test–retest reliability following a 2-month interval between assessments was fair to good for the Jadad Scale and NOS cohort and excellent for the NOS case–control (table 3). Test–retest reliability was slightly higher for the three-item Jadad Scale versus the six-item Jadad Scale. Point estimate ICC(2,1)s and 95% CIs calculated for consistency were similar to the results calculated for absolute agreement.

DISCUSSION

Overview and discussion of key findings

We investigated inter-rater and test–retest reliability for student raters with no previous experience in quality assessment. Our study is novel because, to the best of our knowledge, no other research has examined this issue. The raters used the Jadad Scale and NOS to assess the quality of studies on the topic of ECT and cognitive impairment. Inter-rater reliability was generally poor to fair and test–retest reliability was fair to excellent. Our results highlight the need for researchers to consider rater experience during the quality assessment of articles included in systematic reviews.

For inter-rater reliability, the poor κs on the Jadad Scale pertained to the questions about appropriateness of double blinding and the clarity of reporting withdrawals, inclusion/exclusion criteria and adverse effects. Often, authors did not report methods of blinding and raters had to make judgements about whether to award a point for the question on appropriateness of double blinding. Despite what we communicated during the training session, some raters may have given authors the benefit of the doubt and awarded the point for appropriateness if studies simply reported double blinding, even though another question on the Jadad Scale already asked whether authors reported their studies as blinded. Similarly, differences in rater opinion regarding what constitutes an 'adequate' description of withdrawals, inclusion/exclusion criteria or adverse effects led to poor agreement on these questions. To improve inter-rater agreement among inexperienced raters, we suggest a pilot phase wherein raters rate the quality of a subsample of articles to allow for the identification and clarification of areas of ambiguity.

We recognise that any strategy to improve reliability will be limited by instrument content and structure. Scales with larger numbers of interpretive questions will likely have lower reliability than scales with fewer interpretive questions, regardless of the efforts made to improve reliability.

With regard to the NOS, question-specific inter-rater reliability was poorer than that of the Jadad Scale. We believe that the NOS's poor reliability may be explained

Table 1 Inter-rater reliability for Jadad Scale and Newcastle–Ottawa Scale (NOS): by question

Question—Jadad Scale	κ (95% CI)	Question—NOS cohort	κ (95% CI)	Question—NOS case–control	κ (95% CI)
Randomisation	0.50 (−1.00 to 1.00)	Representativeness of exposed cohort	−0.13 (−0.36 to 0.11)	Case definition adequate	1.00 (1.00 to 1.00)
Appropriate randomisation	0.56 (0.29 to 0.83)	Selection of non-exposed cohort	−0.14 (−0.28 to 0.00)	Cases representative	−0.20 (−0.49 to 0.09)
Double blind	0.41 (0.16 to 0.66)	Exposure ascertainment	0.00 (0.00 to 0.00)	Control selection	0.25 (−0.19 to 0.69)
Appropriate double blind	0.17 (−0.07 to 0.41)	Outcome not present at baseline	0.20 (−0.33 to 0.73)	Control definition	0.14 (−0.54 to 0.82)
Description of withdrawals	0.21 (−0.02 to 0.45)	Comparability of cohorts	0.12 (−0.23 to 0.47)	Case and control comparability	0.00 (0.00 to 0.00)
Description of inclusion/ exclusion criteria	0.27 (−0.03 to 0.57)	Outcome assessment	0.31 (−0.08 to 0.69)	Exposure ascertainment	−0.11 (−0.68 to 0.46)
Description of adverse effects	0.13 (−0.11 to 0.37)	Follow-up long enough	−0.09 (−0.22 to 0.04)	Same ascertainment method for cases and controls	0.60 (−0.07 to 1.00)
Description of statistical analysis	0.49 (0.21 to 0.77)	Follow-up adequate	0.39 (−0.02 to 0.81)	Non-response rate	−0.11 (−0.65 to 0.43)

κ, Kappa.

Reliability of article quality assessment

Table 2 Inter-rater reliability for Jadad and Newcastle–Ottawa Scales: total scale scores within rater pairs

Scale	ICC(2,1) (95% CI), consistency*	ICC(2,1) (95% CI), absolute agreement†
Jadad—six item	0.32 (0.08 to 0.53)	0.32 (0.08 to 0.52)
Jadad—three item	0.35 (0.11 to 0.56)	0.35 (0.11 to 0.56)
Newcastle–Ottawa—cohort	−0.19 (−0.63 to 0.34)	−0.19 (−0.67 to 0.35)
Newcastle–Ottawa—case–control	0.55 (−0.18 to 0.89)	0.46 (−0.13 to 0.92)

*ICC(2,1) where systematic differences between raters are irrelevant.
†ICC(2,1) where systematic differences between raters are relevant.
ICC, intraclass correlation coefficient.

in part by differences in how raters answered interpretive questions, for example, whether exposed cohorts are somewhat or truly representative of the average exposed person in the community (first question on NOS cohort).

Poor question-specific inter-rater agreement on the NOS also reflects an inherent challenge with rating the quality of observational studies compared with RCTs. This challenge is exemplified by the multiplicity of tools that exist to assess observational study quality. Two systematic reviews[26 27] each found over 80 such tools, which varied in design and content. Despite the cornucopia of tools, no gold standard scale exists to rate the quality of observational studies.[28]

Rater disagreements on interpretive questions and inherent challenges with assessing observational study quality explain the negative κs that were calculated for some NOS questions. Negative κs result when agreement occurs less often than predicted by chance alone. This suggests genuine disagreement between raters or an underlying issue with the instrument itself.[29] Indeed, Hartling *et al*[18] reported that raters had difficulty using the NOS because of uncertainty over the meaning of certain questions (eg, representativeness of the exposed cohort, selection of non-exposed cohort) and response options (eg, 'truly' vs 'somewhat' exposed). These difficulties existed despite Hartling *et al*'s use of a pilot training phase. Our raters' difficulties with the interpretative questions might have been a function of issues with the NOS, which could be related to the broader challenge of assessing the quality of observational studies.

Question-specific differences between raters also led to poor inter-rater agreement on total scores for the Jadad Scale and NOS cohort. This may not be evident by comparing the κs and ICC(2,1)s calculated for the Jadad. κs for four of the eight Jadad questions were moderate yet the ICC(2,1) for total score was poor. However, since total scores are computed using raters' answers to all of the questions on a scale (some answers are awarded one point and others zero points), raters who disagree on small numbers of questions (eg, two of the eight questions) will nonetheless show poor agreement on total scores.

Conversely, for the NOS case–control, κs for six of the eight questions were poor yet the ICC(2,1) was fair. In this situation, no 'reliability' relation exists between responses to questions and total scores. For example, rater 1 might answer 'yes' (one point per 'yes' response) and rater 2 might answer 'no' (zero points per 'no' response) to even-numbered questions. For odd-numbered questions, the pattern is reversed. Assuming eight questions, inter-rater reliability at the question level will be poor because the raters did not agree on their responses, but their overall scores will be equivalent.

Many authors base their discussions of study quality in systematic reviews on raters' responses to individual questions on quality assessment scales. Given that we found generally poor inter-rater reliability on answers to questions, the process of resolving conflicts between raters becomes important. Many reviews simply report that raters solved disagreements by consensus without describing specific procedures. We speculate that conflict resolution may occasionally be approached in an ad hoc nature or treated as a nuisance to be dealt with as expeditiously as possible. We suggest the process of conflict resolution should be more of a formalised endeavour requiring raters to set aside some 'resolution time' and articulate their reasons for choosing specific

Table 3 Test–retest reliability for Jadad and Newcastle–Ottawa Scales: comparison of total scale scores for individual raters after two assessments

Scale	ICC(2,1) (95% CI), consistency*	ICC(2,1) (95% CI), absolute agreement†
Jadad—six item	0.56 (0.42 to 0.67)	0.55 (0.41 to 0.67)
Jadad—three item	0.67 (0.55 to 0.76)	0.67 (0.55 to 0.76)
Newcastle–Ottawa—cohort	0.61 (0.24 to 0.82)	0.62 (0.25 to 0.83)
Newcastle–Ottawa—case–control	0.85 (0.55 to 0.95)	0.83 (0.48 to 0.95)

*ICC(2,1) where systematic differences between raters are irrelevant.
†ICC(2,1) where systematic differences between raters are relevant.
ICC, intraclass correlation coefficient.

answers. In the event the raters do not agree, a third party may be asked to listen to each rater's opinion and make a decision. Although space restrictions in journals might prevent authors from reporting such procedures (when they exist) in manuscripts, the move towards publication of systematic review protocols, for example, as mandated by the United States Agency for Healthcare Research and Quality's Effective Health Care Program,[30] provides authors with an opportunity to elaborate on their consensus processes.

Test–retest reliability was better than inter-rater reliability. Individual raters appeared to adopt a uniform approach to assessing the quality of articles assigned to them. Each rater had her or his own understanding of the interpretive questions and applied this point-of-view consistently throughout the rating process. The issue was the difference in interpretations between raters.

Comparison with other studies

To the best of our knowledge, no other study has examined inter-rater and test–retest reliability for a group of novice student quality assessors. Two published studies[31] [32] of rater agreement included persons with different levels of experience, although the focus was on extraction of article data (eg, info on study design, sample characteristics, length of follow-up, definition of outcome and results) rather than quality assessment. Horton *et al*[31] classified rater experience as minimal, moderate or substantial and asked raters to extract data from three studies on insomnia therapy. They found no statistically significant differences in error rates according to experience. Haywood *et al*[32] trained two experienced raters and one inexperienced rater to independently extract data from seven studies. Agreement between raters was largely perfect.

A recent AHRQ methods report had 16 raters assess the quality of 131 cohort studies using the NOS. Rater experience ranged from 4 months to 10 years; 13 raters had formal training in systematic reviews.[18] κs were <0.50 for eight of the nine NOS questions, although the authors did not break down their results by rater experience.

Oremus *et al* examined the inter-rater reliability of the Jadad Scale using three raters (two experienced faculty members and one inexperienced PhD student), who read the methods and results of 42 Alzheimer's disease drug trials.[19] The ICC(2,1) for total scores on the Jadad Scale was 0.90. Al-Harbi *et al*[12] engaged two paediatric surgeons to rate 46 cohort studies that were presented at Canadian Association of Pediatric Surgeons annual meetings and later published in the *Journal of Pediatric Surgery*. The authors did not specify whether the surgeons received training in quality assessment. The ICC between surgeons, calculated on NOS total scores, was 0.94.

The lower inter-rater reliability of the novice student raters in this study, compared with the raters in the Oremus *et al*[19] and Al-Harbi *et al*[12] studies, may be explained by topic familiarity and similarity of expertise. The faculty raters in the Oremus *et al* study had previously worked on a systematic review of Alzheimer's disease medications and their expertise lay in two domains of epidemiology, that is, neuroepidemiology and pharmacoepidemiology. The paediatric surgeons in Al-Harbi *et al* may have possessed at least a general familiarity with the types of cohort studies conducted in their specialty. These characteristics may have predisposed the raters to adopt more uniform opinions on the questions contained in the Jadad and NOS. In contrast, the novice student raters in our study had for the most part not been exposed to systematic reviews and quality assessment in the past. Also, seven of these raters were recent entrants to graduate school, and they came from a variety of undergraduate backgrounds such as medicine, psychology and basic science.

Limitations

Readers should exercise caution when generalising the results of our study to other types of raters. Reliability could differ according to raters' disciplines and levels of training. Reliability in our study also could have been affected by the specific training programme we gave to the students. Additionally, the 10 student raters in this study were a convenience sample that might not represent all raters with similar disciplines and training.

We did not compare the students' rankings with the rankings of more experienced raters (eg, faculty who conduct systematic reviews). Thus, we could not assess the relative differences in reliability between experienced raters and inexperienced students.

Reliability is also partly a function of the instruments used in the quality assessment. Indeed, instruments with many interpretive questions (eg, appropriateness of randomisation and double-blinding, representativeness of exposed cohort or adequacy of case definition) could have poor reliability, despite several phases of training.

Furthermore, the topic under study could influence reliability, as could certain methodological decisions related to the systematic review. For example, the systematic review of ECT and cognition, upon which we based this study, included 28 papers published prior to 1990. Since the style of reporting in older papers does not always facilitate quality assessment or data extraction, systematic reviews that include older papers could present challenges for maintaining acceptable levels of inter-rater and test–retest reliability.

Conclusions

In conclusion, we asked a group of 10 novice students to rate the quality of 78 articles that contained data on cognitive impairment following the use of ECT to treat major depressive disorder. Overall, inter-rater reliability on the Jadad Scale and NOS was poor to fair and test–retest reliability was fair to excellent. We trained the raters prior to the quality assessment exercise yet inter-rater agreement was low for several questions that required a certain degree of interpretation to answer. This was especially so for the NOS and underscores an

inherent greater difficulty with assessing the quality of observational studies compared with RCTs.

In addition to standardised training prior to commencing quality assessment, a pilot rating phase may also be necessary to discuss scale questions that generate disagreement among novice student raters. This procedure could help the raters develop standardised interpretations to minimise disagreement.

While the Cochrane Collaboration has stated that quality scales and scale scores are inappropriate means of ascertaining study quality,[33] our results are relevant because many researchers continue to use the Jadad Scale and NOS in their systematic reviews. Indeed, our work suggests an area of future research. The Cochrane Collaboration has proposed a 'risk of bias' tool to assess the quality of RCTs.[33] The reliability of the risk of bias tool should be assessed in raters with different levels of experience.

Author affiliations
[1]McMaster Evidence-based Practice Centre, McMaster University, Hamilton, Ontario, Canada
[2]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada
[3]McMaster Integrative Neuroscience Discovery and Study (MINDS) Program, Hamilton, Ontario, Canada
[4]Department of Psychiatry and Behavioural Neuroscience, Hamilton, Ontario, Canada

Acknowledgements Special thanks to Eleanor Pullenayegum and Harry Shannon for their helpful comments on an earlier draft of this manuscript.

Contributors MO and CO conceived and designed the study. MO analysed the data. MO, CO, MCM, GBCH and the ECT & Cognition Systematic Review Team interpreted the data. MO drafted the manuscript. CO, MCM, GBCH and the ECT & Cognition Systematic Review Team critically revised the manuscript for important intellectual content. All authors approved the final version of the manuscript.

Funding This study did not receive funds from any sponsor. No person or organisation beyond the authors had any input in study design and the collection, analysis and interpretation of data and the writing of the article and the decision to submit it for publication.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data available.

REFERENCES

1. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
2. Agency for Healthcare Research and Quality (AHRQ). *Systems to Rate the Strength of Scientific Evidence. Evidence Report/ Technology Assessment No. 47*. Rockville, MD: Agency for Healthcare Research and Quality, 2002.
3. Verhagen AP, de Vet HC, de Bie RA, *et al*. The art of quality assessment of RCTs included in systematic reviews. *J Clin Epidemiol* 2001;54:651–4.
4. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *CMAJ* 1988;138:697–703.
5. Wells GA, Shea B, O'Connell D, *et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-analyses*. Ottawa: Ottawa Hospital Research Institute. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed 2 Apr 2012).
6. Jadad AR, Moore RA, Carroll D, *et al*. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1–12.
7. Chalmers TC, Smith H Jr, Blackburn B, *et al*. A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981;2:31–49.
8. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. [Updated March 2011]*. The Cochrane Collaboration, 2011. http://www.cochrane-handbook.org (accessed 23 Apr 2012).
9. Maher CG, Sherrington C, Herbert RD, *et al*. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys Ther* 2003;83:713–21.
10. Kocsis JH, Gerber AJ, Milrod B, *et al*. A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Compr Psychiatry* 2010;51:319–24.
11. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–84.
12. Al-Harbi K, Farrokhyar F, Mulla S, *et al*. Classification and appraisal of the level of clinical evidence of publications from the Canadian Association of Pediatric Surgeons for the past 10 years. *J Pediatr Surg* 2009;44:1013–17.
13. Berard A, Andreu N, Tetrault J, *et al*. Reliability of Chalmers' scale to assess quality in meta-analyses on pharmacological treatments for osteoporosis. *Ann Epidemiol* 2000;10:498–503.
14. Hartling L, Ospina M, Liang Y, *et al*. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ* 2009;339:b4012.
15. Hartling L, Bond K, Vandermeer B, *et al*. Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. *PLoS One* 2011;6: e17242.
16. Tooth L, Bennett S, McCluskey A, *et al*. Appraising the quality of randomized controlled trials: inter-rater reliability for the OTseeker evidence database. *J Eval Clin Pract* 2005;11:547–55.
17. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001;135:982–9.
18. Hartling L, Hamm M, Milne A, *et al. Validity and Inter-Rater Reliability Testing of Quality Assessment Instruments*. Rockville, MD: Agency for Healthcare Research and Quality, 2012.
19. Oremus M, Wolfson C, Perrault A, *et al*. Interrater reliability of the modified Jadad quality scale for systematic reviews of Alzheimer's disease drug trials. *Dement Geriatr Cogn Disord* 2001;12:232–6.
20. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
21. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. 3rd edn. Hoboken, NJ: John Wiley & Sons, 2003.
22. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991.
23. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
24. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 4th edn. Oxford: Oxford University Press, 2008.
25. Fleiss J. *The Design and Analysis of Clinical Experiments*. New York: John Wiley & Sons, 1986.
26. Shamliyan T, Kane RL, Dickinson S. A systematic review of tools used to assess the quality of observational studies that examine incidence or prevalence and risk factors for diseases. *J Clin Epidemiol* 2010;63:1061–70.
27. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666–76.
28. Lang S, Kleijnen J. Quality assessment tools for observational studies: lack of consensus. *Int J Evid Based Healthc* 2010;8:247.
29. Juurlink DN, Detsky AS. Kappa statistic. *CMAJ* 2005;173:16.
30. Agency for Healthcare Research and Quality (AHRQ). *Effective Health Care Program*. Rockville, MD: Agency for Healthcare Research and Quality. http://www.effectivehealthcare.ahrq.gov (accessed 2 Apr 2012).
31. Horton J, Vandermeer B, Hartling L, *et al*. Systematic review data extraction: cross-sectional study showed that experience did not increase accuracy. *J Clin Epidemiol* 2010;63:289–98.
32. Haywood KL, Hargreaves J, White R, *et al*. Reviewing measures of outcome: reliability of data extraction. *J Eval Clin Pract* 2004;10:329–37.
33. Higgins JPT, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.

Open Access Protocol

BMJ Open

Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis

Carolina Oremus,[1] Mark Oremus,[2] Heather McNeely,[3] Bruno Losier,[3] Melissa Parlar,[1] Matthew King,[4] Gary Hasey,[5] Gagan Fervaha,[6] Allyson C Graham,[7] Caitlin Gregory,[8] Lindsay Hanford,[9] Anthony Nazarov,[1] Maria Restivo,[1] Erica Tatham,[9] Wanda Truong,[10] Geoffrey B C Hall,[11] Ruth Lanius,[12] Margaret McKinnon[13]

To cite: Oremus C, Oremus M, McNeely H, *et al*. Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis. *BMJ Open* 2015;**5**: e006966. doi:10.1136/bmjopen-2014-006966

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2014-006966).

Received 20 October 2014
Revised 5 February 2015
Accepted 7 February 2015

For numbered affiliations see end of article.

Correspondence to
Dr Carolina Oremus;
coremus@stjoes.ca

ABSTRACT

Introduction: Depression is the leading cause of disability worldwide, affecting approximately 350 million people. Evidence indicates that only 60–70% of persons with major depressive disorder who tolerate antidepressants respond to first-line drug treatment; the remainder become treatment resistant. Electroconvulsive therapy (ECT) is considered an effective therapy in persons with treatment-resistant depression. The use of ECT is controversial due to concerns about temporary cognitive impairment in the acute post-treatment period. We will conduct a meta-analysis to examine the effects of ECT on cognition in persons with depression.

Methods: This systematic review and meta-analysis has been registered with PROSPERO (registration number: CRD42014009100). We developed our methods following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement. We are searching MEDLINE, PsychINFO, EMBASE, CINAHL and Cochrane from the date of database inception to the end of October 2014. We are also searching the reference lists of published reviews and evidence reports for additional citations. Comparative studies (randomised controlled trials, cohort and case–control) published in English will be included in the meta-analysis. Three clinical neuropsychologists will group the cognitive tests in each included article into a set of mutually exclusive cognitive subdomains. The risk of bias of randomised controlled trials will be assessed using the Jadad scale. We will supplement the Jadad scale with additional questions based on the Cochrane risk of bias tool. The risk of bias of cohort and case–control studies will be assessed using the Newcastle-Ottawa Scale. We will employ the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) to assess the strength of evidence.

Statistical analysis: Separate meta-analyses will be conducted for each ECT treatment modality and cognitive subdomain using Comprehensive Meta-Analysis V.2.0.

INTRODUCTION

According to the WHO, depression is the leading global cause of disability. Approximately, 350 million people suffer from depression worldwide.[1] Despite the availability of numerous psychopharmacological treatments, evidence indicates that only 60–70% of persons who tolerate antidepressants will respond to first-line drug therapy for major depressive disorder (MDD).[2] Furthermore, at least one-third of persons with MDD who receive drug therapy will become treatment resistant.[3] Various definitions have been proposed for treatment-resistant depression (TRD). The European Agency for the Evaluation of Medicinal Products has defined TRD as the failure to respond to two drugs of different classes, provided these drugs are used for a sufficient length of time and at an adequate dose.[4] TRD has also been defined as failing four or more different therapeutic antidepressant regimens, including augmentation, combination and electroconvulsive therapy (ECT).[5]

The aetiology of TRD is unclear. Various clinical factors have been associated with treatment non-response and resistance in MDD,[6 7] including non-adherence to treatment, poor tolerability to antidepressant medications, and medical and psychiatric comorbidity. Researchers have also identified comorbid post-traumatic stress disorder[6] and the presence of early life adversity[7] as important predictors of incomplete treatment response.[6–8]

ECT is considered an effective acute treatment for TRD[9] in either unipolar or bipolar depression.[10] ECT is used primarily when antidepressant medications do not result in

adequate response in TRD.[11] Approximately, 100 000 persons annually receive ECT in the USA.[12] However, the use of ECT remains controversial due to concerns about temporary cognitive impairment in persons with depression who receive acute ECT. Indeed, retrograde and anterograde memory deficits are among the more reliably reported cognitive changes due to ECT.[9] The UK ECT Group also found that differences in ECT treatment modalities (eg, electrode placement, pulse shape, treatment frequency and treatment dosage) had a differential impact on the incidence and duration of cognitive impairment in persons with depression.[9]

Semkovska and McLoughlin[13] examined the issue of cognitive impairment post-ECT in a recent meta-analysis. After pooling results by cognitive test, these authors found that cognitive impairment was limited to a post-treatment period of 3 days. Although Semkovska and McLoughlin[13] did assess risk of bias, these results are not reported in the manuscript nor did they report the grading of the strength of evidence.

The purpose of the present study is to conduct a systematic review and meta-analysis of the effects of ECT on cognition in persons with depression. We seek to quantify the effect of different ECT treatment modalities on the occurrence and duration of cognitive impairment. The present review includes comparative studies only (randomised controlled trials (RCTs), cohort and case–control), which are among the highest levels of evidence. Additionally, the review only includes studies where cognitive function as an outcome is reported using standardised neuropsychological tests or self-report measures that are grouped into mutually exclusive cognitive subdomains.

In contrast to Semkovska and McLoughlin,[13] results in the proposed review are grouped by cognitive subdomains, rather than cognitive tests. The focus on cognitive subdomains is a closer reflection of clinical and research practice. In these settings, multiple tests are available to assess performance within individual cognitive domains (eg, verbal recollective memory). The current literature reflects this heterogeneity, with multiple measures reported across studies to assess key cognitive domains that have become the focus of intense research interest. Inclusion of a wider corpus of measures within common cognitive domains reflects clinical and research practice. In further contrast to Semakovska and McLoughlin, we include studies that actively compare more conservative ECT treatments (eg, unilateral) to less conservative (eg, bilateral) ECT treatments. A primary outcome is post-treatment between-group differences in cognition for persons receiving less conservative versus more conservative ECT treatments. By contrast, Semkovska and McLoughlin[13] compared pretreatment and post-treatment scores on cognitive tests. Although they stratified by some components of treatment modality, the resulting comparisons were within-group differences, rather than between-group (between-treatment) comparisons. From a clinical perspective, it is crucial to determine whether the impact of cognitive impairment differs between treatments. Furthermore, by including studies that measured subjective memory in addition to objective neuropsychological measures of memory, we are able to compare and contrast potential differences in these aspects of memory functioning following treatment. Finally, we provide key data concerning the risk of bias of the included studies and rate the overall strength of evidence.

METHODS

This systematic review and meta-analysis was registered with PROSPERO (registration number: CRD42014009100; http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014009100).[14] We based the review methods on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis[15] statement.

Literature review

We are searching MEDLINE, PsychINFO, EMBASE, CINAHL and the Cochrane Central Register of Controlled Trials from database inception to the end of October 2014. The literature search mirrors the search employed by the UK ECT Review Group.[9] We consulted a medical librarian to add specific search terms to narrow our focus to the identification of articles about cognitive side effects. The final search terms included: electroconvulsive therapy; electroshock therapy; ECT; shock therapy; convulsive therapy; mood disorders; depression; schizophrenia-and-disorders-with-psychotic-features; personality disorders; delirium-dementia, -amnesic, -cognitive-disorders; bipolar disorder; randomized-controlled-trials; random*; cohort-studies; case-control-studies; double-blind-method; single-blind-method; follow-up-studies; attention; orientation, learn*; memory; concentration; cognit*; mental-process*; executive functioning; visuospatial; language; intelligence; intellectual functioning; motor function; neuropsychology. We are also searching the references of published reviews and health technology assessments related to ECT and cognition.[9 10 13 16–18]

Inclusion/exclusion criteria

We are including studies retrieved in the literature search that meet the following criteria:

1. Comparative studies (RCTs, cohort studies and case–control studies) assessing less versus more conservative ECT treatments;
2. Outcomes measured using standardised neuropsychological tests and self-report memory measures with established psychometric properties;
3. Diagnosis of participants with a major depressive episode (Diagnostic and Statistical Manual of Mental Disorders (DSM)-III, DSM-III-R, DSM-IV, DSM-IV-TR, Research Diagnostic Criteria (RDC), International Classification of Diseases (ICD)-9, ICD-10) or endogenous depression; and
4. Published in English.

Study selection and data extraction

Two reviewers are independently applying the inclusion and exclusion criteria to the citations retrieved in the literature search. This screening process is divided into two levels and it is guided by standardised instructions. For the first screening level, reviewers are independently evaluating the titles and abstracts. Citations that fulfil the inclusion criteria are advanced to the second screening level. Advancement also occurs if the reviewer does not find sufficient information to determine whether the citation fulfils the inclusion criteria. For the second screening level, the complete scientific paper is read to determine whether the inclusion criteria are met. At both levels, mutual agreement is required from the reviewers to advance a study. Discrepancies are resolved by consensus. When consensus is not attained, a third reviewer independently reviews the study in question and makes a final decision. We will use weighted κ to measure inter-rater agreement between reviewers at both levels of screening.

Studies that pass the second screening level advance to data extraction. A team of trained reviewers extracts data from the included studies. Standardised forms and training guide the data extraction process. The following information is extracted from each article: study design, mean age, proportion of men and women, diagnosis, co-morbidity, illness duration, illness severity, age of illness onset (in years), number of illness episodes, sample size, ECT description, total number of ECT sessions, comparator group characteristics, length of follow-up, treatment modality and cognitive outcomes. Examples of treatment modalities (less vs more conservative modalities) include bilateral versus unilateral ECT, three times versus twice weekly treatment, ultrabrief versus brief pulse, sine versus pulse, ECT versus pharmacological treatment, ECT versus no treatment and ECT versus sham. The first author of this protocol (CO) reviews the extracted data to verify the accuracy of the work. We are contacting the authors of included studies to obtain information that may be missing from the published reports.

Cognitive subdomains

Since cognitive outcomes in ECT studies are reported using a wide range of measurement instruments that increase the number of variables across and between studies, we grouped these instruments into cognitive subdomains to facilitate data extraction, reporting and analysis. Three experienced clinical neuropsychologists (BL, HM and MM) generated a list of subdomains by reviewing the included papers, identifying the cognitive instruments, and grouping these instruments into cognitive subdomains. Disagreements about domain assignment are resolved by consensus. The cognitive subdomains are: verbal memory-immediate recall, verbal memory-delayed recall, verbal memory-recognition, non-verbal memory-immediate recall, non-verbal memory-delayed recall, non-verbal memory-recognition, working memory, attention, intellectual ability, executive function, processing speed, spatial problem solving, global cognitive status, language, motor and construction/visuospatial. In addition, autobiographical memory and subjective memory as measured by standardised self-report tools are included. Notably, narrative comparison of outcomes assessed by objective and subjective measures is critical, given that patients' subjective report of cognitive performance may differ significantly from that captured by objective measurement.

Assessment of risk of bias

Following data extraction, two reviewers will independently assess the risk of bias of each included study. Discrepancies will be resolved by consensus. If consensus is not reached, a third reviewer will decide. The risk of bias of RCTs will be assessed using the Jadad scale[19] which has six questions comprising the following domains: randomisation, double blinding, tracking of withdrawals and adverse effects, appropriate use of statistics, and inclusion and exclusion criteria. We will supplement the questions on the Jadad scale with additional questions (yes/no responses) about the adequacy of allocation concealment, use of intention-to-treat analysis, justification of sample size, reporting of outliers and selective outcome reporting. Some of these additional questions are based on the Cochrane risk of bias tool[20]; the addition of questions to existing scales has been used in other meta-analyses.[21]

The risk of bias of cohort and case–control studies will be assessed using the Newcastle-Ottawa Scale (NOS).[22] The NOS is divided into two subscales, one for cohort and the other for case–control studies. Both subscales assess the following three domains: selection of study groups, comparability of study groups and ascertainment of exposure or outcome.

Using the responses to the aforementioned scales and questions, reviewers will qualitatively assess the risk of bias for each study as 'low', 'unclear' or 'high'. According to the Cochrane Collaboration, 'low' means any bias is unlikely to substantively alter a study's results, 'unclear' means the bias causes doubts about the results, and high means the bias is likely to threaten the validity of the results.[20]

Grading the strength of evidence

We will use the BMJ Evidence Centre guidelines for Grading of Recommendations Assessment, Development, and Evaluation (GRADE)[23] to judge the overall quality of evidence for specific subdomains. In situations where the group of studies assessing a specific subdomain has a low quality of evidence, one would hold little confidence in the validity of the results. One would also be hesitant to draw firm conclusions or make clinical recommendations based on these results. Future studies—assuming they present a higher quality of evidence—might provide a stronger basis from which to draw conclusions or make clinical recommendations.

We will use GRADE to rate the evidence separately for each cognitive subdomain. We will begin by assigning four points to each subdomain if the evidence is largely based on RCT data, or two points if the evidence is largely based on observational study data. We will then assess four other categories, that is, quality, consistency, directness and precision, and add or deduct points for each category in accordance with GRADE guidelines.[23] The additions or deductions reflect preset criteria for assessing how the components of each category contribute to the overall quality of evidence. The final point total serves as the overall GRADE[23] score: scores of 4 or more indicate high quality of evidence, a score of 3 would indicate moderate quality, 2 would suggest low quality and less than 2 would indicate very low quality. The level of confidence to make clinical recommendations based on the evidence would be stronger for higher overall scores.

GRADE's 'quality' category will include the risk of bias assessments. The Cochrane guidelines for ascertaining risk of bias across studies will be used to synthesise the risk of bias findings for individual studies.[20] These guidelines classify groups of studies according to low, unclear or high risk of bias. We will deduct points on the quality category as follows: low risk of bias (−1), unclear risk of bias (−2), high risk of bias (−3).

Statistical analysis

After all data have been extracted from the included studies, the investigators will examine the extraction tables and determine whether meta-analysis is possible. We will only conduct meta-analyses on studies that are relatively homogeneous in terms of participants (eg, age, sex, comorbidity). In the event between-study heterogeneity precludes a meta-analysis, or only permits us to conduct a meta-analysis on a subset of studies, we will undertake a narrative synthesis[24] of all of the included studies.

Studies that are sufficiently homogeneous in terms of participants will be meta-analysed. We will conduct separate meta-analyses for each cognitive subdomain. Within each subdomain, we will stratify the analyses by study design (RCT, observational, RCT and observational combined). The summary estimates computed in the meta-analyses will compare the differences in post-ECT cognitive impairment between groups receiving less versus more conservative ECT treatments. Initially, these comparisons will take the form of mean between-group differences in scale score. Differences in scale score are, however, difficult to interpret across disparate scales because of variations in score ranges (eg, a mean difference of 1.0 is larger on a scale that ranges from 0 to 5 relative to a scale that ranges from 1 to 100). Even standardised mean differences can be difficult to interpret clinically because no threshold exists to mark the minimum important difference in score. Therefore, we will report the study-specific and summary estimates as odds ratios (ORs) in all forest plots. ORs greater than 1.0 will indicate that persons receiving less conservative modalities have greater odds of developing cognitive impairment than persons receiving more conservative modalities. ORs less than 1.0 will show the reverse; ORs equal to 1.0 will suggest no difference between modalities.

We will record all study-specific outcomes as means and SDs or, if unavailable, as mean differences. Borenstein *et al*'s[25] formulae, implemented through Comprehensive Meta-Analysis V.2.0 software,[26] will transform all entered data into ORs and generate forest plots. Forest plots will be computed using a fixed-effects model. We will test statistical heterogeneity for each meta-analysis using the I^2 statistic. If the I^2 value is 50% or higher, then we will recompute the forest plot using a random-effects model. Comprehensive Meta-Analysis will generate funnel plots to enable the assessment of publication bias.

Author affiliations
[1]McMaster Integrative Neuroscience Discovery and Study (MINDS) & Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
[2]School of Public Health and Health Systems, University of Waterloo, Waterloo, Ontario, Canada
[3]Clinical Neuropsychology Service & Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
[4]Department of Psychology, Ryerson University, Toronto, Ontario, Canada
[5]Mood Disorders Program & Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
[6]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada
[7]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada
[8]Faculty of Medicine, Undergraduate Medical Education, University of Toronto, Toronto, Ontario, Canada
[9]McMaster Integrative Neuroscience Discovery and Study (MINDS), McMaster University, Hamilton, Ontario, Canada
[10]Department of Psychology, University of Calgary, Calgary, Ontario, Canada
[11]Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada
[12]Department of Psychiatry, University of Western Ontario, London, Ontario, Canada
[13]Mood Disorders Program, St. Joseph's Healthcare Hamilton, Homewood Research Institute, & Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada

Contributors CO, MO, HM, BL, MK, GH, RL and MM conceived and designed the study. CO conducted the literature search. CO and MP are selecting studies for inclusion in the review (tittle and abstract and full-text screening). CO, MP, GF, ACG, CG, LH, AN, MR, ET, WT and GBCH are extracting data from the included studies and they will assess the risk of bias of included studies. CO drafted the protocol manuscript. All authors critically revised and commented the manuscript for important intellectual content.

Funding This study was funded by a Canadian Institutes of Health Research (CIHR) grant (reference number: MOP-97914) to MM and RL.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. The World Health Organization. *The global burden of disease: 2004 update, table A2: burden of disease in DALYs by cause, sex and income group in WHO regions, estimates for 2004*. Geneva, Switzerland: WHO, 2008. http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_AnnexA.pdf (accessed 6 Mar 2014).
2. Souery D, Papakostas GI, Trivedi MH. Treatment resistant depression. *J Clin Psychiatry* 2006;67(Suppl 6):16–22.
3. Fava M, Davidson KG. Definition and epidemiology of treatment-resistant depression. *Psychiatr Clin North Am* 1996;19:179–200.
4. Mathew SJ. Treatment-resistant depression: recent developments and future directions. *Depress Anxiety* 2008;25:989–92.
5. Keller MB. Issues in treatment-resistant depression. *J Clin Psychiatry* 2005;66(Suppl 8):5–12.
6. Nemeroff CB, Heim CM, Thase ME, *et al*. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proc Natl Acad Sci USA* 2003;100:14293–6.
7. Klein DN, Arnow BA, Barkin JL, *et al*. Early adversity in chronic depression: clinical correlates and response to pharmacotherapy. *Depress Anxiety* 2009;26:701–10.
8. Nanni V, Uher R, Danese A. Childhood maltreatment predicts unfavourable course of illness and treatment outcome in depression: a meta-analysis. *Am J Psychiatry* 2012;169:141–51.
9. UK ECT Review Group. Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. *Lancet* 2003;361:799–808.
10. Kellner CH, Greenberg RM, Murrough JW, *et al*. ECT in treatment-resistant depression. *Am J Psychiatry* 2012;169:1238–44.
11. Lisanby SH. Electroconvulsive therapy for depression. *N Engl J Med* 2007;357:1939–45.
12. Abrams R. *Electroconvulsive therapy*. 4th edn. Oxford, UK: Oxford University Press, 2002.
13. Semkovska M, McLoughlin DM. Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. *Biol Psychiatry* 2010;68:568–77.
14. Oremus C, Oremus M, McNeely H, *et al*. Effects of electroconvulsive therapy on cognition in depression: protocol for a systematic review and meta-analysis. PROSPERO 2014:CRD42014009100. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014009100
15. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–34.
16. Greenhalgh J, Knight C, Hind D, *et al*. Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies. *Health Technol Assess* 2005;9:1–156, iii–iv.
17. Dunne RA, McLoughlin DM. Systematic review and meta-analysis of bifrontal electroconvulsive therapy versus bilateral and unilateral electroconvulsive therapy in depression. *World J Biol Psychiatry* 2012;13:248–58.
18. Charlson F, Siskind D, Doi SA, *et al*. ECT efficacy and treatment course: a systematic review and meta-analysis of twice vs thrice weekly schedules. *J Affect Disord* 2012;138:1–8.
19. Jadad AR, Moore RA, Carroll D, *et al*. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1–12.
20. Higgins JPT, Green S. eds. *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011. http://www.cochrane-handbook.org
21. Oremus M, McKelvie R, Don–Wauchope A, *et al*. A systematic review of BNP and NT-proBNP in the management of heart failure: overview and methods. *Heart Fail Rev* 2014;19:413–19.
22. Wells GA, Shea B, O'Connell D. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm (accessed 7 Oct 2011).
23. BMJ Evidence Centre. What is GRADE?: http://clinicalevidence.bmj.com/x/set/static/ebm/learn/665072.html (accessed 19 Jan 2015).
24. Rodgers M, Sowden A, Petticrew M, *et al*. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. *Evaluation* 2009;15:49–73.
25. Borenstein M, Hedges LV, Higgins JPT, *et al*. *Introduction to meta-analysis*. John Wiley & Sons, Ltd, 2009.
26. Biostat: Comprehensive Meta-analysis software, version 2, Englewood, NJ, 2005.

CHAPTER 4:

EFFECTS OF ELECTROCONVULSIVE THERAPY ON COGNITIVE FUNCTIONING IN PATIENTS WITH DEPRESSION: A SYSTEMATIC REVIEW AND META-ANALYSIS RESULTS & DISCUSSION

Foreword to Chapter 4

This chapter contains the results of the systematic review and meta-analysis of ECT's effect on cognition. The protocol for the systematic review and meta-analysis has been published in a peer-reviewed journal, and it is reproduced in chapter 3 above: "Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis" (*BMJ Open* 2015 11;5:e006966).

The meta-analysis found evidence to suggest poorer cognitive performance in persons who received bilateral versus unilateral ECT, especially in the 8- to 30-day period post-ECT. However, most of the results were not statistically significant and the strength of evidence was weak. Despite decades of research into ECT and potential cognitive adverse effects, a large amount of work remains to be done to provide a firmer scientific understanding of the intensity and length of cognitive sequelae post-ECT.

The work described in chapter 4 will be submitted to *The Lancet* in 2016. For *The Lancet* submission, elements of the published protocol in chapter 3 above will be added to the manuscript to provide introduction and methods sections.

4. Overview

The literature search retrieved 2,640 citations, of which 213 (8%) were duplicates (Figure 4.1). A further 2,039 citations (77%) were excluded at title and abstract screening, thereby leaving 388 citations (15%) for full-text screening. At full-text screening, examination of the published study reports led to the exclusion of 174 articles (45%). A total of 214 articles were included in the review. Twenty-seven articles and one book were included in the systematic review of unilateral versus bilateral ECT. These publications reported data on 19 RCTs (Brakemeier, Berman, Prudic, Zwillenberg, & Sackeim, 2011; Daniel, Weiner, & Crovitz, 1983a; D'Elia, 1970; Devanand, Fitzsimons, Prudic, & Sackeim, 1995; Fleminger, de Horne, & Nott, 1970; Horne, Pettinati, Sugerman, & Varga, 1985; C. H. Kellner et al., 2010; Levy, 1968; McCall, Dunn, Rosenquist, & Hughes, 2002; Ranjkesh, Barekatain, & Akuchakian, 2005; Rosenberg & Pettinati, 1984; Sienaert, Vansteelandt, Demyttenaere, & Peuskens, 2009, 2010; Sobin et al., 1995; Stoppe, Louza, Rosa, Gil, & Rigonatti, 2006; Stromgren & Juul-Jensen, 1975; Taylor & Abrams, 1985a; Tew et al., 2002), eight cohort studies (Ashton & Hess, 1976; Cannicott & Waggoner, 1967; Loo, Sainsbury, Sheehan, & Lyndon, 2008; O'Connor et al., 2008; Schat et al., 2007a; Squire & Chace, 1975; Squire & Slater, 1983; Strain et al., 1968), and one case-control study (Weeks, Freeman, & Kendell, 1980). The book (D'Elia, 1970) reported data on two different RCTs (identified using a single citation in the text). We included 18 studies (Ashton & Hess, 1976; Cannicott & Waggoner, 1967; Daniel et al., 1983a; D'Elia, 1970; Fleminger et al., 1970; Horne et al., 1985; C. H.

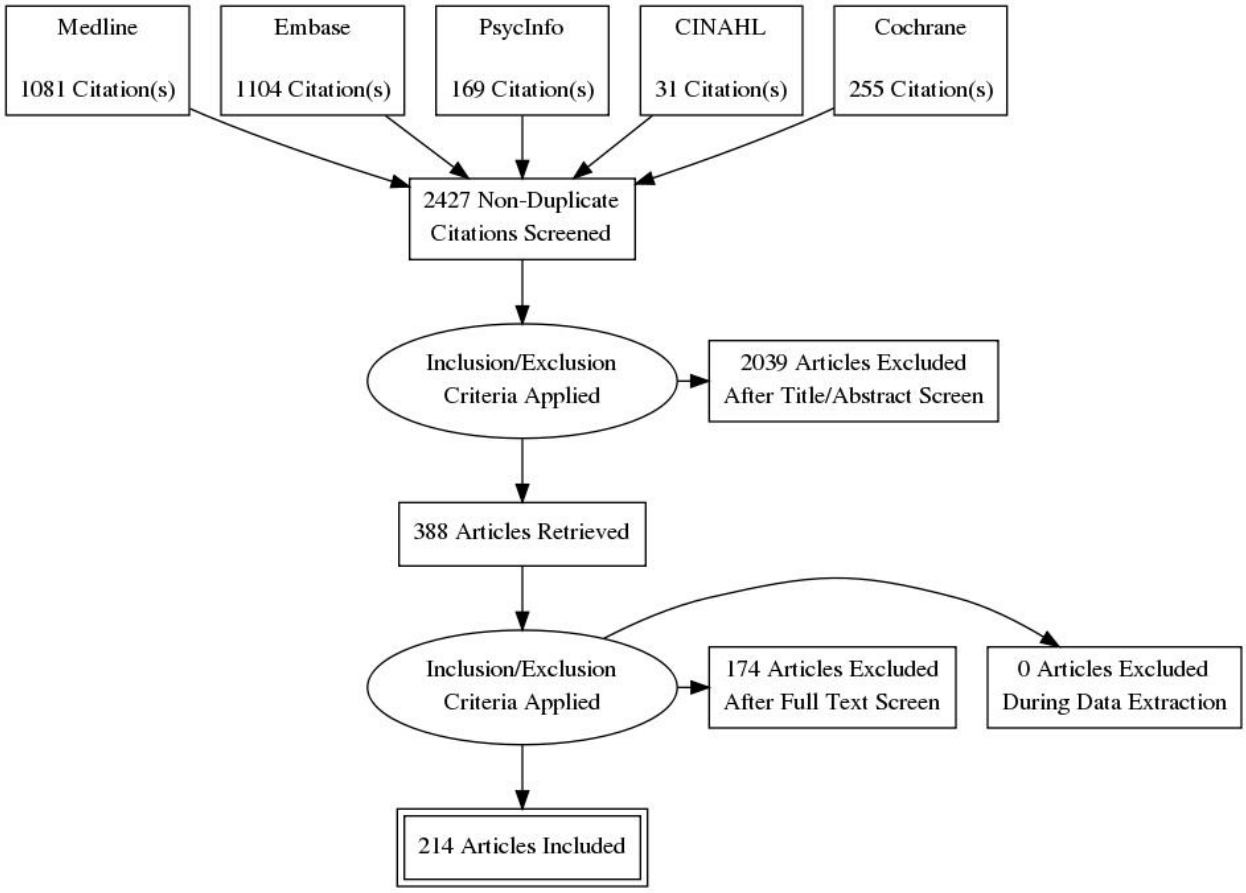
Kellner et al., 2010; Levy, 1968; McCall et al., 2002; Ranjkesh et al., 2005; Rosenberg & Pettinati, 1984; Sienaert et al., 2009; Stoppe et al., 2006; Strain et al., 1968; Stromgren & Juul-Jensen, 1975; Taylor & Abrams, 1985a; Tew et al., 2002; Weeks et al., 1980) in the meta-analysis (see Section 4.1 below and the forest plots in Additional file 1). The meta-analysis encompassed 10 different cognitive domains (see Section 4.1 below). Studies were excluded from the meta-analysis if they did not report on unilateral versus bilateral ECT or if they did not contain data in a format that could be entered into the meta-analysis software. Otherwise, the eligibility criteria for including studies in the review did not change from what was reported in the protocol in Chapter 2.

In actuality, several meta-analyses were performed in this thesis. The meta-analyses were stratified by cognitive domain and by time period, i.e., pre-ECT, 1-7 days post-ECT, 8-30 days post-ECT, 31-183 days post-ECT, and more than 183 days post ECT. Since the purpose of the research was to combine multiple studies and obtain summary overall effects, meta-analyses could only be performed when at least two studies provided data for a particular stratum, limiting significantly our ability to analyze results at 31-183 days ECT and more than 183 days post ECT, where no analyses were possible.

Due to the wide variety of therapeutic variants of ECT (i.e., pulse shape, pulse width, electrode placement, dosage), this research grouped all treatment modalities according to the placement of the electrodes into either ‘bilateral’ or ‘unilateral’ ECT. These groupings made intuitive sense because all forms of ECT are fundamentally bilateral or unilateral. Indeed, bilateral and unilateral ECT are the two classic and still most common types of

ECT. Had the ECT variants not been grouped accordingly, the number of strata would have increased beyond cognitive domain and time period to include treatment modality. With a relatively small number of included studies already in the meta-analyses, increasing the number of strata would have potentially excluded some studies with treatment variants not seen in other papers or obfuscated the clarity of the take-home messages that were otherwise more clearly compartmentalized in a straightforward manner by cognitive domain and time period.

Figure 2. PRISMA Flow Chart



4.1. Results – Meta-analysis: Summary Odds Ratios and 95% Confidence Intervals

At least two studies provided data for the following strata: autobiographical memory 8-30 days; attention 8-30 days; global cognitive status pre-ECT, 8-30 days, and 31-183 days; non-verbal memory – delayed recall pre-ECT and 8-30 days; verbal memory – delayed recall pre-ECT; verbal memory – delayed recall 8-30 days; verbal memory – immediate and delayed recall pre-ECT; subjective memory pre-ECT and 8-30 days post-ECT; executive function pre-ECT and 8-30 days; and motor pre-ECT. Specific findings are described in Sections 4.1.1 – 4.1.10 below.

4.1.1. Autobiographical Memory

At 8-30 days, the results from three studies (Daniel et al., 1983a; C. H. Kellner et al., 2010; Strain et al., 1968) showed poorer cognitive performance that was statistically significant in the bilateral compared to unilateral ECT group (OR: 1.9; 95% CI: 1.1 to 3.5 [$p > 0.05$]).

4.1.2. Attention

At 8-30 days, the results from the analysis of two independent groups within the same study (Horne et al., 1985) showed poorer cognitive performance in the bilateral compared to unilateral ECT group (OR: 1.6 95% CI: 0.7 to 4.0 [$p > 0.05$]).

4.1.3. Global Cognitive Status

Prior to ECT, five studies (Horne et al., 1985; Ranjkesh et al., 2005; Sienaert et al., 2010; Taylor & Abrams, 1985a; Tew et al., 2002) showed baseline cognitive performance that

was statistically significantly worse in persons who would be assigned subsequently to bilateral as compared to unilateral ECT (OR: 1.6; 95% CI: 1.0 to 2.8 [p > 0.05]). At 8-30 days, the results from six studies (Horne et al., 1985; C. H. Kellner et al., 2010; Ranjkesh et al., 2005; Sienaert et al., 2009; Taylor & Abrams, 1985a; Tew et al., 2002) revealed worse cognitive performance in the bilateral ECT group (OR: 2.0; 95% CI: 0.6 to 2.3 [p > 0.05]), while four studies (C. H. Kellner et al., 2010; Ranjkesh et al., 2005; Sienaert et al., 2009, 2010) showed the(Horne et al., 1985; Levy, 1968) same results at 31-183 days (OR: 1.4; 95% CI: 1.0 to 3.9 [p < 0.05]).

4.1.4. Non-verbal Memory – Delayed Recall

Two studies from the same book (D'Elia, 1970) indicated poorer cognitive performance in persons assigned subsequently to the bilateral as compared to unilateral ECT study groups (OR: 1.5; 95% CI: 0.8 to 3.1 [p > 0.05]). At 8-30 days, the results from four studies (Ashton & Hess, 1976; C. H. Kellner et al., 2010; McCall et al., 2002; Weeks et al., 1980) showed poorer cognitive performance in the bilateral as compared to unilateral ECT group (OR: 2.6; 95% CI: 1.0-6.7 [p > 0.05]).

4.1.5. Verbal Memory – Delayed Recall

Prior to ECT, three studies, including two of the studies published in the book (D'Elia, 1970; McCall et al., 2002) reported worse cognitive performance in persons who were about to receive bilateral versus unilateral ECT (OR: 1.4; 95% CI: 0.8 to 2.5 [p > 0.05]).

4.1.6. Verbal Memory – Immediate Recall

At 8-30 days post-ECT, the results from two studies (Horne et al., 1985; Weeks et al., 1980) showed worse cognitive performance that was statistically significant in the bilateral as compared to unilateral ECT group (OR: 5.7; 95% CI: 1.5 to 21.6 [p > 0.05]).

4.1.7. Verbal Memory – Immediate and Delayed Recall

Two studies (Horne et al., 1985; Levy, 1968) reported worse cognitive performance in persons who would be assigned subsequently to bilateral ECT (OR: 2.9; 95% CI: 0.9 to 9.8 [p > 0.05]). At 8-30 days post-ECT, five studies (Cannicott & Waggoner, 1967; Fleminger et al., 1970; Horne et al., 1985; C. H. Kellner et al., 2010; Levy, 1968) found worse cognitive impairment in persons that was statistically significant who received bilateral as compared to unilateral ECT (OR: 5.7; 95% CI: 1.5 to 21.6 [p > 0.05]).

4.1.8. Subjective Memory

As with verbal memory – immediate and delayed recall, worse cognitive performance was found in the bilateral as compared to unilateral ECT group prior to ECT (OR: 1.4; 95% CI: 0.6 to 3.7 [p > 0.05]) (Cannicott & Waggoner, 1967; Levy, 1968) and at 8-30 days post-ECT (OR: 4.1; 95% CI: 1.2 to 13.7 [p < 0.05]) (Levy, 1968; Rosenberg & Pettinati, 1984; Stromgren & Juul-Jensen, 1975).

4.1.9. Executive Function

Assignment to the bilateral as compared to unilateral ECT group was associated with worse cognitive performance pre-ECT (1.3 (95% CI: 0.5 to 3.7 [p > 0.05]) (Horne et al.,

1985) and at 8-30 days post-ECT (OR: 1.3; 95% CI: 0.8 to 2.4 [p > 0.05]) (Horne et al., 1985; C. H. Kellner et al., 2010).

4.1.10. Motor

Persons assigned to the bilateral ECT group showed worse pre-ECT motor functioning than did persons assigned to unilateral ECT (OR: 2.5; 95% CI: 0.9 to 7.0 [p > 0.05]) .

4.2. Risk of Bias

The risk of bias for the RCTs, measured using the Jadad scale (Jadad et al., 1996) and the Cochrane risk of bias tool, (Higgins et al., 2011) was medium (Table 4.1). Only two studies had a low risk of bias (Charles H. Kellner, Tobias, & Wiegand, 2010; Sobin et al., 1995), nine had an unclear risk of bias (Brakemeier et al., 2011; D'Elia, 1970; Horne et al., 1985; Levy, 1968; McCall et al., 2002; Ranjkesh et al., 2005; Stoppe et al., 2006; Stromgren & Juul-Jensen, 1975), and eight had a high risk of bias (Daniel et al., 1983a; Devanand et al., 1995; Fleminger et al., 1970; Rosenberg & Pettinati, 1984; Sienaert et al., 2009, 2010; Taylor & Abrams, 1985a; Tew et al., 2002). The principal problems with many of the studies were inadequate or unclear appropriateness of randomization or double-blinding (where blinding was present), inadequate or unclear allocation concealment, unclear use of intent-to-treat analysis, no justification of sample size, and no reporting of outliers.

Turning to the cohort studies, three had a low risk of bias (Loo et al., 2008; O'Connor et al., 2008; Schat et al., 2007a) and five had an unclear risk of bias (Ashton & Hess, 1976; Cannicott & Waggoner, 1967; Squire & Chace, 1975; Squire & Slater, 1983;

Strain et al., 1968) on the Newcastle-Ottawa Scale (Lo, Mertz, & Loeb, 2014) (Table 4.2). The primary issues with the cohort studies were uncertainty over the representativeness of the cohort, uncertainty over the comparability of study groups, and no description of outcome assessment.

The lone case-control study (Weeks et al., 1980) performed well on the Newcastle-Ottawa Scale. The only major issue was a lack of description of control selection.

4.3. Grading the Strength of Evidence

The strength of evidence (Table 4.3) was remarkably consistent across the studies that were meta-analyzed. In all cases, inconsistency and indirectness were not issues because the odds ratios were consistently above 1.0, indicating worse cognitive outcomes for bilateral versus unilateral ECT. Additionally, all of the results were obtained through direct comparison between the two types of ECT. Unilateral and bilateral ECT were not indirectly compared to one another through some third mechanism. Imprecision was an issue because most 95% confidence intervals were quite large, mainly due to the small samples sizes in most included studies. Imprecision generates wide confidence intervals and large p-values, which renders interpretation of results more difficult because one cannot clearly determine where the true effect lies (e.g., how far from the null value does an odds ratio lie). When the studies contributing evidence to each domain were taken together, many had unclear or high risk of bias, which counterbalanced the low risk of bias in other studies. Thus, the risk of bias was rated as serious or very serious for eight of 10 cognitive domains.

Due to the serious nature of the imprecision and risk of bias, the overall quality of evidence was very low for four domains (global cognitive status, autobiographical memory, subjective memory, verbal memory – immediate delayed recall), low for four domains (attention, non-verbal memory – delayed recall, verbal memory – delayed recall, motor), and moderate for two domains (executive function, verbal memory – immediate recall) (Table 4.3).

4.5. Discussion

Overall, we meta-analyzed 18 studies across 10 different cognitive domains. The two major time points in the extracted studies were pre-ECT and 8 to 30 days post ECT. In the pre-ECT groups, the point estimated odds ratios were primarily within a range of 1.3 to 1.6, with two point estimates at 2.5 (motor) and 2.9 (verbal memory-immediate and delayed recall). In the 8- to 30-day timeframe post-ECT, the point estimated odds ratios ranged from 1.3 to 10.5, with five odds ratios exceeding 2.0 (double the odds of worse cognitive performance in bilateral versus unilateral ECT in the domains of global cognition, non-verbal delayed recall, verbal memory immediate and delayed recall, subjective memory, and verbal memory immediate recall). A recent review (Semkovska & McLoughlin, 2010) suggested the cognitive impact of ECT does not last beyond three days. Here, we found worse cognitive performance for bilateral versus unilateral ECT at 8-30 days post-ECT, suggesting that this treatment modality is associated with worse cognitive outcomes for the following domains in the month following ECT treatment: autobiographical memory, attention, global cognitive status, non-verbal memory –

delayed recall, verbal memory – delayed recall, subjective memory, and executive function. Notably, persons assigned to bilateral ECT showed worse cognitive performance pre-ECT in the following domains: global cognitive status, non-verbal memory – delayed recall, verbal memory – delayed recall, verbal memory – immediate and delayed recall, subjective memory, executive function, and motor. These pre-ECT results suggest that patients entering into bilateral ECT treatment may be more vulnerable to cognitive impairment prior to treatment entry than those undergoing unilateral ECT.

Our findings must be taken within the context of the data from the meta-analyzed studies. Only four summary odds ratios were statistically significant at the 5% level, and all of them pertained to 8- to 30-day periods in the domains of subjective memory (OR: 4.1; 95% CI: 1.2 to 13.7), verbal memory immediate and delayed recall (OR: 5.7; 95% CI: 1.5 to 21.6), verbal memory immediate recall (OR: 10.5; 95% CI: 2.0 to 53.0), and executive function (OR: 1.9; 95% CI: 1.1 to 3.5). None of the other summary odds ratios attained statistical significance. The lack of significance was likely due to small sample sizes in many of the studies. When combined with the predominantly 'low' ratings from the GRADE analysis, the lack of statistical significance indicates that further evidence could well change the findings of this meta-analysis.

We note that our findings differ from the findings of an earlier meta-analysis examining cognitive performance in ultrabrief versus brief pulse ECT (Tor et al., 2015). In this earlier study, the summary standard mean differences (SMDs) for cognitive function all unequivocally favoured ultra-brief pulse ECT: retrograde memory (SMD:

0.38; 95% CI: 0.15 to 0.61 [5 studies]); anterograde memory—learning (SMD: 0.45; 95% CI: 0.22 to 0.68 [2 studies]); anterograde memory—delayed recall (SMD: 0.56; 95% CI: 0.40 to 0.73 [3 studies]); global cognitive function (SMD: 0.36; 95% CI: 0.09 to 0.63 [2 studies]). The difference in findings can be explained by the different thrust of this earlier meta-analysis (Tor et al., 2015), which was conducted to specifically examine two variants of ECT, while the meta-analysis in this thesis examined a broader spectrum of ECT modalities. Still, both meta-analyses found cognitive performance to be poorer in persons who received more intense variants of ECT.

The meta-analysis reporting cognitive adverse effects to last only within three days post-ECT combined studies according to cognitive test (Semkovska & McLoughlin, 2010). In the ≤ 3-day period post-ECT, the authors reported statistically significant decreases on 72% of the cognitive tests. Given that many studies in this area include multiple tests to measure the same cognitive domain, study-specific features such as selection or information bias can impact all of the measures of cognition, thus leading to an over-inflation of effects in a meta-analysis. For example, a study that recruits more severely depressed persons who might be indicated to receive a more intense form of ECT could produce results showing a high incidence of reduced cognitive performance post-ECT. However, this finding would be partially the result of the type of participants selected for inclusion in the study. If such a study employed several measures of the same cognitive domain, whereas another study employed just one measure in that same domain while also recruiting a more balanced sample in terms of TRD severity, then the results of

the first study could have a larger effect on the overall conclusions of a meta-analysis if the analysis was combining studies by measure. The suspicion is this type of occurrence was present in the other meta-analysis (Semkovska & McLoughlin, 2010). Indeed, methods recommendations suggest meta-analysts should avoid vote-counting to draw conclusions about summary treatment effects (Higgins et al., 2011). This is because vote counting ignores substantive differences between studies and often reduces a conclusion about effects to whether study results meet the $p < 0.05$ threshold. From this discussion, the 72% figure cited in the meta-analysis should be interpreted with caution.

In the meta-analysis reported in this thesis, the problem of weighting studies by numbers of tests was avoided by meta-analyzing by domain. When a single study reported more than one test per domain, decisions had to be made on how to approach the issue to avoid double-counting the same participant results in the meta-analysis. Without methods guidance in the literature, the decision was made to choose a test was based on sample size to increase the power of the meta-analysis (i.e., taking the test for whom the authors report results for the most participants). When sample sizes across tests were the same, the next option was to select tests that were similar to the tests employed by the other studies in the same domain (to enhance comparability).

The findings of this and other meta-analyses, despite different objectives and methods, suggest that more intense types of ECT appear to be associated with greater odds of worsening cognitive performance. However, the evidence to date does not provide a clear picture of the extent to which the poorer performance persists over time,

nor does it suggest which cognitive domains may be more susceptible to injury following ECT. Clearly, further research is needed to elucidate these relationships and provide clinicians with guidance regarding the type of ECT to select as a treatment for persons with TRD. The clinician's treatment choice is guided by a risk-benefit analysis: what type of ECT will best suit my patients and will the benefits of the treatment outweigh the risks of potentially worsening cognitive performance? The current state of the evidence does not help the clinician grapple with this question. Further research using clearly defended ECT protocols and a small set of generally accepted cognitive tests should be launched to investigate the question in more detail. The psychiatric community should come to consensus on a standard battery of cognitive tests to enhance the applicability and comparability of results to strengthen the body of evidence. Moreover, to reduce bias in the conduct of research studies and to enhance the clarity of presenting results, researchers should follow recommended guidelines for reporting results, e.g., CONSORT for RCTs (Moher et al., 2010), STROBE for observational studies (Elm et al., 2007), and PRISMA for systematic reviews and meta-analyses (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). Abrams describes how the impact of potentially important ECT-cognition research can be dampened by methodological flaws and poor reporting (Abrams, 2007).

Notably, previous meta-analyses have revealed small to moderate effect sizes for cognitive deficits in depression in the absence of ECT treatment. A meta-analysis by Rock et al., 2014 searched PubMed and Google Scholar between 1980 to December 2012

and included 24 studies comparing currently depressed patients to healthy controls and six studies comparing remitted depressed patients to healthy controls. The authors reported moderate deficits in memory, executive function and attention (Cohen's d effect sizes ranging from -0.34 to -0.65) in currently depressed vs. healthy controls; and memory deficits (Cohen's d ranging from 22 to 0.54) in remitted depressed patients. However, the I^2 for sixteen out of twenty four studies pool results range from 56 to 82 indicating substantial statistical heterogeneity between studies.

Another meta-analysis by Lee et al., 2011 searched PubMed and PsychInfo databases from 1990 to February 2011 summarizing 13 studies. The authors reported that patients in their first episode of depression performed worse than healthy controls in attention (SMD: 0.36, 95% CI: 0.13-0.59; I^2=0%), working memory (SMD: 0.16, 95% CI: -0.20 - 0.51; I^2=61%), verbal learning and memory (SMD: 0.13, 95% CI: 0.18-0.45; I^2=81%), visual learning and memory (SMD: 0.53, 95% CI: -0.05 – 1.11; I^2=88%). Unfortunately, the heterogeneity of the pooled results for working memory, verbal and learning memory and visual and learning memory heterogeneity range from 61% to 88% which warns caution in interpreting the findings.

In addition, a meta-analysis from Wagner et al., 2012 included 15 studies investigating severity of executive dysfunctions in persons with MDD in comparison to healthy controls and 3 before and after antidepressant treatment studies. The authors found that healthy controls had better cognitive functioning than persons with MDD in semantic and phonemic memory 0.92 SD and 0.71 SD, I^2=52%; Stroop interference test

1.18 SD, Trail Making Test B 1.109 SD. These previous meta-analyses did not strictly adhere to PRISMA guidelines, nor did they evaluate the risk of bias and did not measure the strength of evidence. Interpretation of these findings of should be done with caution due to moderate to high levels of heterogeneity (> 50%). Taken together, however, these meta-analyses suggest that depression alone is associated with changes in cognitive functioning. The extent to which these alterations may be exacerbated by ECT remains to be elucidated.

The current meta-analysis has many strengths, including adherence to current methods guidance (Moher et al., 2009) and the use of risk of bias and strength of evidence assessments to help interpret the findings. The major limitation is the small number of studies that could be meta-analyzed, although this is more of a function of the vast amount of clinical heterogeneity seen in the included studies. Studies varied in terms of participants (e.g., age, co-morbidities), treatment modalities, cognitive measures, and frequency and timing of follow-ups (Table 4.4).

A major challenge was selecting studies that were 'similar enough' to include in the meta-analysis. The thesis author pooled studies whose design and operational characteristics were as homogeneous as possible. Failing to account for clinical heterogeneity when making a priori decisions about which studies to include in a meta analysis can produce statistical heterogeneity, which exaggerates the variability of summary treatment effects (Gagnier, Moher, Boon, Beyene, & Bombardier, 2012). The author also had to group studies under two broad categories of treatment to prevent the

unwieldy growth of multiple strata, which would have made summary interpretations of results difficult. Some of the reasons for variability in the included studies are the use of different neuropsychological tests to measure a multiplicity of cognitive domains, multiple time points following ECT (including immediately post-ECT), substantial clinical heterogeneity in the study samples, problems with the data (i.e., many data were not normally distributed), and differences in how data were reported (i.e., different measures of association).

The non-meta-analyzed studies generally pointed to the same conclusions as the subset of studies that were included in the meta-analysis. Within-group comparisons were not undertaken in the meta-analysis because many studies, in particular those published prior to 2000, did not report baseline sample characteristics in great depth. Also, many studies were not designed to investigate cognitive performance as a primary outcome and some of these studies looked at cognition as an adverse effect and measured it only at follow-up. For these same reasons, further stratifying the analyses by baseline cognitive status was not possible.

In conclusion, while our meta-analysis found evidence to suggest poorer cognitive performance in persons who receive bilateral versus unilateral ECT, most of our results were not statistically significant and the strength of evidence was weak. More research is required to obtain a deeper understanding of the association between ECT and cognition. Perhaps most striking, given decades of research on this topic, is the large amount of

remaining work to be done to provide a firmer scientific basis to guide clinical decision-making

4.5. Tables and Figures

Table 3. Risk of Bias Assessment - RCT-

Author, Year Country	Randomized		Appropriate randomization		Double blind		Appropriate double blinding		Description withdrawals/ drop-outs		Clear description inclusion / exclusion criteria		Adverse effects described		Statistical analysis described		Adequacy of Allocation Conceal-ment		Use of Intention-to-treat Analysis		Justification of Sample Size		Reporting of Outliers		Selective Outcome Reporting		Total Jadad Score	Total Cochrane Score	Total Jadad + Cochrane Score
	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score			
Brakemeier, 2011, USA	Y	1	Y	1	N	0	N	0	Y	1	Y	1	N	0	Y	1	N	0	Y	1	N	0	N	0	N	1	5	2	7
d'Elia, 1970, Sweden Study IV	Y	1	U	0	Y	1	U	0	N	0	Y	1	Y	1	Y	1	N	0	Y	1	N	0	N	0	N	1	5	2	7
d'Elia, 1970, Sweden Study V	Y	1	U	0	Y	1	U	0	N	0	Y	1	Y	1	Y	1	N	0	Y	1	N	0	N	0	N	1	5	2	7
Daniel WF, 1983 USA	Y	1	U	0	N	0	N	0	Y	1	N	0	N	0	N	0	N	0	N	0	N	0	N	0	N	1	2	1	3
Devanand, 1995, USA	Y	1	U	0	Y	1	N	-1	N	0	Y	1	Y	1	Y	1	N	0	N	0	N	0	N	0	N	1	4	1	5
Fleminger, 1970 England	Y	1	N	-1	Y	1	Y	1	Y	1	N	0	N	0	Y	1	N	0	N	0	N	0	N	0	N	1	4	1	5
Horne R, 1985, USA	Y	1	Y	1	Y	1	Y	1	N	0	Y	1	N	0	Y	1	N	0	N	0	N	0	N	0	N	1	6	1	7
Kellner C.H., 2010, USA	Y	1	U	1	Y	1	Y	1	Y	1	Y	1	N	0	Y	1	N	0	Y	1	Y	1	N	0	N	1	7	3	10
Levy R, 1968, UK	Y	1	Y	1	Y	1	N	0	N	0	Y	1	N	0	N	0	N	0	Y	1	N	0	N	0	N	1	4	2	6

Author, Year Country	Randomized		Appropriate randomization		Double blind		Appropriate double blinding		Description withdrawals/ drop-outs		Clear description inclusion / exclusion criteria		Adverse effects described		Statistical analysis described		Adequacy of Allocation Conceal-ment		Use of Intention-to-treat Analysis		Justification of Sample Size		Reporting of Outliers		Selective Outcome Reporting		Total Jadad Score	Total Cochrane Score	Total Jadad + Cochrane Score
	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score	Y/N	Score			
McCall, 2002, USA	Y	1	U	0	Y	1	Y	1	N	0	Y	1	N	0	Y	1	N	0	N	0	Y	1	N	0	N	1	5	2	7
Ranjkesh, 2005, Iran	Y	1	U	0	Y	1	U	0	Y	1	Y	1	y	1	N	0	N	0	Y	1	N	0	N	0	N	1	5	2	7
Rosenberg, 1984, USA	Y	1	N	-1	Y	1	N	-1	N	0	Y	1	N	0	N	0	N	0	N	0	N	0	N	0	N	1	1	1	2
Sand Stromgren and Juul-Jensen	Y	1	U	0	N	0	N	0	Y	1	Y	1	N	0	Y	1	N	0	Y	1	N	0	N	0	N	1	4	2	6
Sienaert, 2009 Belgium	Y	1	U	0	N	0	N	-1	Y	1	Y	1	N	0	N	0	N	0	N	0	N	0	N	0	N	1	2	1	3
Sienaert 2010 belgium	Y	1	U	0	N	0	N	0	Y	1	Y	1	N	0	Y	1	N	0	N	0	N	0	N	0	N	1	4	1	5
Sobin, 1995, USA	Y	1	U	0	Y	1	Y	1	Y	1	Y	1	Y	1	Y	1	N	0	Y	1	N	0	N	0	N	1	7	2	9
Stoppe A, 2006, Brazil	Y	1	N	0	N	0	N	0	N	0	Y	1	N	0	Y	1	N	0	Y	1	N	1	N	0	N	1	3	3	6
Taylor & Abrams, 1985, Britain	Y	1	U	0	N	0	N	-1	Y	1	Y	1	N	0	N	0	N	0	N	0	N	0	N	0	N	1	2	1	3
Tew JD, 2002, USA	Y	1	U	0	Y	1	Y	1	N	0	N	0	N	0	Y	1	N	0	N	0	N	0	N	0	N	1	4	1	5

Table 4. Risk of Bias Assessment

Author, year Country	Is the case definition adequate?		Representativeness of the cases		Selection of controls		Definition of controls?		Comparability of cases and controls		Ascertainment of exposure		Ascertainment same for cases and controls		Non-response rate		Total Score
Ashton and Hess, 1976, Australia	d	0	a	1	a	1	a	1	c	0	a	1	a	1	a	1	6
Cannicott SM, 1967, USA	c	0	a	1	b	1	b	0	c	0	d	0	a	1	a	1	4
Loo, 2008, Australia	a	1	a	1	b	1	b	1	b	2	a	1	b	0	c	0	7
O'Connor M, 2008 USA	a	1	a	1	b	1	a	1	a	1	a	1	a	1	b	1	8
Schat, 2007, Netherlands	a	1	a	1	b	1	b	0	b	2	a	1	a	1	c	0	7
Squire 1983 USA	b	1	a	1	b	1	a	1	c	0	c	0	a	1	b	1	6
Squire, 1975 USA	c	0	a	1	a	1	b	0	c	0	d	0	a	1	b	1	4
Strain, 1968, USA	d	0	a	1	a	1	a	1	a	1	a	1	a	1	a	1	7

Table 5 GRADE

Cognitive Deficit: Bilateral versus Unilateral Electroconvulsive Therapy						
Quality assessment						
№ of participants (studies) Follow-up	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Overall quality of evidence**
Global Cognitive Status						
6 RCTs	very serious [1]	not serious	not serious	serious [2]	none	⊕◯◯◯ VERY LOW
Autobiographical Memory						
2 RCTs, 1 cohort	serious [3]	not serious	not serious	serious [2,4]	none	⊕◯◯◯ VERY LOW
Attention						
1 RCT	serious [4]	not serious	not serious	serious [2]	none	⊕⊕◯◯ LOW

Executive Function						
2 RCTs	not serious [5]	not serious	not serious	serious [2,7]	none	⊕⊕⊕◯ MODERATE
Non-verbal Memory - Delayed Recall						
3 RCTs, 1 case control	serious [6]	not serious	not serious	serious [2,9]	none	⊕⊕◯◯ LOW

Verbal Memory - Delayed Recall						
2 RCTs	serious [7]	not serious	not serious	serious [2,11]	none	⊕⊕◯◯ LOW
Verbal Memory - Immediate Recall						
2 RCTs	not serious	not serious	not serious	serious [2,13]	none	⊕⊕⊕◯ MODERATE

Subjective Memory						
3 RCTs, 1 cohort	very serious [8]	not serious	not serious	serious [2]	none	⊕◯◯◯ VERY LOW
Verbal Memory - Immediate Delayed						
4 RCTs, 1 cohort	very serious [9]	not serious	not serious	serious [2]	none	⊕◯◯◯ VERY LOW
Motor						
1 RCT	serious [110]	not serious	not serious	serious [2]	none	⊕⊕◯◯ LOW

CI: Confidence interval

1. Two studies high risk of bias; two studies unclear risk of bias
2. Wide confidence intervals
3. One study high risk of bias; one study unclear risk of bias
4. Unclear risk of bias
5. One study unclear risk of bias; one study low risk of bias
6. Three studies unclear risk of bias; two studies low risk of bias
7. Two studies unclear risk of bias
8. One study high risk of bias; three studies unclear risk of bias
9. One study high risk of bias; three studies unclar risk of bias
10. One study unclear risk of bias

Table 6. Summary table

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Ashton R, 1976, Australia	Cohort (prospective)	Endogenous or Reactive Depression (N = 15) Note: May be either recruitment or randomization sample	BL Temporal (n = 8)	Right UL – Non-Dominant - Temporal Parietal (n = 7)	Non-Verbal Memory-Recognition	16-point abstract shapes - Recognition	**8-30 Days Post 1st ECT;** BL (n = 8); UL (n = 7)	Y	
Bidder TG, 1970, USA	RCT	Depressed – Non-Schizophrenic (N = 96) Note: Could be either randomization of recruitment sample	Bilateral (n = 46)	Right UL – Temporal Parietal (n = 50)	Autobiographical Memory	Personal Data Sheet	**Pre ECT;** BL (n = 46); RUL (n = 50) **Change from Pre ECT to 2 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 2 ECT treatments to 4 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 4 ECT treatments to**	N	No information given on ECT scheduling; can not determine the number of days since the 1st ECT for follow up times

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
							6 ECT Treatments; BL (n = 46); RUL (n = 50) **10 Days After ECT Completion;**		
					Non-Verbal Memory-Delayed Recall	BVRT	**Pre ECT;** BL (n = 46); RUL (n = 50) **Change from Pre ECT to 2 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 2 ECT treatments to 4 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 4 ECT treatments to 6 ECT Treatments;** BL (n = 46); RUL (n = 50) **10 Days After ECT Completion;** **30 Days After ECT Completion;** **1 year After ECT Completion;**		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Verbal Memory - Immediate Recall	PAL	**Pre ECT;** BL (n = 46); RUL (n = 50) **Change from Pre ECT to 2 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 2 ECT treatments to 4 ECT Treatments;** BL (n = 46); RUL (n = 50) **Change from 4 ECT treatments to 6 ECT Treatments;** BL (n = 46); RUL (n = 50) **10 Days After ECT Completion;** **30 Days After ECT Completion;** BL (n = 9); RUL (n = 12) **1 year After ECT Completion;**		
Brakemeier, 2011, USA	RCT	Major Depressive Disorder (N	Bilateral – Frontal Temporal –	Right UL – 6.0x Seizure Threshold–	Autobiographical Memory	AMI	**Pre ECT:** Can not determine with certainty	N	

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		= 90)	Ultra Brief Pulse Width (n =) BL Temporal – 2.5x Seizure Threshold – Brief Pulse Width (n =)	Ultra Brief Pulse Width (n =) Right UL – 6.0x Seizure Threshold– Ultra Brief Pulse Width (n =)	Global Cognitive Status	MMSE	**8-30 Days Post 1st ECT;** Can not determine with certainty		
					Subjective Memory	CFQ			
						GSE - My			
						SMCQ			
					Non Verbal Memory- Delayed Recall	RFT			
					Verbal Memory - Immediate and Delayed Recall	WAIS - III			
Cannicott SM, 1967, USA	Cohort (prospective)	Major Depressive Disorder (N = 34) Note: randomization, not recruitment, sample	BL (n = 10)	UL – Non-Dominant (n = 14)	Spatial Problem Solving	Block Design	**Pre ECT;** BL (n = 10); UL – Non-Dominant (n = 14) **8-30 Days Post 1st ECT;** BL (n = 10); UL – Non-Dominant (n = 14)	Y	
					Subjective Memory	Recent Memory			
					Verbal memory- Immediate and Delayed Recall	Immediate Memory			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		Criteria for inclusion available in Daniel et al. (1982); could be either randomization or recruitment sample	= 6) BL -Frontal Temporal – Sinusoidal (n = 7)	Parietal – Brief Pulse (n = 4) UL - Non-Dominant - Temporal Parietal – Sinusoidal (n = 5)			Sinusoidal (n = 7); UL — Brief Pulse (n = 4); UL – Sinusoidal (n = 5)		
D'Elia G, 1970, Sweden (Study 3)	RCT	Unipolar Depression (n = 31) Bipolar Depression (n = 4)	BL -Frontal Temporal(n = 25)	UL – Non-Dominant – Temporal Parietal (n = 28)	Non-Verbal Memory-Delayed Recall	Delayed Reproduction - 30 Figure Test	**Pre ECT:** BL (n = 25); UL (n = 28) **31-183 Days Post 1st ECT:** BL (n = 25); UL (n = 28)	Y	
					Non-Verbal Memory Immediate Recall	Immediate Reproduction - 30 Figure Test			
					Verbal Memory-Delayed Recall	Delayed Reproduction - 30 Word Pair Test			
						Delayed Reproduction - 30 Personal Data Test			
					Verbal Memory Immediate Recall	Immediate Reproduction - 30 Word Pair Test			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
						Immediate Reproduction - 30 Personal Data Test			
D'Elia G, 1970, Sweden (Study 4)	RCT	Recruited (N = 108) Schizophrenic psychosis (n = 9) Manic depressive episode (n = 32) Paranoid psychosis (n = 4) Psychoneurosis (n = 29)	BL - Frontal Temporal (n = 24)	UL – Non-Dominant – Temporal Parietal (n = 25) UL – Dominant – Temporal Parietal (n = 25)	Non Verbal Memory - Delayed Reproduction	Delayed Reproduction - 30 Figure Test	**Within 24hrs Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25) **1-7 Days Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25)	Y	
					Non Verbal Memory - Immediate Reproduction	Immediate Reproduction - 30 Figure Test	**Pre ECT:** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25) **1-7 Days Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25)		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Verbal Memory - Delayed Recall	Delayed Reproduction - 30 Word Pair Test Delayed Reproduction - Personal Data Test	**Within 24hrs Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25) **1-7 Days Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25)		
					Verbal Memory - Immediate Recall	Immediate Reproduction - 30 Word Pair Test Immediate Reproduction - Personal Data Test	**Pre ECT:** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25) **1-7 Days Post 1st ECT;** BL (n = 24); UL – Non-Dominant (n = 25); UL – Dominant (n = 25)		
Devanand DP, 1995, USA*	RCT	Major Depressive Disorder (N = 100)	BLTemporal – Just Above Seizure Threshold (n =	Right UL – Temporal Parietal - Just Above Seizure	Subjective Memory	Cognitive Symptoms	**8-30 Days Post 1st ECT;**	N	No record in forest plots or manuscript

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
			23) BLTemporal – 2.5x Seizure Threshold (n = 27)	Threshold (n = 23) Right UL – Temporal Parietal - 2.5x Seizure Threshold (n = 23)		Memory Complaints	**8-30 Days Post 1st ECT;**		under meta-analysis Number of subjects per treatment condition at 8-30 days post 1st ECT can not be determined with certainty
						Subjective Memory (Disorientation)	**8-30 Days Post 1st ECT;**		
Halliday, 1968, …	RCT	(N = 52) Note: randomized, not recruitment, sample	BL (n = 18)	UL – Left Hemisphere n = 16) UL – Right Hemisphere (n = 18)	Attention	Digits Backward	**Change in Score After 4th ECT Treatment;** BL (n = 18); UL-Left Hemisphere (n = 16); UL – Right Hemisphere (n = 18) **Change in Score After 3 Months;** BL (n = 17); UL-Left Hemisphere (n = 11); UL – Right Hemisphere (n = 16)	N	Does not provide any information on ECT scheduling
						Digits Forward			
					Autobiographical Memory	Remote Memory Test			
					Non-Verbal Memory (?)	Rey Davis Board (repeated as a test of retention)			
					Verbal Memory – Immediate and Delayed Recall	Learn the meaning of five unfamiliar words (errors)			
					What would the delayed	Delayed Recall			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					recall test be categorized as?				
Horne RL, 1985, USA	RCT	Major Depressive Disorder (N = 48)	BL Bi-Temporal – Placebo (n = 12) BL Bi-Temporal – Dexamethasone (n = 12)	UL – Non-Dominant – Temporal Parietal – Placebo (n = 12) UL – Non-Dominant – Temporal Parietal – Dexamethasone (n = 12)	Attention	Digits Backward Digits Forward Trail Making B	**0-30 Days Post 1st ECT;** BL (n = 22); UL (n = 21)	Y	Degrees of freedom on Table 5 suggest differences in the number of patients who completed each test
						Random Numbers	**Pre ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12) **0-30 Days Post 1st ECT;** BL (n = 22); UL (n = 21) **8-30 Days Post 1st ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12)		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Executive Functioning	Trail Making B(s)	**Pre ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12) **8-30 Days Post 1st ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12)		
					Global Cognitive Status	Number of Objects Recalled			
					Motor	Pegboard (errors)			
					Short Term Memory	Short story (recall immediately)	**8-30 Days Post 1st ECT;** BL (n = 22); UL (n = 21)		
						No. of objects recalled			
						Paired Words			
						Pegboard (errors)	**0-30 Days Post 1st ECT;** BL (n = 22); UL (n = 21)		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Verbal Memory-Immediate and Delayed Recall	Paired Words	**Pre ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12) **8-30 Days Post 1st ECT;** Dexamethasone; BL (n = 22); UL (n = 21) Placebo; BL (n = 12); UL (n = 12)		
					Verbal Memory-Immediate Recall	Short Story			
						Digits Backward			
						Digits Forward			
Kellner CH, 2010, USA	RCT	Recruited (N = 274) Major Depressive Disorder (N = 237) Note: Unipolar and bipolar included	BL – Bifrontal – 1.5x Seizure Threshold (n = 81) BL - Bitemporal - 1.5x Seizure Threshold (n = 72)	Right UL – Temporal Parietal – 6x Seizure Threshold (n = 77)	Attention	Trail Making A	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 65); BL - Bitemporal (n = 56); UL (n = 57)	Y	
					Autobiographical Memory	AMI	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 69); BL - Bitemporal (n = 60); UL (n = 60)		
					Executive Functioning	Category Fluency	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 66); BL - Bitemporal (n =		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
							58); UL (n = 59)		
						COWAT	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 68); BL - Bitemporal (n = 57); UL (n = 60)		
						D-KEFS	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 61); BL - Bitemporal (n = 52); UL (n = 52)		
						Stroop Color Word Test	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 62); BL - Bitemporal (n = 56); RUL (n = 55)		
						Trail Making B	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 63); BL - Bitemporal (n = 54); UL (n = 57)		
					Global Cognitive Status	CGI Severity Scale	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 71); BL – Bitemporal (n = 61); UL (n = 62)		
						MMSE			
					Non Verbal Memory-Delayed Recall	CFT-Delay	**8-30 Days Post 1st ECT;** BL – Bifrontal (n = 66); BL - Bitemporal (n = 56); UL (n = 59)		
					Verbal Memory-	RAVLT	**8-30 Days Post 1st ECT;** BL – Bifrontal		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Immediate and Delayed Recall		(n = 65); BL - Bitemporal (n = 69); UL (n = 50)		
Levy R, 1968, UK	RCT	Depression (N = 40) Note: randomized, not recruited, sample	BL (n = 20)	UL – Non-Dominant (n = 20)	Subjective Memory	General Events Recent Personal Events	**Pre ECT;** BL (n = 20); UL (n = 20) **8-30 Days Post 1st ECT;** BL (n = 20); UL (n = 20)	Y	
					Verbal Memory-Immediate and Delayed Recall	Paired Associates Learning Test WMS			
Lisanby, 2000, USA	RCT	Major Depressive Disorder (N = 55)	BLTemporal – Just Above Seizure Threshold BLTemporal – 2.5x Seizure Threshold	Right UL – Temporal Parietal – Just Above Seizure Threshold Right UL – Temporal Parietal – 2.5x Seizure Threshold	Autobiographical Memory	AMI	**Pre ECT**; N= 52 **1 Week After Randomization;** N = 52 **8 Weeks After ECT**: N = 31	N	No information provided on number of participants per treatment condition No information on number of treatments given (ECT given 3 times/week) therefore follow up since 1st ECT
						PIMT	**Pre ECT**; N= 55 **1 Week After Randomization;** N = 55 **8 Weeks After ECT**: N = 33		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
									can not be determined
McElhiney, 1995, USA	RCT	Major Depressive Disorder (N = 75)	BLTemporal – Just Above Seizure Threshold BLTemporal – 2.5x Seizure Threshold	Right UL – Temporal Parietal – Just Above Seizure Threshold Right UL – Temporal Parietal – 2.5x Seizure Threshold	Autobiographical Memory	AMI	**Pre ECT;** N = 75 **8-30 Days Post 1st ECT;** BL-Low (n = 18); BL – High (n = 20); RUL – Low (n = 16); RUL – High (n = 18) **31-183 Days Post 1st ECT;** N = 45	N	No information provided on number of participants per treatment condition at baseline or 2 month follow up
McCall et al, 2002, USA	RCT	Major Depressive Disorder (N = 77) Note: randomized, not recruitment, sample	BLTemporal – 1.5x Seizure Threshold (n = 34)	Right UL – Temporal Parietal – 8.0x Seizure Threshold (n = 40)	Non Verbal Memory-Delayed Recall	RFT	**Pre ECT;** BL (n = 32); UL (n = 39) **8-30 Days Post 1st ECT;** BL (n = 36); UL (n = 34) **8-30 Days Post 1st ECT;** BL (n = 33); UL (n = 32) **31-183 Days Post 1st ECT;** BL (n = 23); UL (n = 40)	Y	
					Verbal Memory-	RAVLT	**Pre ECT;** BL (n = 35); UL (n = 39)		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Delayed Recall		**8-30 Days Post 1st ECT;** BL (n = 36); UL (n = 36) **8-30 Days Post 1st ECT;** BL (n = 35); UL (n = 35) **31-183 Days Post 1st ECT;** BL (n = 26); UL (n = 35)		
O'Connor M, 2008 USA	Cohort (prospective)	Major Depressive Disorder (N = 22)	UL BL (n = 11)	UL (n = 11)	Verbal Memory - Immediate and Delayed Recall	WMS (story)	**Pre ECT;** UL BL (n = 11); UL (n = 11) Can not determine follow up time/sample size with certainty	N	Authors report on average time to follow up (30.9 days from baseline)
Pettinatti HM, 1984 USA		Major Depressive Disorder (N = 28) Note: Could be either recruitment or randomization sample	BL – Bitemporal (n = 15)	UL – Non-Dominant – Temporal Parietal (n = 13)	Subjective Memory	Self-Rating Scale of Memory Function	**Pre ECT;** BL (n = 15); UL (n = 13) **Post 5th ECT;** BL (n = 15); UL (n = 13)	N	No information given on ECT scheduling; can not determine the number of days since the 1st ECT for follow up times
Prudic J, 1994, USA	RCT	Major Depressive	BLTemporal – Just Above	Right UL – Temporal	Autobiographical	Word Recall	**5 minutes After Orientation was**	N	No information

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
(Study 1)		Disorder (N = 52)	Seizure Threshold	Parietal – Just Above Seizure Threshold	Memory	Word Recall and Recognition	**Achieved;**		available on number of patients per treatment condition at randomizatio n or follow-up
Prudic J, 1994, USA (Study 2)	RCT	Major Depressive Disorder (N = 100)	BLTemporal – Just Above Seizure Threshold (n = 23) BLTemporal – 2.5x Seizure Threshold (n = 27)	Right UL – Temporal Parietal – Just Above Seizure Threshold (n = 23) Right UL – Temporal Parietal – 2.5x Seizure Threshold (n = 23)	Autobiographical Memory	Word Recall	**5 minutes After Orientation was Achieved;**	N	No information available on number of patients per treatment condition at follow-up
						Word Recall and Recognition			
					?	Verbal Fluency - Category			
					?	Verbal Fluency - Letter			
					?	Word Finding - Abstract			
					?	Word Finding - Concrete			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Ranjkesh F, 2005, Iran	RCT	Major Depressive Disorder (N = 45) Note: unipolar and bipolar included	BL – Bifrontal – 1.5x Seizure Threshold (n = 15) BL – Bitemporal – 1x Seizure Threshold (n = 15)	Right UL – Temporal Parietal – 5.0x Seizure Threshold (n = 15)	Global Cognitive Status	MMSE	**Pre ECT:** BL - Bifrontal (n = 13); BL - Bitemporal (n = 14); UL (n = 12) **8-30 Days Post 1st ECT;**	Y	No information available for number of subjects per treatment condition at 8-30 days post 1st ECT
Rosenberg J, 1984, USA	RCT	Major Affective Disorder (n = 34) Schizoafective Disorder (n = 1) Note: Could be either recruitment or randomization sample	BL – Bitemporal (n = 21)	Unilaeral – Non-Dominant – Temporal Parietal (n = 14)	Subjective Memory	0-1 Memory complaints 2-5 memory complaints	**8-30 Days Post 1st ECT;** **8-30 Days Post 1st ECT;**	Y	
Sackeim, 1993, USA	RCT	MDD	BL- low-Bifrontal-Temporal –	RUL-low-initial ST, (n=23),	Global Cognitive Status	MMSE 10-item orientation	**8-30 days**	N	66 participants were tested at 2 month

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
			Initial ST (n=23) BL- high-Bifrontal-Temporal – 2.5 ST (n=27) 3x/week (n=20)	RUL- high-2.5 ST (n=23) RUL-d'Elia 3x week		interview			follow-up; 70 participants werte followed up for a year Percentage change from baseline
					Non-Verbal Memory-Recognition	Shape and Face Recall and Recognition	**8-30 days** **31-183 days**		
					Verbal Memory - Recognition	Anterograde Paired Words	**8-30 days** **31-183 days**		
					Verbal Memory – Immediate Recall	Buschke Selective Reminding Test	**8-30 days**		
					Autobiographical Memory	AMI	**8-30 days**		
					Subjective Memory	SSMQ	**8-30 days**		
Sackeim, 2000, USA	RCT	MDD	BL- high-Bifrontal-Temporal – 150% initial ST 3x/week (n=20)	RUL-low-50% initial ST, (n=20), RUL-moderate - 150% initial ST (n=20)	Global Cognitive Status	MMSE	**8-30 days:** (n=80) **31-183 days:** (n=80)	N	Values are the percentage of items not recalled or recognized during postictal assessment that were recalled or
					Non-Verbal Memory-Immediate Recall	Rey-Osterrieth Complex Figure	**8-30 days:** (n=80) **31-183 days:** (n=80)		
					Verbal Memory - Recognitio	Anterograde Paired Words	Immediate recognition,		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
				RUL- high-500% initial ST (n=20) RUL-d'Elia 3x week	n		Delayed recognition – 4hrs post-ECT **8-30 days:** (n=80) **31-183 days:** (n=80)		recognized prior to the treatment.
					Verbal Memory – Immediate Recall	Buschke Selective Reminding Test (	**8-30 days:** (n=80) **31-183 days:** (n=80)		
					Autobiographical Memory	AMI	**8-30 days:** (n=80)		
					Subjective Memory	Randt Memory Test SSMQ	**8-30 days:** (n=80)		
Sackeim, 2007, USA	Cohort (prospective)	Recruited (N = 398) Major Depressive Disorder (sample size given for ITT sample	BL	Right UL	Attention	Stroop Color Word Test	**Pre ECT;** **1-7 Days Post 1st ECT:** N = 199 **6 month follow – up (Can not determine exact time of follow up);** N = 162	N	All values for participant count per cognitive measure given as total patients completed across treatment

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		only) Schozioaffective Disorder – Depressed (sample size given for ITT sample only)				CPT	**Pre ECT;** N = 242 **1-7 Days Post 1st ECT:** N = 167 **6 month follow – up (Can not determine exact time of follow up);** N = 153		conditions (i.e. BL and UL No information given on number of subjects randomized or recruited to each treatment condition
					Autobiographical Memory	AMI – Short Form	**Pre ECT;** N = 242 **1-7 Days Post 1st ECT:** N = 243 **6 month follow – up (Can not determine exact time of follow up);** N = 186		
					Global Cognitive Status	MMSE	**Pre ECT;** N = 347 **1-7 Days Post 1st ECT:** N = 257 **6 month follow – up (Can not determine exact time of follow up);** N = 191		
					Motor	Simple – RT	**Pre ECT;** **1-7 Days Post 1st**		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
							ECT: N = 211 6 month follow – up (Can not determine exact time of follow up); N = 152		
						Choice – RT	Pre ECT; 1-7 Days Post 1st ECT: N = 227 6 month follow – up (Can not determine exact time of follow up); N = 169		
						Stroop Color Word Test - RT	Pre ECT; 1-7 Days Post 1st ECT: N = 199 6 month follow – up (Can not determine exact time of follow up); N = 162		
					Non Verbal Memory-Delayed Recall	CFT - Copy	Pre ECT; 1-7 Days Post 1st ECT: N = 218 6 month follow –		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
							up (Can not determine exact time of follow up); N = 148		
						CFT - Reproduction	**Pre ECT;** **1-7 Days Post 1st ECT:** N = 203 **6 month follow – up (Can not determine exact time of follow up);** N = 139		
						BSRT - Learn	**Pre ECT;** **1-7 Days Post 1st ECT:** N = 236 **6 month follow – up (Can not determine exact time of follow up);** N = 175		
						BSRT - Delay	**Pre ECT;** **1-7 Days Post 1st ECT:** N = 225 **6 month follow – up (Can not determine exact time of follow up);**		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
							N = 171		
Sand Stromgren L, 1975, Denmark	RCT	Endogenous Depression (N = 100) Note: randomized, not recruitment, sample	BL (n = 48)	UL – Non-Dominant (n = 52)	Subjective Memory	Memory Quotient (Post 6 Tx) Memory Quotient(Post 12 Tx)	**8-30 Days Post 1st ECT;** BL (n = 48); UL (n = 52)	Y	
Schat et al, 2007, Netherlands	Cohort (prospective)	Recruited (N = 104) Unipolar Depression (N = 96)	BL - 1.5x Seizure Threshold (n = 79)	Right UL – 2.5x Seizure Threshold (n = 17)	Autobiographical Memory	Groninger Intelligence Test - Word Fluency Test 1 Groninger Intelligence Test - Word Fluency Test 2	**Pre ECT;** BL (n = 67); UL (n = 16) **8-30 Days Post 1st ECT;** **31-183 Days Post 1st ECT;** **184-365 days Post 1st ECT;**	Y	No information available for number of subjects per treatment condition at 8-30 Days Post 1st ECT, 31-183 Days Post 1st ECT, or 184-365 days Post 1st ECT.
					Intellectual Ability	Rivermead Behavioural Memory Test	**Pre ECT;** BL (n = 64); UL (n = 16) **8-30 Days Post 1st ECT;** **31-183 Days Post 1st ECT;** **184-365 days Post 1st ECT;**		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Sienaert P, 2009, Belgium	RCT	Major Depressive Disorder (N = 81) Note: Randomization, not recruitment, sample; Bipolar and unipolar depression, with and without psychotic features included	BL – Bifrontal – 1.5x Seizure Threshold (n = 41)	Right UL – Temporal Parietal – 6x Seizure Threshold (n = 40)	Global Cognitive Status	MMSE	**Pre ECT;** BL (n = 32); UL (n = 32) **8-30 Days Post 1st ECT;** BL (n = 32); UL (n = 32) **31-183 Days Post 1st ECT;** BL (n = 32); UL (n = 32)	Y	Only data pertaining to participants that completed the study are available
Sienaert, 2010, Belgium	RCT	Major Depressive Disorder (N = 81) Note: unipolar and bipolar, with and without psychotic symptoms	BLTemporal – 1.5x Seizure Threshold (n = 41)	UL – 6x Seizure Threshold (n = 40)	Attention	CPT	**Pre ECT;** **31-183 Days Post 1st ECT (1 week post treatment);** **31-183 Days Post 1st ECT (6 weeks post treatment);**	N	Degrees of freedom vary from test to test Number of subjects per treatment condition at Pre ECT, 31-183 Days Post 1st ECT (1 week post treatment), or
						Trail Making Test - A			
					Autobiographical Memory	AMI			
					Executive Functioning	Trail Making Test- B			
						Wisconsin Card Sorting			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
						Task - Categories			31-183 Days Post 1st ECT(6 weeks post treatment) can not be determined with certainty.
					Global Cognitive Status	MMSE			
					Subjective memory	SSEMF			
					Verbal Memory-Immediate and Delayed Recall	RAVLT			
					Working Memory	Letter Number Sequencing Test			
Sobin C., 1995, USA	RCT	Major Depressive Disorder with psychotic features (n = 32) Major Depressive Disorder without psychotic features (n = 15)	BLTemporal – Just Above Seizure Threshold BLTemporal – 2.5x Seizure Threshold	Right UL– Just Above Seizure Threshold Right UL– 2.5x Seizure Threshold	Autobiographical Memory	AMI	**Pre ECT;** 71/71 patients **31-184 Days Post 1st ECT (1 week post randomized phase;** 45/71 patients **31-184 Days Post 1st ECT(8 weeks post treatment completion);** Can not determine with certainty	N	Number of participants randomized to each treatment condition not given

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		Bipolar Disorder (n = 24)			Global Cognitive Status	MMSE	**Pre ECT;** 67/71 patients **31-184 Days Post 1st ECT (1 week post randomized phase;** 67/71 patients **31-184 Days Post 1st ECT(8 weeks post treatment completion);** 44/45 patients		
Squire LR, 1975, USA	Cohort (retrospective)	Depressive Neurosis (n=19) Primary Affective Disorder (n=9) Involutional Melancholia (n=4) Schizophrenic Depressive	BL – Temporal Parietal (n = 16)	Right UL – Temporal Parietal (n = 12) Note: Number that agreed to participate No ECT (n = 14)	Subjective Memory	SSEMF	**184-365 days Post 1st ECT;** BL (n = 16); UL (n = 10)	Y	Can not determine follow up periods

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		(n=3) Manic Depressive (n=3)							
Squire, 1983, USA	Cohort (prospective)	Primary Affective Disorder/Major Depressive Disorder (n = 50) Manic Depressive Illness (depressed phase) (n = 16) Depressive Neurosis (n 8) Schizoaffective Disorder (n = 3) Unspecified Personality Disorder (n = 1)	BL – Bitemporal (n = 35)	Right UL (n = 28) Depressed Controls (n = 19)	Subjective Memory	Memory Self Rating Scale	**Pre ECT;** Can not determine with certainty **8-30 Days Post 1st ECT;** Can not determine with certainty **184-365 Days Post 1st ECT;** Can not determine with certainty **1 Year Post 1st ECT:** BL (n = 31); UL (n = 28)	N	Sample Size for follow up pertains to 3 year follow up exclusively

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Stoppe A, 2006, Brazil	RCT	Major Depressive Disorder (N = 39) Note: Randomization, not recruitment, sample	BL (n = 22)	Right UL – 5-6x Seizure Threshold (n = 17)	Attention	WAIS-Revised - Digits Forward	**0-183 Days Post 1st ECT;** BL (n = 14); UL (n = 14)	Y	
					Executive Functioning	WAIS-Revised - Digits Backwards	**0-183 Days Post 1st ECT;** BL (n = 14); UL (n = 14)		
						OM-Animal Verbal Fluency	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
					Global Cognitive Status	BNT - Correct with Phonemic Clues	**0-183 Days Post 1st ECT;** BL (n = 14); UL (n = 14)		
						BNT - Correct with Visual Cues			
						BNT - Naming			
						BNT - Number of Phonemic Clues			
						BNT - Number of Visual Cues			
						BNT - Paragnosias			
						OM - Evocation	**0-183 Days Post 1st ECT;** BL (n =		

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
						OM - Intrusions	13); UL (n = 13)		
						OM - Storage			
						MMSE	**0-183 Days Post 1st ECT;** BL (n = 12); UL (n = 14)		
					Intellectual Ability	WAIS-Revised - Vocabulary	**0-183 Days Post 1st ECT;** BL (n = 12); UL (n = 12)		
					Motor	WAIS-Revised - Block Design	**0-183 Days Post 1st ECT;** BL (n = 12); UL (n = 12)		
						Clock Drawing	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
					Subjective Memory	Autobiographic memory (does not recall)	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
						Autobiographic memory (partial recall)			
						Autobiographic memory (question does not apply)			
						Autobiographic memory			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
						(recalls)			
						RBMT - Classification	**0-183 Days Post 1st ECT;** BL (n = 12); UL (n = 12)		
						RBMT - Standard			
					Verbal Memory – Delayed Recall	OM - Delayed Recall	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
					Verbal Memory – Immediate Recall	OM - Consistent Recall	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
					Verbal Memory - Recognition	OM - Recognition	**0-183 Days Post 1st ECT;** BL (n = 13); UL (n = 13)		
Strain, J.J, 1968, USA	Cohort (prospective)	Recruited (N = 102) Manic-Depressive Depressed (n = 42) Neurotic Depressive (n = 39) Psychotic Depressive (n = 11)	BL – Temporal Parietal (n = 46)	Right UL – Temporal Parietal (n = 50)	Autobiographical Memory	Distribution of Post-ECT Recent Memory Test Scores of Individual Patients	**8-30 Days Post 1st ECT;** 4 ECT Treatments: BL (n = 5); UL (n = 2) 6 ECT Treatments: BL (n = 13); UL (n = 12) 8 ECT Treatments: BL (n = 17); UL (n =	Y	

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
		Involutional Depressive (n = 4)					13) <u>10 ECT Treatments:</u> BL (n = 5); UL (n = 9) <u>12 ECT Treatments:</u> BL (n = 2); UL (n = 9)		
						Personal Data Sheet - Recent Memory	**0-30 Days Post 1st ECT;** BL (n = 42); UL (n = 45) **1-183 Days Post 1st ECT;** BL (n = 42); UL (n = 45)		
						Personal Data Sheet – Remote Memory			
					Non-Verbal Memory- Delayed Recall	BVRT	**0-30 Days Post 1st ECT;** BL (n = 42); UL (n = 45) **1-183 Days Post 1st ECT;** BL (n = 42); UL (n = 45)		
					Verbal Memory - Immediate Recall	PAL			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Taylor & Abrams, 1985, Britain	RCT	Melancholia (N = 37) Note: Pertains to the number of patients that completed all testing phases	BLTemporal	Right UL – Non-Dominant	Global Cognitive Status	MMSE	**Pre ECT;** Can not determine with certainty **8-30 Days Post 1st ECT;** BL (n = 15); UL (n = 22)	Y	Sample size makes reference to the number of patients who completed all testing phases
Tew JD, 2002, USA	RCT	Major Depressive Episode (N = 24)	BL - 2.5x Seizure Threshold (n = 11)	Right UL - 5.5x Seizure Threshold (n = 13)	Global Cognitive Status	MMSE	**Pre ECT;** BL (n = 8); UL (n = 7) **8-30 Days Post 1st ECT;** BL (n = 8); UL (n = 7)	Y	
Weeks D, 1980, Edinburgh	Case-Control	Major Depressive Disorder (N = 51)	BL - Bitemporal (n = 36)	UL – Non-Dominant – Temporal Parietal (n = 15)	Non-Verbal Memory - Delayed Recall	Delayed Recall	**8-30 Days Post 1st ECT;** BL (n =); UL (n =)	Y	
					Non-Verbal Memory- Immediate Recall	PAL - Visual Design			
					Verbal Memory - Immediate Recall	PAL - Auditory Verbal			

Note: RCT – Randomized Control Trial; AMI – Autobiographical Memory Index; MMSE1 – Modified Mental State Examination; SMCQ - Squire Memory Complaint Questionnaire; GSE-my – Global Self Evaluation-Memory; RFT – Rey Figure Test; WAIS - Weshler Adult Intelligence Scale; COWAT - Controlled Oral Word Association Test; D-KEFS - Delis-Kaplan Executive Function System Sorting Test; WMS – Weshler Memory Scale;

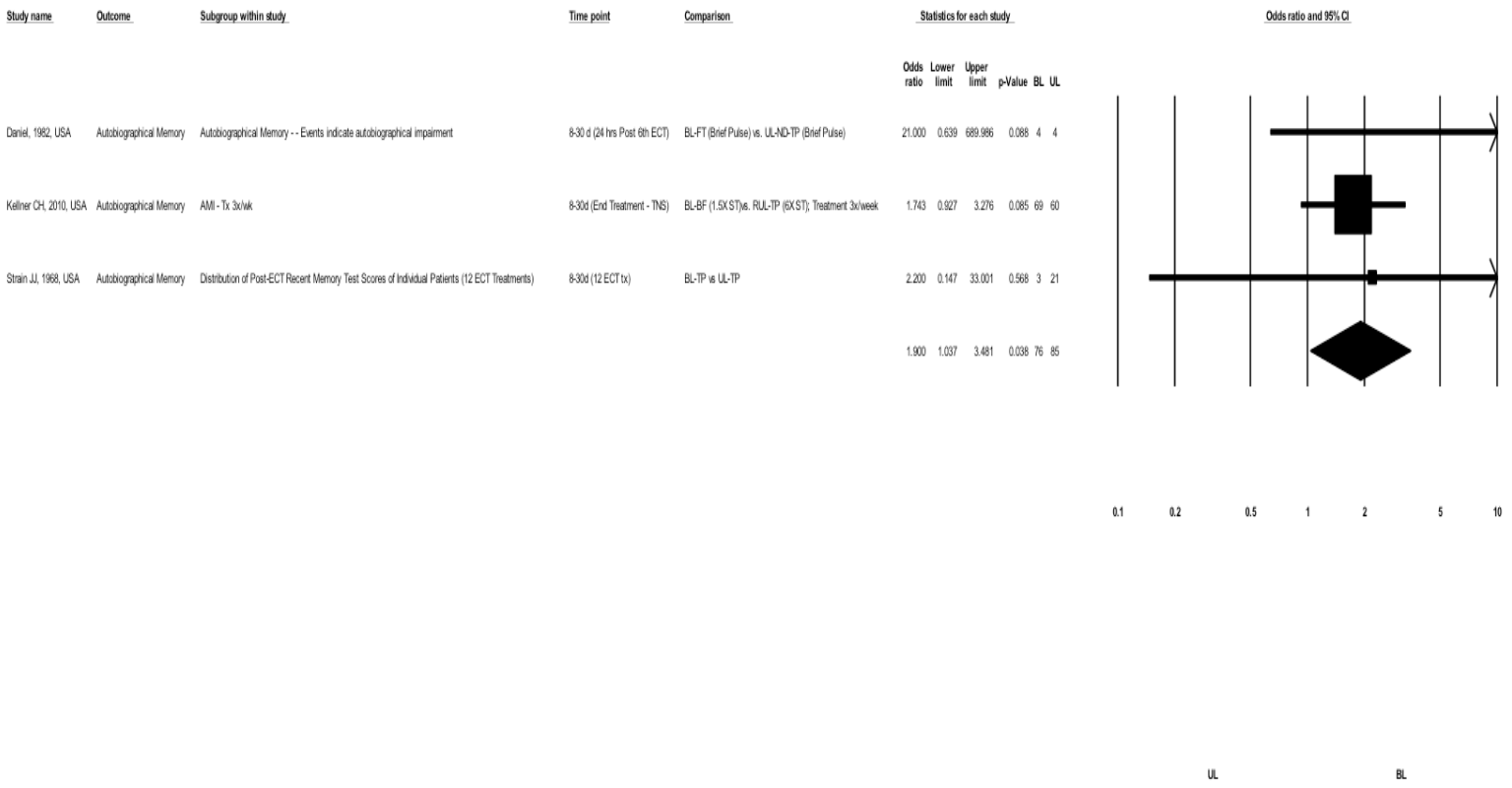
CGI Severity Scale – Clinical Global Impression Severity Scale; CFT-Delay – Cognitive Function Test-Delay; RAVLT – Rey Auditory Verbal Learning Test; CPT - Continuous Performance Task; SSEMF - Subjective Self-Evaluation of Memory Function; BNT – Boston Naming Test; OM - Fluid Object Memory Evaluation; PAL - Paired Associates Learning Test; BVRT - Benton Visual Retention Test; RBMT - Rivermead Behavior Memory Test; CFQ - Cognitive Failures Questionnaire; BSRT – Bushke Selective Reminding Test; PIMT - Personal and Impersonal Memory Test

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
Fleminger J.J., 1970, UK	RCT	Major Depressive Disorder (n =)	BL (n = 12) Left UL (n = 12)	Right UL (n = 12)	Verbal Memory-Immediate and Delayed Recall	WMS - Associate Learning Test	**8-30 Days Post 1st ECT;**	Y	
Loo, 2008, Australia	Cohort (prospective)	Major Depressive Disorder (n = 58) Bipolar Disorder (n = 36)	Right UL – Ultra Brief Pulse Width – 6.0x Seizure Threshold (n = 72)	Right UL – 6.0x Seizure Threshold (n = 22)	Non Verbal Memory-Delayed Recall	Rey Figure	**Pre ECT:** UL (n = 22); UL Ultra Brief Pulse Width (n = 72) **8-30 Days Post 1st ECT (after 6 treatments);** UL (n = 18); UL Ultra Brief Pulse Width (n = 59) **8-30 Days Post 1st ECT (end of ECT);** UL (n =	Y	
						CFT			
					Verbal Memory-Delayed Recall	RAVLT			
					Executive Functioning	COWAT			
						Digits Forward			
						Stroop Color Word Test			

Author, year, country	Study design	Diagnosis	Intervention	Comparator	Cognitive Domain	Cognitive Tests	Participants (n) at different follow-up times	Data entered in meta-analysis	Comments
					Autobiographical Memory	AMI	16); UL Ultra Brief Pulse Width (n = 59)		
						AMI – Short Form	**Pre ECT:** UL (n = 5); UL Ultra Brief Pulse Width (n = 27) **8-30 Days Post 1st ECT (after 6 treatments);** UL (n = 3); UL Ultra Brief Pulse Width (n = 27) **8-30 Days Post 1st ECT (end of ECT);** UL (n = 3); UL Ultra Brief Pulse Width (n = 27)		

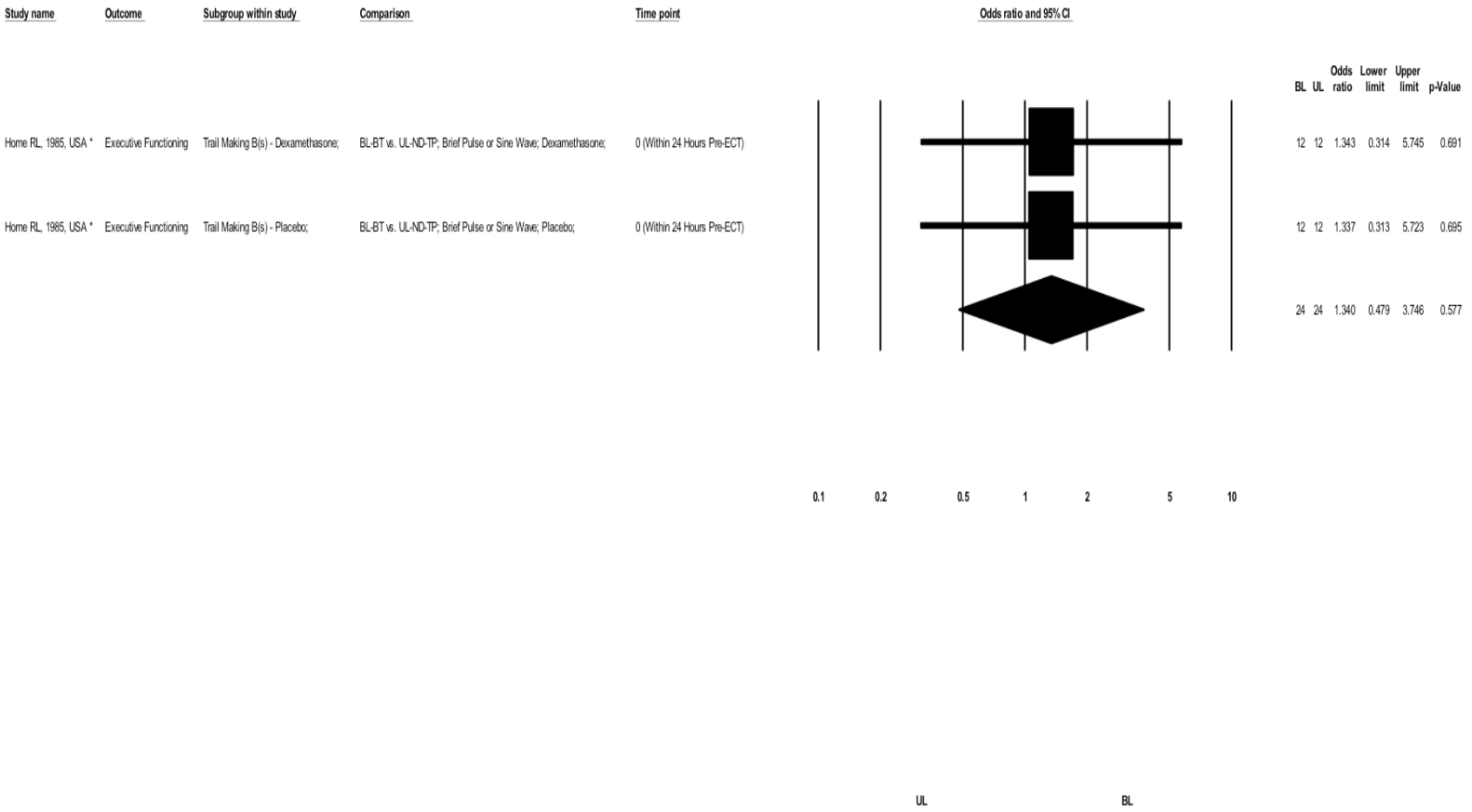
Forest Plots

Figure 3. Autobiographical Memory – 8 to 30 days post-ECT



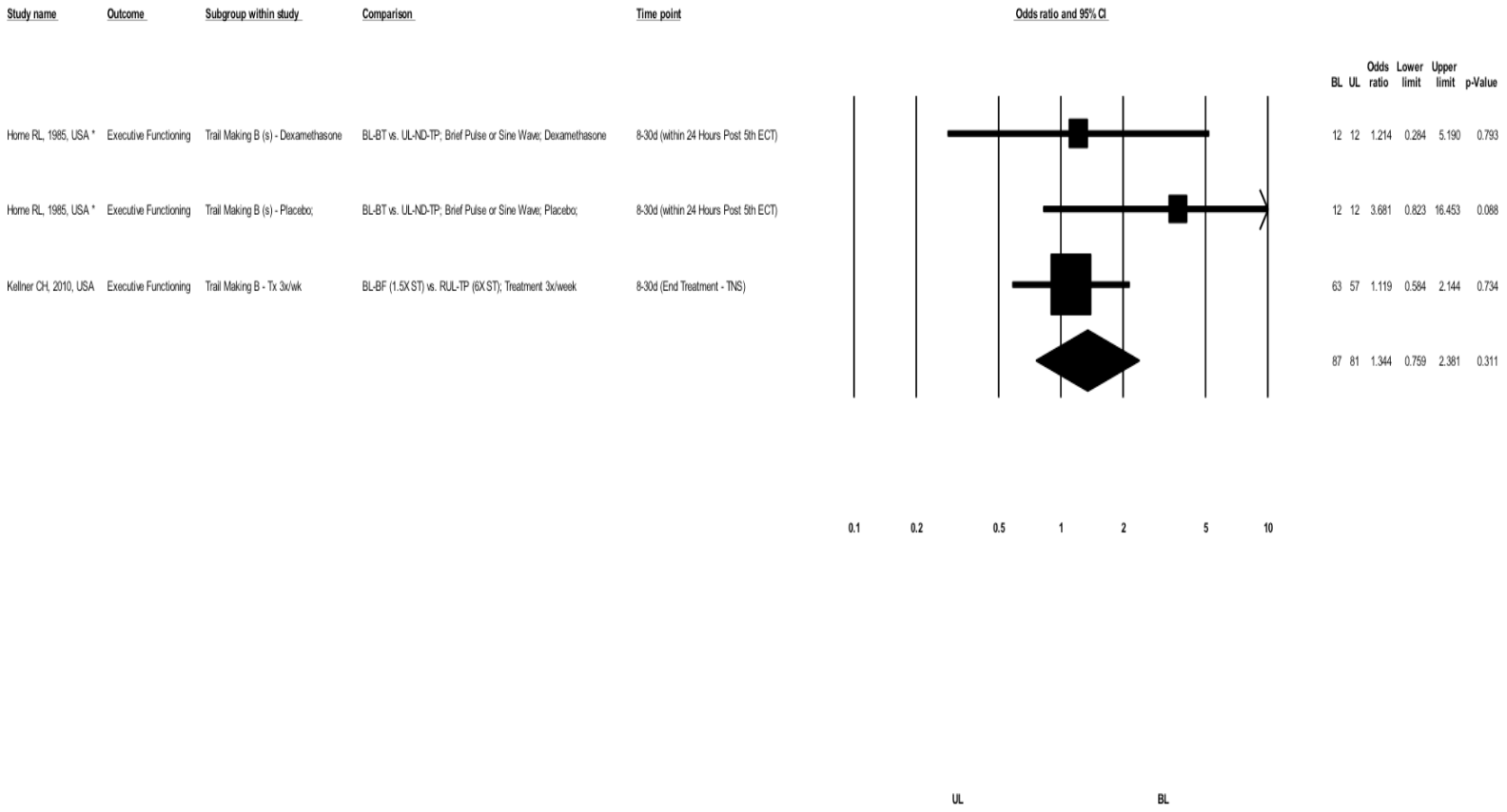
Q=1.92, df=2, p=0.386, I^2=0

Figure 4. Executive Function – pre-ECT



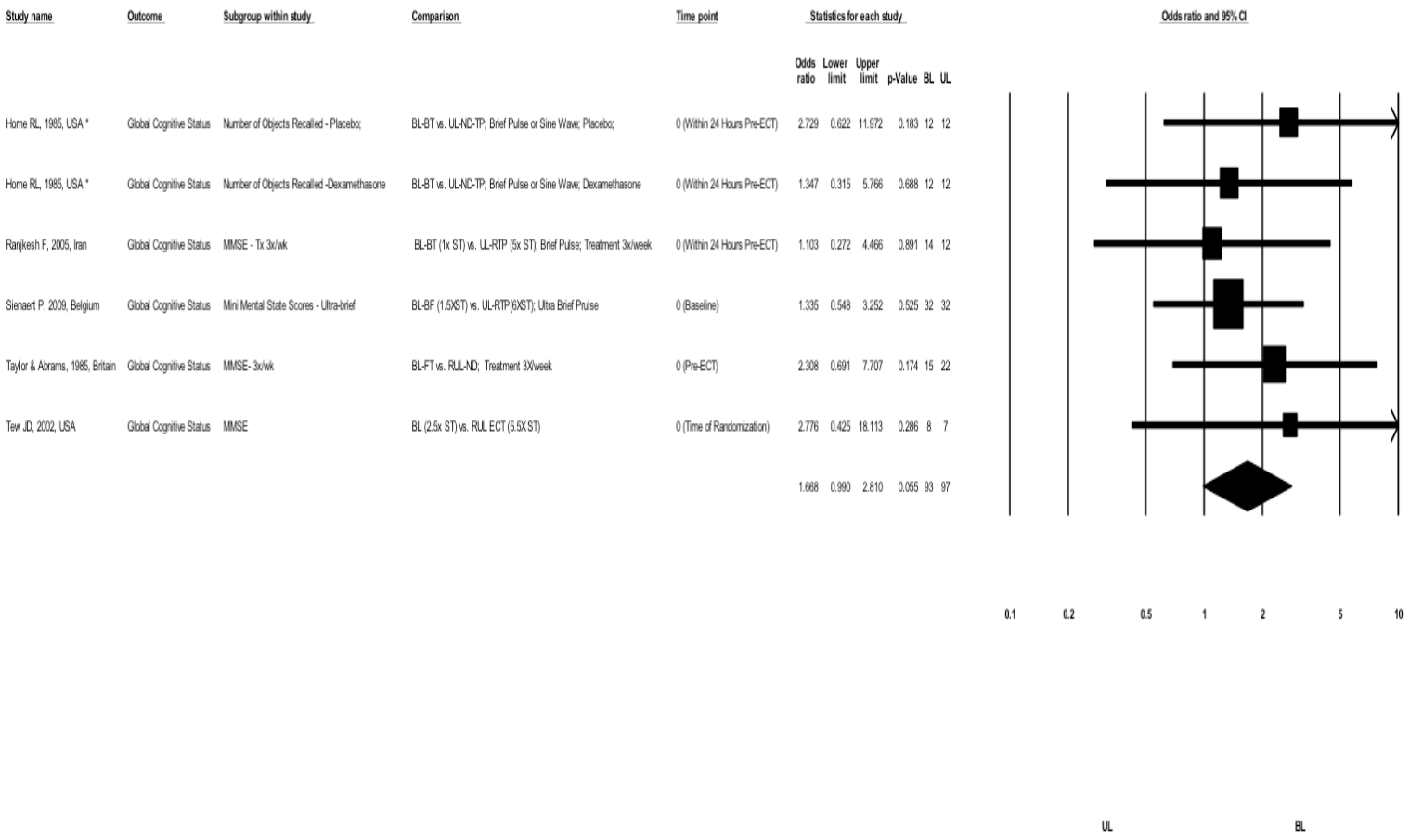
Q-value=0, df=1, p=0.997, I^2=0

Figure 5. Executive Function – 8 to 30 days post-ECT



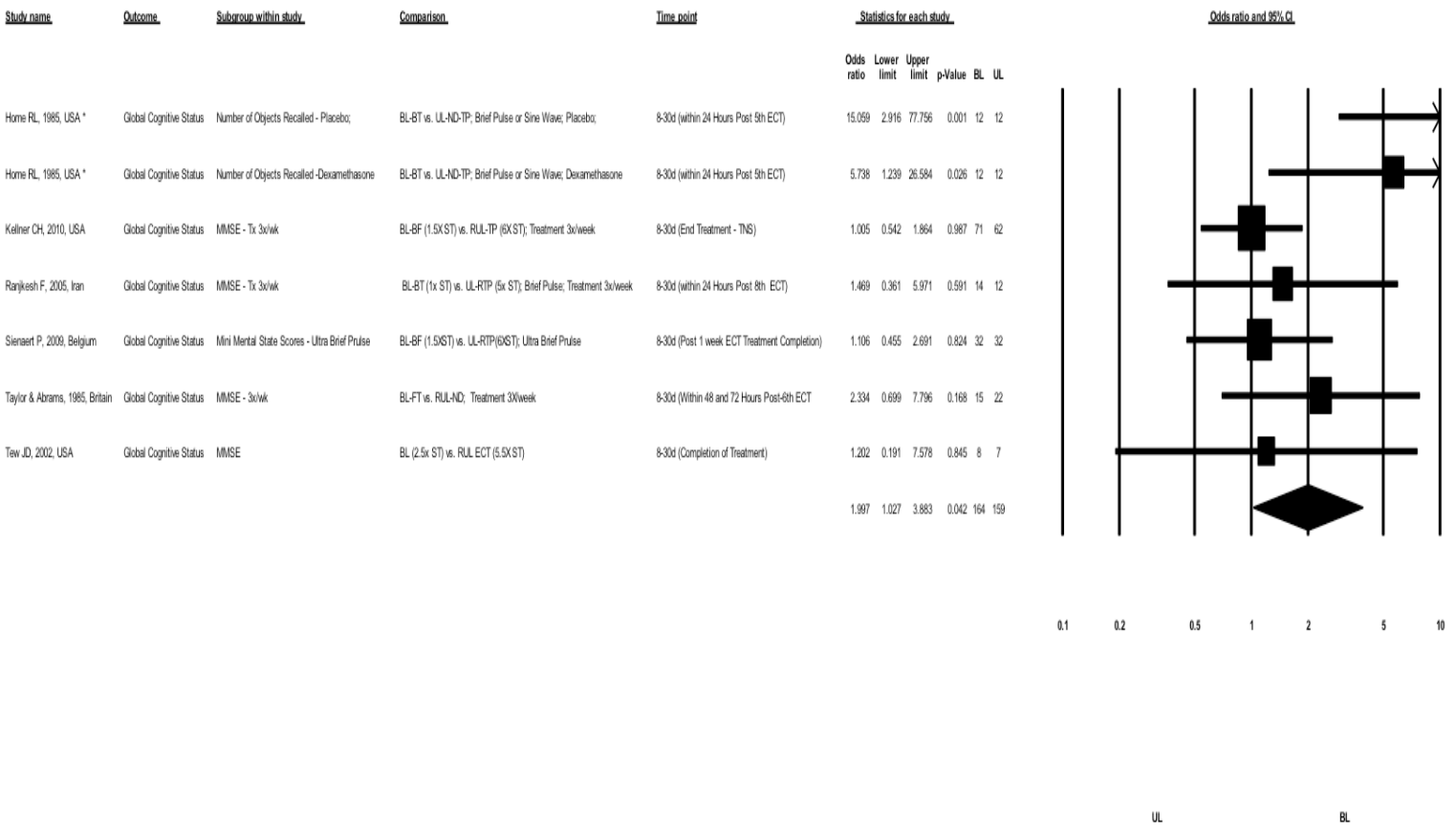
Q-value=2.061, df=2, p=0.357, I^2=2.95

Figure 6. Global Cognitive Status – pre-ECT



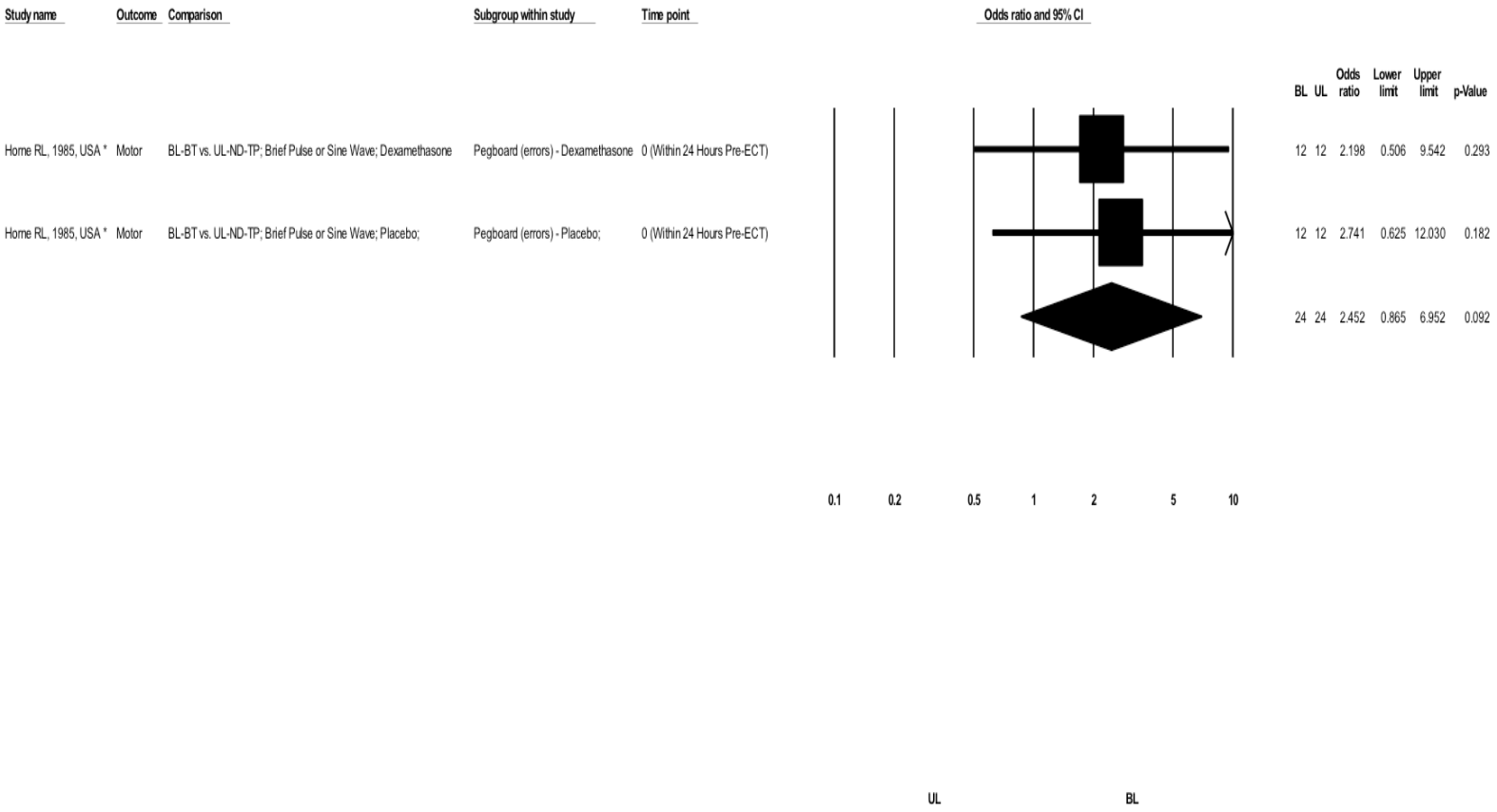
Q-value=1.647, df=5, p=0.896, I^2=0

Figure 7. Global Cognitive Status – 8 to 30 days post-ECT



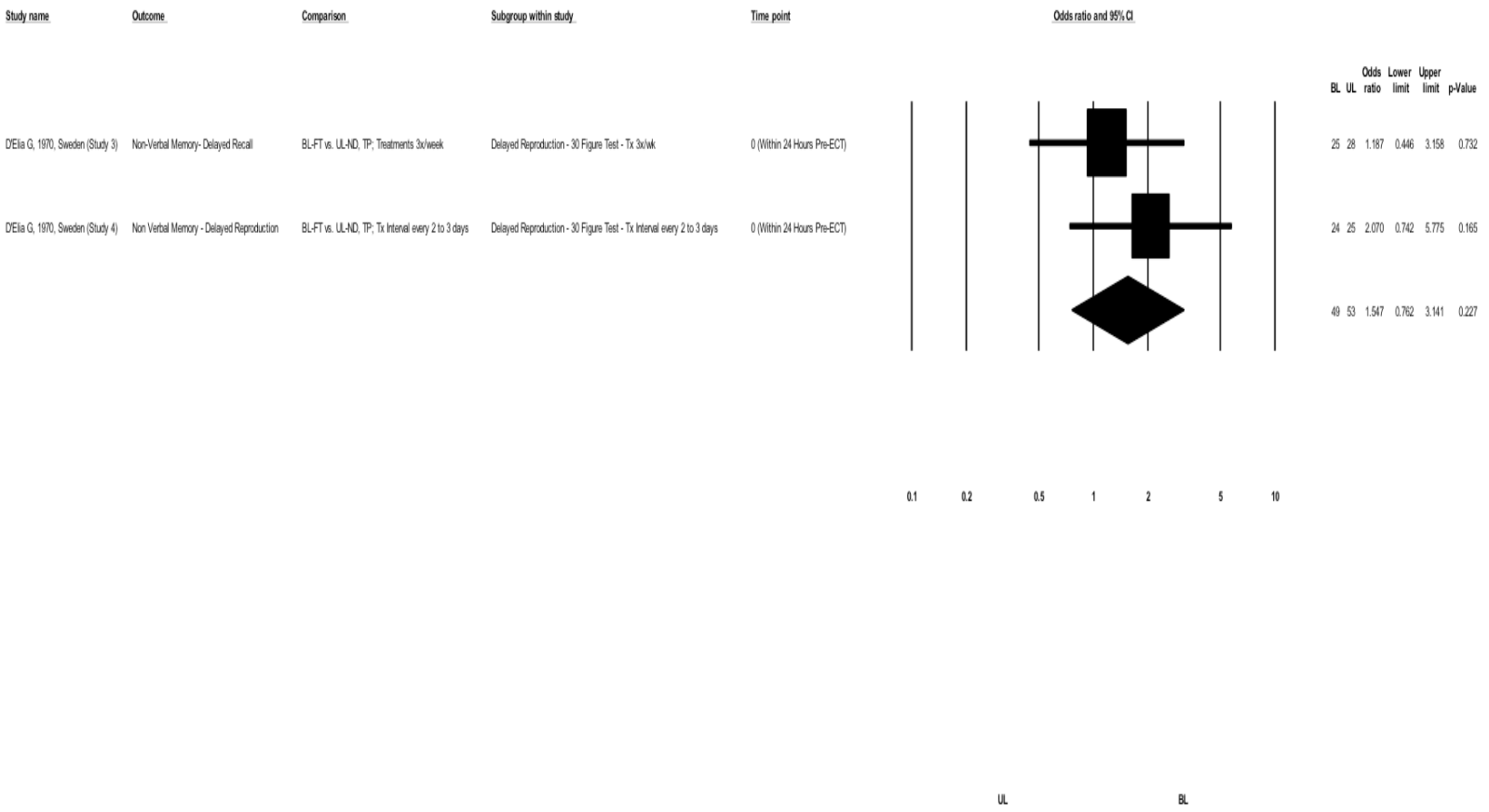
Q-value=13.131, df=6, p=0.041, I^2=54.307

Figure 8. Motor – pre-ECT



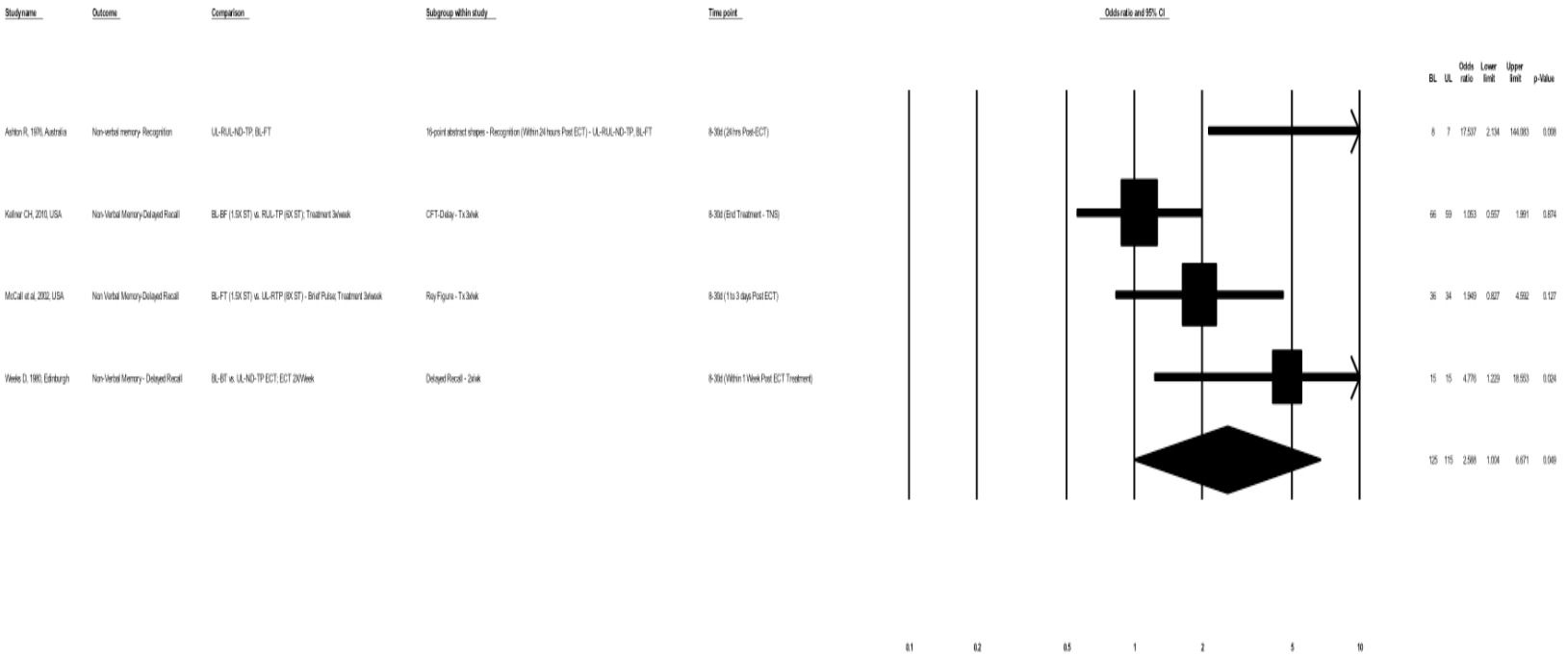
Q-value=0.043, df=1, p=0.835, I^2=0

Figure 9.Non-verbal Memory, Delayed Recall – pre-ECT



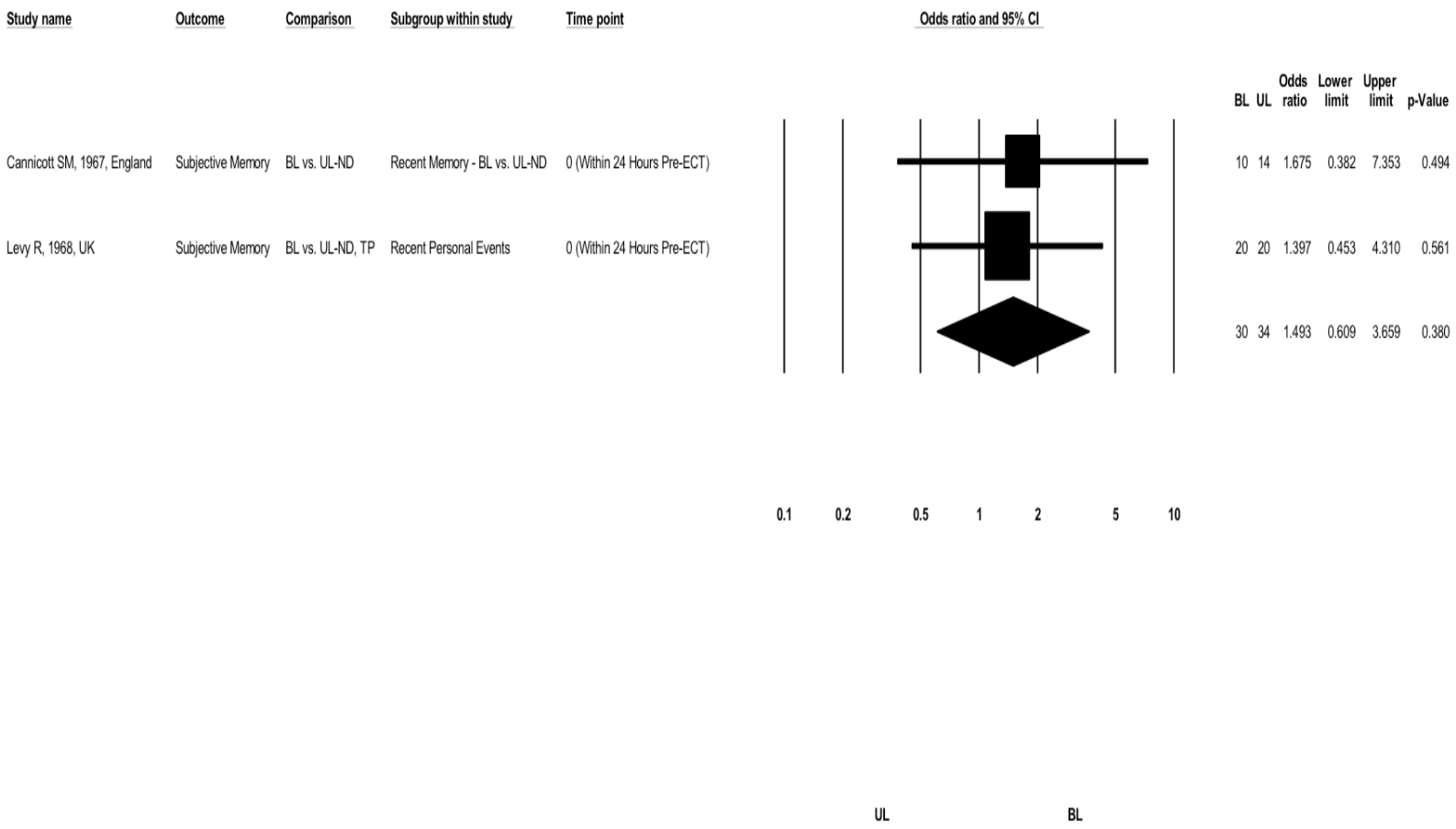
Q-value=0.591, df=1, p-value=0.442, I^2=0

Non-verbal Memory, Delayed Recall – 8 to 30 days post-ECT



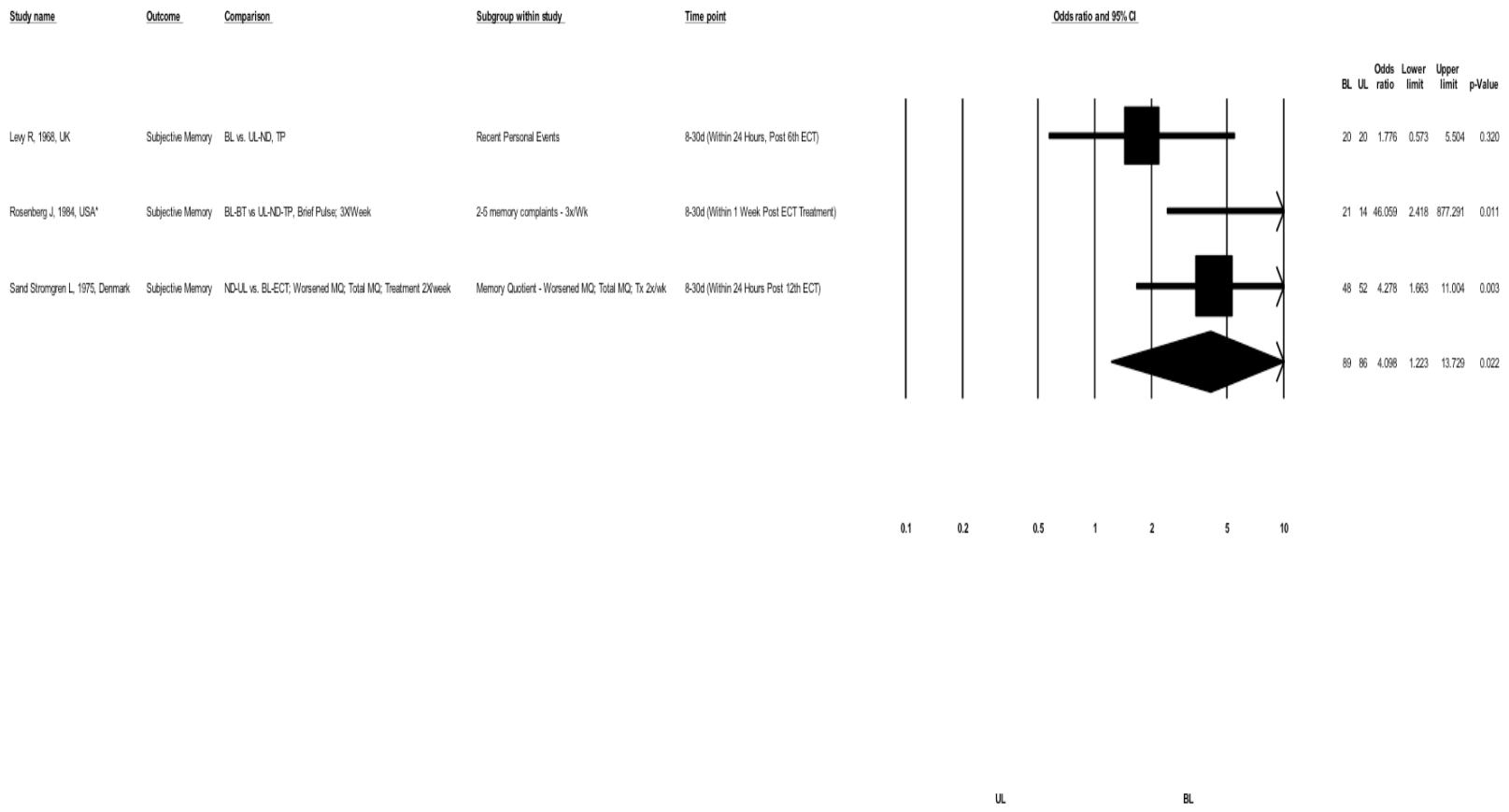
Q-value=9.205, df=3, p-value=0.027, I^2=67.409

Figure 10. Subjective Memory – pre-ECT



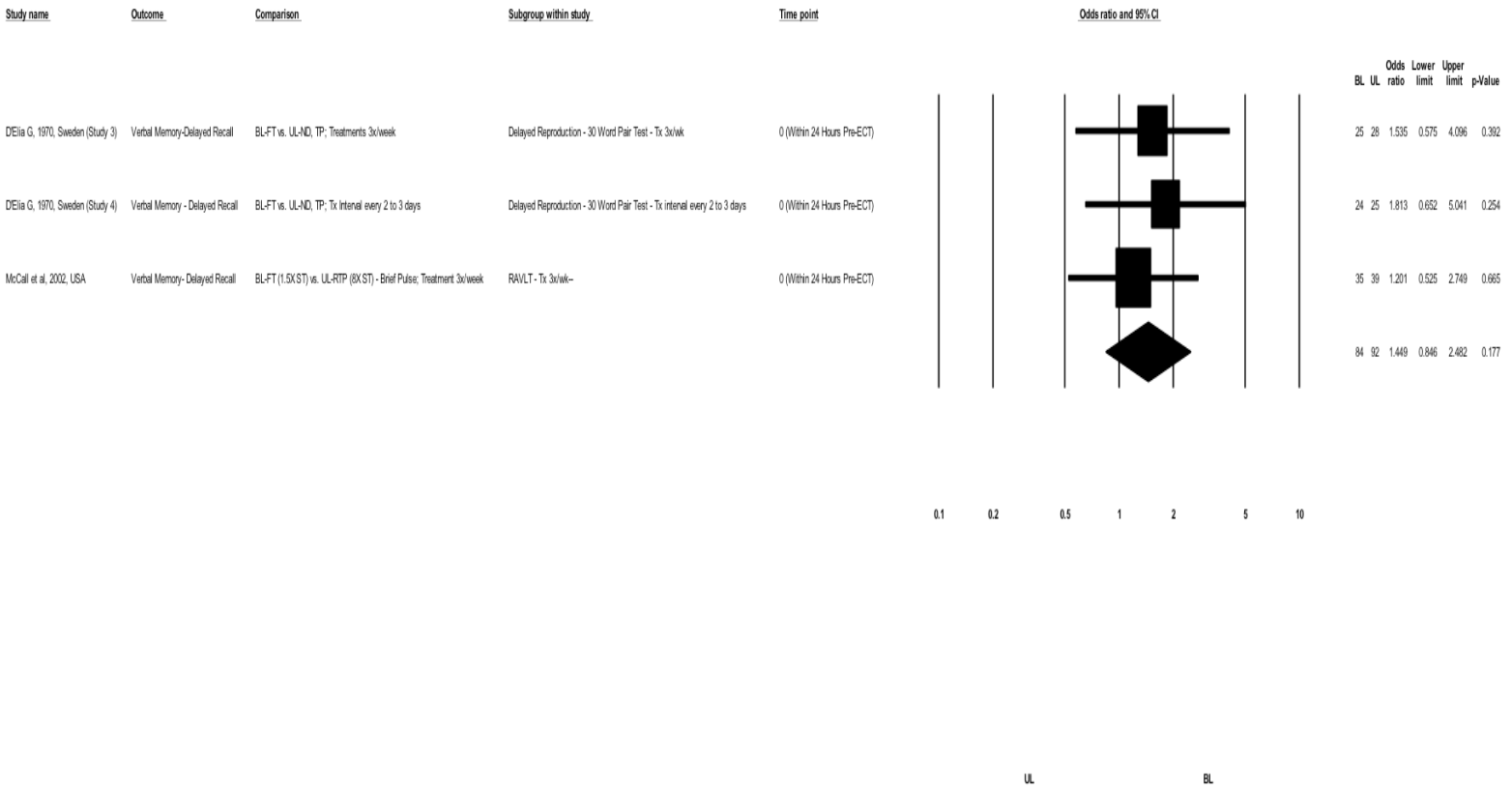
Q-value=0.037, df=1, p-value=0.848, I^2=0

Figure 11. Subjective Memory – 8 to 30 days post-ECT



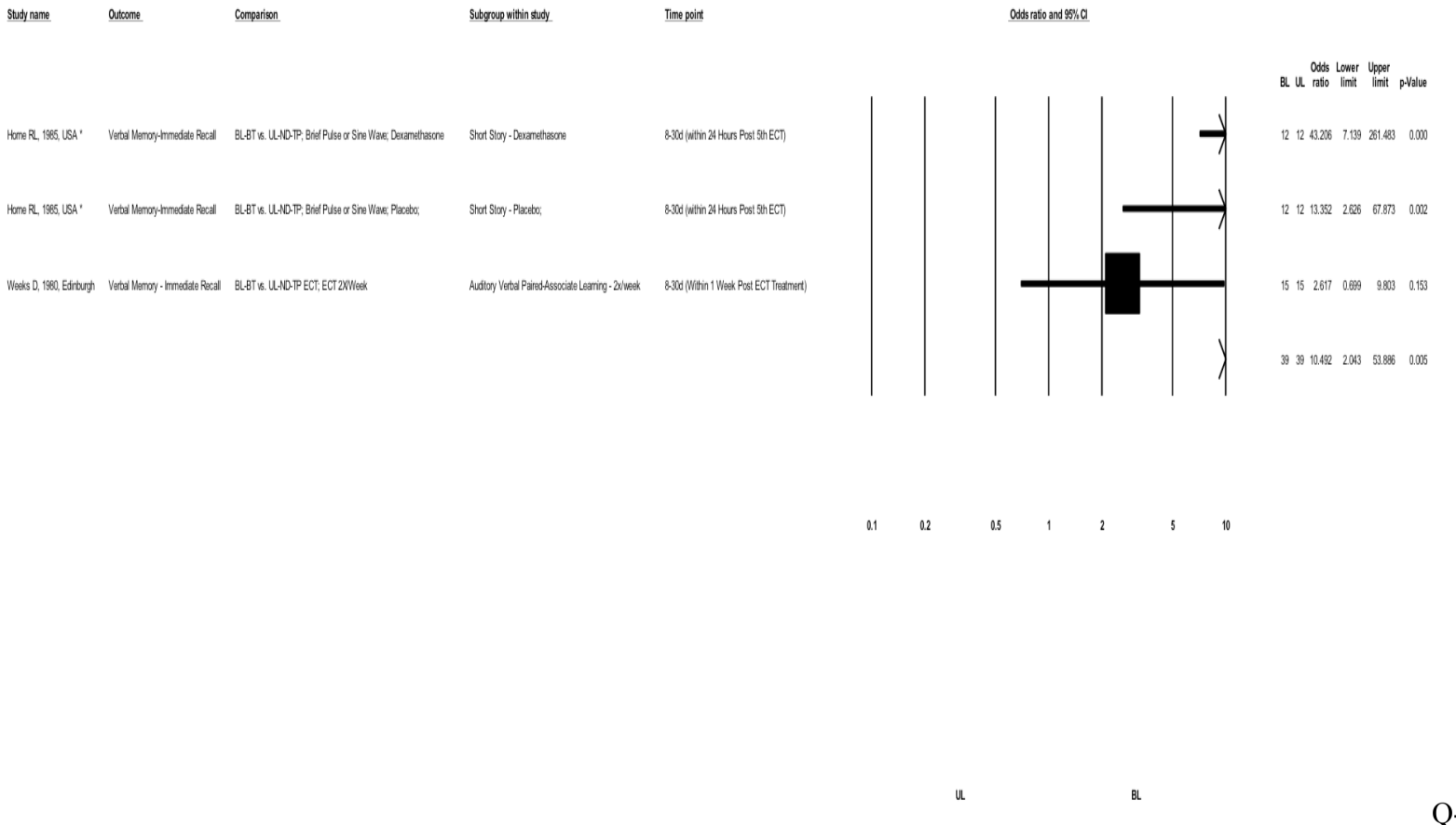
Q-value=4.49, df=2, p-value=0.106, I2=55.490

Figure 12. Verbal Memory, Delayed Recall – pre-ECT



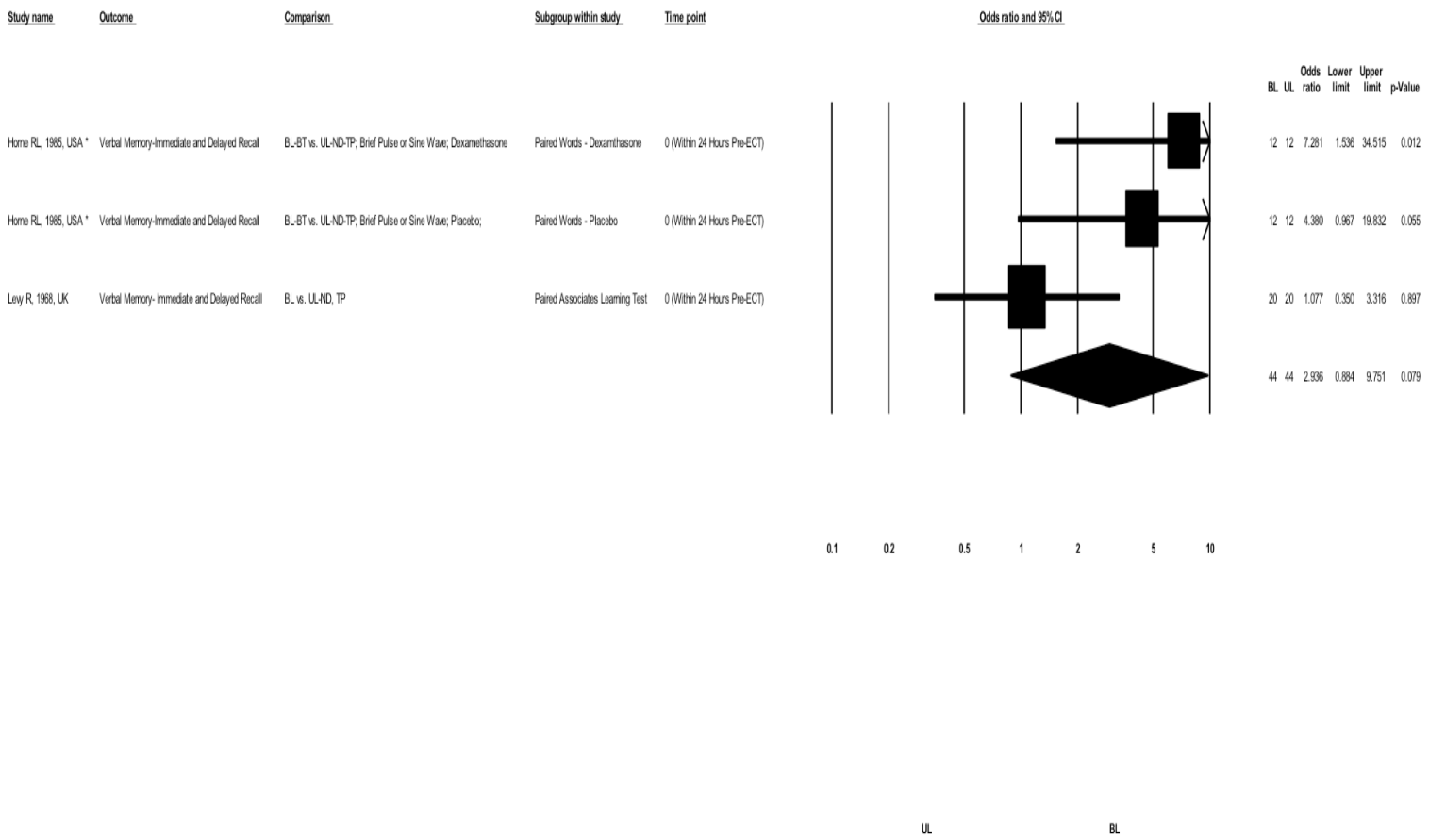
Q-value:0.395, df=2, p-value=0.821 I^2=0

Figure 13. Verbal Memory, Immediate Recall – 80 to 30 days post-ECT



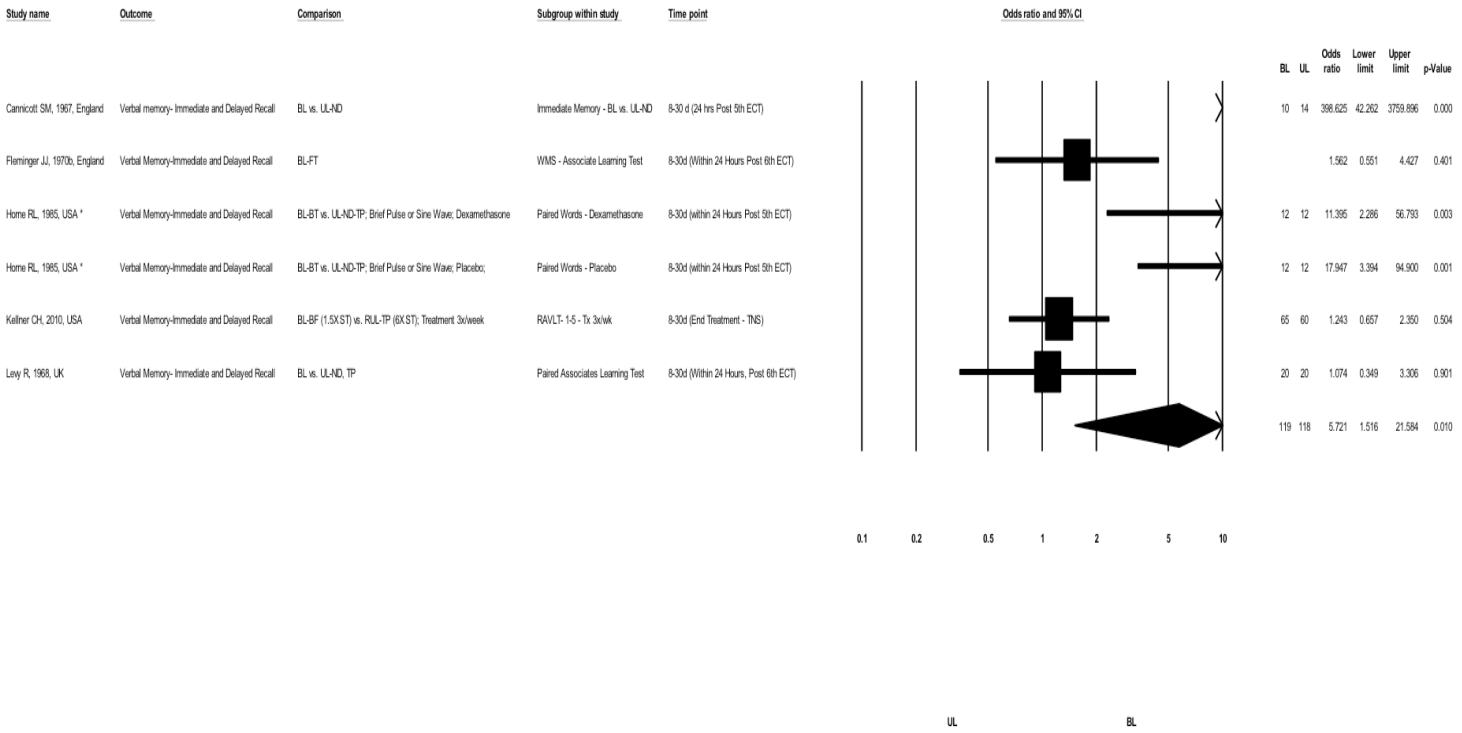
Q-value=6.485, df=2, p-value=0.039, I^2=69.16

Figure 14. Verbal Memory, Immediate and Delayed Recall – pre-ECT



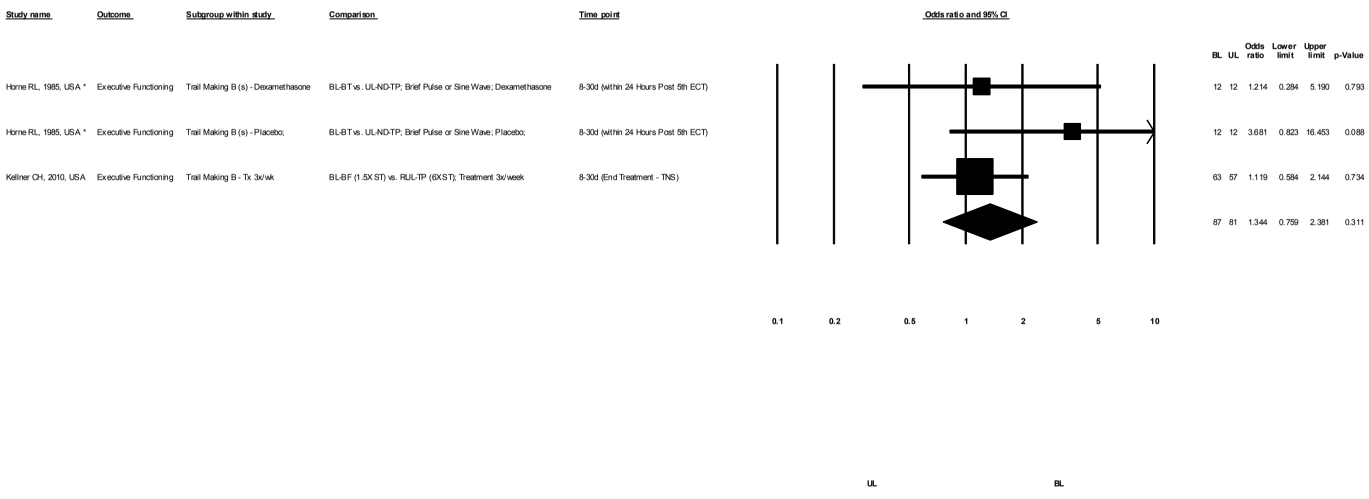
Q-value=4.495, df=2, p-value=0.106, I^2=55.510

Figure 15Verbal Memory, Immediate and Delayed Recall – 8 to 30 days post-ECT



Q-value=35.80, df=5, p<0.001, I^2=86.03

Figure 16. Executive Function



8 to 30 days

Q-value=2.061, df=2, p=0.357, I2=2.95

References

Abrams, R. (2007). Does bilateral ECT cause persistent cognitive impairment? *The Journal of ECT*, *23*(2), 61–62. http://doi.org/10.1097/yct.0b013e3180544997

Ashton, R., & Hess, N. (1976). Amnesia for random shapes following unilateral and bilateral electronvulsive shock therapy. *Perceptual and Motor Skills*, *42*(2), 669–670.

Bidder, T. G., Strain, J. J., & Brunschwig, L. (1970). Bilateral and unilateral ECT: follow-up study and critique. *The American Journal of Psychiatry*, *127*(6), 737–745.

Brakemeier, E.-L., Berman, R., Prudic, J., Zwillenberg, K., & Sackeim, H. A. (2011). Self-evaluation of the cognitive effects of electroconvulsive therapy. *The Journal of ECT*, *27*(1), 59–66. http://doi.org/10.1097/YCT.0b013e3181d77656

Cannicott, S. M., & Waggoner, R. W. (1967). Unilateral and bilateral electroconvulsive therapy. A comparative study. *Archives of General Psychiatry*, *16*(2), 229–232.

Daniel, W. F., Weiner, R. D., & Crovitz, H. F. (1983a). Autobiographical amnesia with ECT: an analysis of the roles of stimulus wave form, electrode placement, stimulus energy, and seizure length. *Biological Psychiatry*, *18*(1), 121–126.

Daniel, W. F., Weiner, R. D., & Crovitz, H. F. (1983b). Autobiographical amnesia with ECT: an analysis of the roles of stimulus wave form, electrode placement, stimulus energy, and seizure length. *Biological Psychiatry*, *18*(1), 121–126.

D'Elia, G. (1970). *Unilateral Electroconvulsive Therapy*. Copenhagen: Munksgaard.

Devanand, D. P., Fitzsimons, L., Prudic, J., & Sackeim, H. A. (1995). Subjective side effects during electroconvulsive therapy. *Convulsive Therapy*, *11*(4), 232–240.

Elm, E. von, Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*, *335*(7624), 806–808. http://doi.org/10.1136/bmj.39335.541782.AD

Fleminger, J. J., de Horne, D. J., & Nott, P. N. (1970). Unilateral electroconvulsive therapy and cerebral dominance: effect of right- and left-sided electrode placement on verbal memory. *Journal of Neurology, Neurosurgery, and Psychiatry*, *33*(3), 408–411.

Gagnier, J. J., Moher, D., Boon, H., Beyene, J., & Bombardier, C. (2012). Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. *BMC Medical Research Methodology*, *12*, 111. http://doi.org/10.1186/1471-2288-12-111

Halliday, A. M., Davison, K., Browne, M. W., & Kreeger, L. C. (1968). A comparison of the effects on depression and memory of bilateral E.C.T. and unilateral E.C.T. to the dominant and non-dominant hemispheres. *The British Journal of Psychiatry : The Journal of Mental Science*, *114*(513), 997–1012.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., … Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928. http://doi.org/10.1136/bmj.d5928

Horne, R. L., Pettinati, H. M., Sugerman, A. A., & Varga, E. (1985). Comparing bilateral to unilateral electroconvulsive therapy in a randomized study with EEG monitoring. *Archives of General Psychiatry*, *42*(11), 1087–1092.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*, *17*(1), 1–12.

Kellner, C. H., Knapp, R., Husain, M. M., Rasmussen, K., Sampson, S., Cullum, M., … Petrides, G. (2010). Bifrontal, bitemporal and right unilateral electrode placement in ECT: randomised trial. *The British Journal of Psychiatry : The Journal of Mental Science*, *196*(Journal Article), 226–234. http://doi.org/10.1192/bjp.bp.109.066183

Kellner, C. H., Tobias, K. G., & Wiegand, J. (2010). Electrode placement in electroconvulsive therapy (ECT): A review of the literature. *The Journal of ECT*, *26*(3), 175–180. http://doi.org/10.1097/YCT.0b013e3181e48154

Levy, R. (1968). The clinical evaluation of unilateral electroconvulsive therapy. *The British Journal of Psychiatry : The Journal of Mental Science*, *114*(509), 459–463.

Lisanby, S. H., Maddox, J. H., Prudic, J., Devanand, D. P., & Sackeim, H. A. (2000). The effects of electroconvulsive therapy on memory of autobiographical and public events. *Archives of General Psychiatry*, *57*(6), 581–590.

Lo, C. K.-L., Mertz, D., & Loeb, M. (2014). Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Medical Research Methodology*, *14*, 45. http://doi.org/10.1186/1471-2288-14-45

Loo, C. K., Sainsbury, K., Sheehan, P., & Lyndon, B. (2008). A comparison of RUL ultrabrief pulse (0.3 ms) ECT and standard RUL ECT. *The International Journal of Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, *11*(7), 883–890. http://doi.org/10.1017/S1461145708009292

McCall, W. V., Dunn, A., Rosenquist, P. B., & Hughes, D. (2002). Markedly suprathreshold right unilateral ECT versus minimally suprathreshold bilateral ECT: antidepressant and memory effects. *The Journal of ECT*, *18*(3), 126–129.

McElhiney, M. C., Moody, B. J., Steif, B. L., Prudic, J., Devanand, D. P., Nobler, M. S., & Sackeim, H. A. (1995). Autobiographical Memory and Mood: Effects of Electroconvulsive Therapy. *Neuropsychology*, *9*(4), 501–517.

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., … Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.

Journal of Clinical Epidemiology, *62*(10), 1006–1012. http://doi.org/10.1016/j.jclinepi.2009.06.005

O'Connor, M., Lebowitz, B. K., Ly, J., Panizzon, M. S., Elkin-Frankston, S., Dey, S., … Pearlman, C. (2008). A dissociation between anterograde and retrograde amnesia after treatment with electroconvulsive therapy: a naturalistic investigation. *The Journal of ECT*, *24*(2), 146–151. http://doi.org/10.1097/YCT.0b013e318158792f

Pettinati, H. M., & Rosenberg, J. (1984). Memory self-ratings before and after electroconvulsive therapy: depression-versus ECT induced. *Biological Psychiatry*, *19*(4), 539–548.

Prudic, J., Olfson, M., Marcus, S. C., Fuller, R. B., & Sackeim, H. A. (2004). Effectiveness of electroconvulsive therapy in community settings. *Biological Psychiatry*, *55*(3), 301–312.

Ranjkesh, F., Barekatain, M., & Akuchakian, S. (2005). Bifrontal versus right unilateral and bitemporal electroconvulsive therapy in major depressive disorder. *The Journal of ECT*, *21*(4), 207–210.

Rosenberg, J., & Pettinati, H. M. (1984). Differential memory complaints after bilateral and unilateral ECT. *The American Journal of Psychiatry*, *141*(9), 1071–1074.

Sackeim, H. A., Prudic, J., Devanand, D. P., Nobler, M. S., Lisanby, S. H., Peyser, S., … Clark, J. (2000). A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities. *Archives of General Psychiatry*, *57*(5), 425–434.

Sackeim, H. A., Prudic, J., Fuller, R., Keilp, J., Lavori, P. W., & Olfson, M. (2007). The cognitive effects of electroconvulsive therapy in community settings. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *32*(1), 244–254. http://doi.org/10.1038/sj.npp.1301180

Sackheim HA, Prudic J, Devanand DP, Kiersky JE, Fitzsimons L, Moody BJ, … Sttembrino JM. (1993). Effects of stimulus intensity and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy. *New England Journal of Medicine*, *328*(12), 839–846.

Schat, A., van den Broek, W. W., Mulder, P. G., Birkenhager, T. K., van Tuijl, R., & Murre, J. M. (2007a). Changes in everyday and semantic memory function after electroconvulsive therapy for unipolar depression. *The Journal of ECT*, *23*(3), 153–157. http://doi.org/10.1097/yct.0b013e318065aa0c

Schat, A., van den Broek, W. W., Mulder, P. G., Birkenhager, T. K., van Tuijl, R., & Murre, J. M. (2007b). Changes in everyday and semantic memory function after electroconvulsive therapy for unipolar depression. *The Journal of ECT*, *23*(3), 153–157. http://doi.org/10.1097/yct.0b013e318065aa0c

Semkovska, M., & McLoughlin, D. M. (2010). Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. *Biological Psychiatry*, *68*(6), 568–577. http://doi.org/10.1016/j.biopsych.2010.06.009

Sienaert, P., Vansteelandt, K., Demyttenaere, K., & Peuskens, J. (2009). Randomized comparison of ultra-brief bifrontal and unilateral electroconvulsive therapy for major depression: clinical efficacy. *Journal of Affective Disorders*, *116*(1-2), 106–112. http://doi.org/10.1016/j.jad.2008.11.001

Sienaert, P., Vansteelandt, K., Demyttenaere, K., & Peuskens, J. (2010). Randomized comparison of ultra-brief bifrontal and unilateral electroconvulsive therapy for major depression: cognitive side-effects. *Journal of Affective Disorders*, *122*(1-2), 60–67. http://doi.org/10.1016/j.jad.2009.06.011

Sobin, C., Sackeim, H. A., Prudic, J., Devanand, D. P., Moody, B. J., & McElhiney, M. C. (1995). Predictors of retrograde amnesia following ECT. *The American Journal of Psychiatry*, *152*(7), 995–1001.

Squire, L. R., & Chace, P. M. (1975). Memory functions six to nine months after electroconvulsive therapy. *Archives of General Psychiatry*, *32*(12), 1557–1564.

Squire, L. R., & Slater, P. C. (1983). Electroconvulsive therapy and complaints of memory dysfunction: a prospective three-year follow-up study. *The British Journal of Psychiatry : The Journal of Mental Science*, *142*(Journal Article), 1–8.

Stoppe, A., Louza, M., Rosa, M., Gil, G., & Rigonatti, S. (2006). Fixed high-dose electroconvulsive therapy in the elderly with depression: a double-blind, randomized comparison of efficacy and tolerability between unilateral and bilateral electrode placement. *The Journal of ECT*, *22*(2), 92–99.

Strain, J. J., Brunschwig, L., Duffy, J. P., Agle, D. P., Rosenbaum, A. L., & Bidder, T. G. (1968). Comparison of therapeutic effects and memory changes with bilateral and unilateral ECT. *The American Journal of Psychiatry*, *125*(3), 50–60.

Stromgren, L. S., & Juul-Jensen, P. (1975). EEG in unilateral and bilateral electroconvulsive therapy. *Acta Psychiatrica Scandinavica*, *51*(5), 340–360.

Taylor, M. A., & Abrams, R. (1985a). Short-term cognitive effects of unilateral and bilateral ECT. *The British Journal of Psychiatry : The Journal of Mental Science*, *146*(Journal Article), 308–311.

Taylor, M. A., & Abrams, R. (1985b). Short-term cognitive effects of unilateral and bilateral ECT. *The British Journal of Psychiatry : The Journal of Mental Science*, *146*(Journal Article), 308–311.

Tew, J. D., Jr, Mulsant, B. H., Haskett, R. F., Dolata, D., Hixson, L., & Mann, J. J. (2002). A randomized comparison of high-charge right unilateral electroconvulsive therapy and bilateral electroconvulsive therapy in older depressed patients who failed to respond to 5 to 8 moderate-charge right unilateral treatments. *The Journal of Clinical Psychiatry*, *63*(12), 1102–1105.

Tor, P.-C., Bautovich, A., Wang, M.-J., Martin, D., Harvey, S. B., & Loo, C. (2015). A Systematic Review and Meta-Analysis of Brief Versus Ultrabrief Right Unilateral Electroconvulsive Therapy for Depression. *The Journal of Clinical Psychiatry*, *76*(9), e1092–1098. http://doi.org/10.4088/JCP.14r09145

Weeks, D., Freeman, C. P., & Kendell, R. E. (1980). ECT: III: Enduring cognitive deficits? *The British Journal of Psychiatry : The Journal of Mental Science*, *137*(Journal Article), 26–37.

CHAPTER 5:

INTEGRATIVE DISCUSSION, FUTURE DIRECTIONS, AND CONCLUSIONS

5.1. Integrative Discussion

The previous three chapters formed the core components of the thesis, which examined two important issues in the treatment of MD. The first issue was medication adherence and the second issue was cognitive impairment following ECT for TRD. Specifically, the thesis described the preliminary results of an RCT conducted to investigate the comparative efficacy of PIMM/SAM versus SPP in improving medication adherence in persons admitted to the inpatient unit of a mood disorders program. As well, the thesis reported the results of a systematic review and meta-analysis of cognitive deficits in persons who received bilateral versus unilaterial ECT. The current chapter presents an integrative discussion of these three chapters, plus directions for future research. Chapter 2 described the RCT and showed that the PIMM/SAM program may have some benefits versus SPP in improving medication adherence. This is important given the challenge of medication adherence in persons with MD. To date, no definitive program exists to boost adherence in persons with MD. Findings from a review of other programs found that a multidimensional approach, rather than education alone, is necessary to improve adherence (Chong et al., 2011). PIMM/SAM is an example of such an approach because it combines education (informing persons about the usage and importance of their medications), participant empowerment (allowing persons with MD to describe their understanding of their medications, allowing them to choose tailored and individualized reminders to take their medications), and participant-nurse interaction (participants explain the usage and importance of their medications to their nurses on a daily basis) to

equip persons with MD with the knowledge to understand why they should take their medications. Further, PIMM/SAM provides these persons with daily reinforcement and reminders about their medications to habitualize the taking of pharmacotherapy in line with the physicians' prescribing.

Even among persons who adhere to treatment, therapeutic non-response is an issue in MD. The problem is especially prevalent in the treatment of MDD (Souery et al., 2006). ECT is one modality used to treat persons with TRD, but reports of cognitive impairment following ECT have led to questions about the benefits versus risks of the treatment. To examine the body of evidence for this issue, we conducted a systematic review and meta-analysis of primary studies that measured cognition in groups of persons who received bilateral or unilateral ECT. Chapter 3 contains two published articles related to the systematic review and meta-analysis, namely a preliminary methods report on the test-retest and inter-rater reliability of risk of bias assessments of ECT studies (M. Oremus et al., 2012) and the protocol for our systematic review and meta-analysis (C. Oremus et al., 2015).

Our findings in the test-retest and inter-rater reliability study suggested the need for improved rater training during the important phase of assessing the risk of bias in published studies (study quality). Risk of bias assessment is an important component of any systematic review or meta-analysis because the conclusions one can draw from the body of evidence on a particular subject must be moderated by the degree of bias in the evidence. High levels of bias in a majority of published studies suggest the body of

evidence contains many invalid conclusions. Therefore, clinicians should be hesitant to use the evidence to inform practice decisions. In preparation for the review reported in this thesis, raters were trained in the use of relevant risk of bias scales such as the Jadad scale and the Newcastle-Ottawa scale. Further, the raters received instruction on the nuances of study design to enable them to validly extract data from the published study reports. Assessing risk of bias and extracting study data are challenging processes that require careful rater training. Besides engaging raters with enough basic knowledge about research design and study bias, raters must be taught how read published study reports. Many published reports are poorly written and lacking in important details such as recruitment and follow-up procedures, and many describe outcomes selectively. An algorithm (Hartling, Bond, Santaguida, Viswanathan, & Dryden, 2011) even exists to help readers (and systematic reviewers) identify study designs in journal articles because authors do not always clearly report such basic details clearly. The test-retest and interrater reliability study provided guidance to train the raters for the ECT systematic review and meta-analysis to understand the questions on the risk of bias scales in the context of the review topic. Additionally, the raters were trained to extract data on a sample of studies included in the systematic review and meta-analysis.

The protocol for the systematic review and meta-analysis describes the methods used to obtain the results reported in Chapter 4. The protocol is written in conformity with existing guidelines for such documents (Moher et al., 2015) and the systematic review and meta-analysis itself conforms to current methods guidance (Moher et al., 2009a).

Further, the review is registered with the PROSPERO international registry of systematic reviews (Booth et al., 2012) (ID # CRD42014009100). The concern with methods and registration permits the scientific community to assess whether the review methods are transparent, reproducible, and valid. Also, registration reduces overlap and duplication because researchers (and policy makers who often commission systematic reviews) can determine whether a similar review is already ongoing (Stewart, Moher, & Shekelle, 2012).

Chapter 4 contains the results of the review of cognitive impairment following bilateral versus unilateral ECT. Although many summary odds ratios were not statistically significant in the forest plots, the results consistently indicated that cognitive impairment was worse for persons who received bilateral ECT at 8 to 30 days post treatment, regardless of the specific treatment modality, when compared to persons who received unilateral ECT. Furthermore, clinically, these results confirm that unilateral ECT should be the initial treatment option in TRD unless a specific persons' clinical presentation suggests differently (i.e., a person has not previously responded to unilateral). A challenge with summarizing the available evidence of ECT and cognition is the substantial amount of clinical heterogeneity in studies published in this area. For example, authors often use many tests to measure the same cognitive domains. This can overdramatize a positive or negative finding because several different instruments point to the same conclusion or, one or two of the tests in a large cognitive battery measuring the same domain may produce a chance finding that contradicts the other findings. A

similar issue is the lack of consensus regarding the most suitable cognitive tests to employ. Some studies selectively reported certain cognitive outcomes. Other challenges with comparing studies in this area concern dissimilar types of treatment modalities, length of follow up, follow-up time point, and study groups.

Other systematic reviews and meta-analyses have examined cognitive impairment following ECT. While the findings are consistent with the review reported in Chapter 4, the previous reviews suffered from some methodological gaps that limit the applicability of their conclusions. For example, Tor et al. found better cognitive outcomes in persons who received ultrabrief pulse instead of brief pulse right unilateral ECT in six studies (Tor et al., 2015). Dunne and McLoughliin studied the efficacy and side effects of bifrontal to bitemporal ECT versus unilateral ECT in eight depression trials and found a lower post-treatment decline on the Mini-Mental State Examination for bifrontal versus bitemporal ECT, but not unilateral (Dunne & McLoughlin, 2012). Semkovska and McLoughlin meta-analyzed 84 studies covering eight cognitive domains and concluded that cognitive impairment did not persist beyond three days post-ECT (Semkovska & McLoughlin, 2010). The UK ECT Review Group identified three studies and found inconclusive results with respect to cognition (The UK ECT Review Group, 2003). None of these reviews assessed the risk of bias of included studies, nor did they grade the strength of evidence or register their reviews in a registry such as PROSPERO.
Overall, this thesis is an important contribution to the field of treating MD. The RCT provides early evidence for the efficacy of a program to increase medication adherence in

inpatients and the methodologically rigorous systematic review shows evidence of cognitive impairment in persons who are treated with bilateral versus unilateral ECT.

5.2. Future Directions

The RCT is part of a sequential explanatory mixed methods study examining the efficacy of PIMM/SAM versus SPP. The experience with the study to date shows that recruitment and follow-up of study participants from a mood disorders program is possible, despite the challenges of conducting research in this group of people. Recruitment of participants into the RCT will continue beyond the thesis and the planned follow-up is 12 months per person. Besides the battery of tests described in Chapter 2, another outcome will be the time to re-hospitalization in each group. Between-group differences in time to re-hospitalization will be compared using Kaplan-Meier survival curves and Cox proportional hazards regression. All statistical analyses will be performed using R software (R Core Team, 2013).

An additional component of the evaluation of PIMM/SAM and SPP will be to qualitatively explore reasons for low adherence in both study groups. To do so, a purposeful sample of participants who score seven or less on the MARS at the 12-month follow-up time point will be randomly invited to participate in a qualitative study. Trained interviewers will use a semi-structured, qualitative one-on-one interview to ask the following questions:

Did participants feel they received adequate instruction with respect to taking medications?

Did participants feel that they were better empowered to take responsibility for their health and well-being following participation in the protocol?

Were participants more satisfied with the quality of care they received during their hospital stay?

What might have been added to the program to make taking medications easier? Further questions may be added to the interview, depending on the findings of the RCT. The interviews will be audio-recorded and transcribed verbatim.

Content analysis (Krippendorff, 2004) will guide the thematic categorization of participants' interview responses. Two independent reviewers will read and re-read the transcripts and identify key themes related the medication adherence. After the reviewers identify an initial list of themes, they will meet, group similar themes together, and develop a codebook of themes. The reviewers will then separately analyze each interview again, applying the themes to the text and identifying additional themes as required. The reviewers will then meet and reconcile differences and develop the final list of themes. NVivo 10 (QRS International Pty Ltd, Doncaster, Australia) will be used to organize and code the data. The reviewers will keep an audit trail of detailed notes to record the development and evolution of the themes. Recruitment of participants will continue until saturation. A sample size of 15 participants per group should be the maximum number of persons that will be needed to reach saturation (Guest, Bunce, & Johnson, 2006).

The final component of the PIMM/SAM versus SPP study will be an economic evaluation, undertaken from the healthcare system perspective, to compare the costs of

first re-hospitalization between each study group. When a participant is re-hospitalized for the first time at any point during the 12-month follow-up period, research staff will conduct a chart review of the entire re-hospitalization to identify all of the direct medical resources consumed during the re-hospitalization. The chart review will cover the entire length of the re-hospitalization, even if this length exceeds the 12-month follow-up period for the participant in question. Direct medical resources include physician, nurse, and other healthcare professionals' time, use of disposable supplies, lab tests, and medications. Costs will be attached to these resources using Ontario Hospital Insurance Plan billing rates, prescribed wage rates, Ontario Drug Benefit Program reimbursement rates, and market rates for disposables. The cost of first re-hospitalization will be calculated for each participant. A generalized linear model with a log link and gamma distribution will be used to compare the difference in cost between study groups. In this model, the cost of re-hospitalization will be the dependent variable and the randomization group will be the independent variable.

The meta-analysis found some results to suggest the existence of worse cognitive performance in persons who received bilateral versus unilateral ECT. However, the evidence was generally weak and a substantial amount of remaining work needs to be done to provide a firmer scientific basis to guide clinical decision-making. Further research in this area would benefit from clearly defended ECT protocols and consensus on a standard battery of cognitive tests to promote applicability and comparability across

studies and strengthen the body of evidence regarding ECT's impact on cognitive performance.

5.3. Conclusions

This thesis showed that persons admitted to an inpatient mood disorders clinic, who received a structured medication training program, may have better adherence-related outcomes than persons who received standard prescribing practice. The study to evaluate PIMM/SAM is ongoing and it is expected to generate further evidence in support of the active program. The systematic review and meta-analysis employed current methods and showed that cognitive performance was worse in persons who received bilateral versus unilateral ECT in some cognitive domains both prior to receipt of treatment and at 8 to 30 days post-treatment.

References - Chapters 1, 3 (foreword) and 5

Abrams, R. (2007). Does bilateral ECT cause persistent cognitive impairment? *The Journal of ECT*, *23*(2), 61–62. http://doi.org/10.1097/yct.0b013e3180544997

American Psychiatric Association, American Psychiatric Association, & DSM-5 Task Force. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Retrieved from http://dsm.psychiatryonline.org/book.aspx?bookid=556

Ashton, R., & Hess, N. (1976). Amnesia for random shapes following unilateral and bilateral electronvulsive shock therapy. *Perceptual and Motor Skills*, *42*(2), 669–670.

Bauer, I. E., Wu, M.-J., Frazier, T. W., Mwangi, B., Spiker, D., Zunta-Soares, G. B., & Soares, J. C. (2016). Neurocognitive functioning in individuals with bipolar disorder and their healthy siblings: A preliminary study. *Journal of Affective Disorders*, *201*, 51–56. http://doi.org/10.1016/j.jad.2016.04.026

Bauer, J., Hageman, I., Dam, H., Baez, A., Bolwig, T., Roed, J., … Jorgensen, M. B. (2009). Comparison of propofol and thiopental as anesthetic agents for electroconvulsive therapy: a randomized, blinded comparison of seizure duration, stimulus charge, clinical effect, and cognitive side effects. *The Journal of ECT*, *25*(2), 85–90. http://doi.org/10.1097/YCT.0b013e31818a0203

Bergfeld, I. O., Mantione, M., Hoogendoorn, M. L. C., Ruhé, H. G., Notten, P., van Laarhoven, J., … Denys, D. (2016). Deep Brain Stimulation of the Ventral Anterior Limb of the Internal Capsule for Treatment-Resistant Depression: A

Randomized Clinical Trial. *JAMA Psychiatry*, *73*(5), 456–464. http://doi.org/10.1001/jamapsychiatry.2016.0152

Berlim, M. T., McGirr, A., Van den Eynde, F., Fleck, M. P. A., & Giacobbe, P. (2014). Effectiveness and acceptability of deep brain stimulation (DBS) of the subgenual cingulate cortex for treatment-resistant depression: a systematic review and exploratory meta-analysis. *Journal of Affective Disorders*, *159*, 31–38. http://doi.org/10.1016/j.jad.2014.02.016

Booth, A., Clarke, M., Dooley, G., Ghersi, D., Moher, D., Petticrew, M., & Stewart, L. (2012). The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic Reviews*, *1*, 2. http://doi.org/10.1186/2046-4053-1-2

Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders*, *113*(1), 1–20. http://doi.org/10.1016/j.jad.2008.06.009

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: John wiley & Sons, Ltd.

Brakemeier, E.-L., Berman, R., Prudic, J., Zwillenberg, K., & Sackeim, H. A. (2011). Self-evaluation of the cognitive effects of electroconvulsive therapy. *The Journal of ECT*, *27*(1), 59–66. http://doi.org/10.1097/YCT.0b013e3181d77656

British Medical Journal. (2016). *Medical Milestones*. London: British Medical Journal. Available at: http://www.bmj.com/content/medical-milestones. Accessed on: July 3.

Brunoni, A. R., Baeken, C., Machado-Vieira, R., Gattaz, W. F., & Vanderhasselt, M.-A. (2014). BDNF blood levels after electroconvulsive therapy in patients with mood disorders: a systematic review and meta-analysis. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, *15*(5), 411–418. http://doi.org/10.3109/15622975.2014.892633

Canadian Agency for Drugs and Technologies in Health (CADTH). (2014). *CADTH Common Drug Review: Procedure for the CADTH Common Drug Review*. Ottawa: CADTH.

Canadian Agency for Drugs and Technologies in Health (CADTH). (2016). *CADTH - Evidence Driven*. Ottawa: CADTH. Available at: https://www.cadth.ca/. Accessed on: July 3.

Cannicott, S. M., & Waggoner, R. W. (1967). Unilateral and bilateral electroconvulsive therapy. A comparative study. *Archives of General Psychiatry*, *16*(2), 229–232.

Chong, W. W., Aslani, P., & Chen, T. F. (2011). Effectiveness of interventions to improve antidepressant medication adherence: a systematic review. *International Journal of Clinical Practice*, *65*(9), 954–975. http://doi.org/10.1111/j.1742-1241.2011.02746.x

Coldham, E. L., Addington, J., & Addington, D. (2002). Medication adherence of individuals with a first episode of psychosis. *Acta Psychiatrica Scandinavica*, *106*(4), 286–290.

Daniel, W. F., Weiner, R. D., & Crovitz, H. F. (1983). Autobiographical amnesia with ECT: an analysis of the roles of stimulus wave form, electrode placement, stimulus energy, and seizure length. *Biological Psychiatry*, *18*(1), 121–126.

Darling, C. A., Olmstead, S. B., Lund, V. E., & Fairclough, J. F. (2008). Bipolar disorder: medication adherence and life contentment. *Archives of Psychiatric Nursing*, *22*(3), 113–126. http://doi.org/10.1016/j.apnu.2008.02.004

D'Elia, G. (1970). *Unilateral Electroconvulsive Therapy*. Copenhagen: Munksgaard.

Department of Psychiatry. (Available at: http://www.psych.med.umich.edu/ect/common-side-effects.asp. Accessed on July19). *Electroconvulsive Therapy Program: Common Side Effects*. Ann Arbor, MI: University of Michigan Health System.

Depp, C. A., Mausbach, B. T., Eyler, L. T., Palmer, B. W., Cain, A. E., Lebowitz, B. D., … Jeste, D. V. (2009). Performance-based and subjective measures of functioning in middle-aged and older adults with bipolar disorder. *The Journal of Nervous and Mental Disease*, *197*(7), 471–475. http://doi.org/10.1097/NMD.0b013e3181ab5c9b

Devanand, D. P., Fitzsimons, L., Prudic, J., & Sackeim, H. A. (1995). Subjective side effects during electroconvulsive therapy. *Convulsive Therapy*, *11*(4), 232–240.

Dickersin, K., Straus, S. E., & Bero, L. A. (2007). Evidence based medicine: increasing, not dictating, choice. *BMJ*, *334*(suppl 1), s10–s10. http://doi.org/10.1136/bmj.39062.639444.94

Dickerson, F. B., Boronow, J. J., Stallings, C. R., Origoni, A. E., Cole, S., & Yolken, R. H. (2004). Association between cognitive functioning and employment status of persons with bipolar disorder. *Psychiatric Services (Washington, D.C.)*, *55*(1), 54–58. http://doi.org/10.1176/appi.ps.55.1.54

Dougherty, D. D., & Rauch, S. L. (2007). Somatic Therapies for Treatment- Resistant Depression: New Neurotherapeutic Interventions. *Psychiatric Clinics of North America*, *30*(1), 31–37. http://doi.org/10.1016/j.psc.2006.12.006

Dunbar-Jacob, J., & Mortimer-Stephens, M. K. (2001). Treatment adherence in chronic disease. *Journal of Clinical Epidemiology*, *54*(12, Supplement 1), S57–S60. http://doi.org/10.1016/S0895-4356(01)00457-7

Dunkin, J. J., Leuchter, A. F., Cook, I. A., Kasl-Godley, J. E., Abrams, M., & Rosenberg-Thompson, S. (2000). Executive dysfunction predicts nonresponse to fluoxetine in major depression. *Journal of Affective Disorders*, *60*(1), 13–23.

Dunne, R. A., & McLoughlin, D. M. (2012). Systematic review and meta-analysis of bifrontal electroconvulsive therapy versus bilateral and unilateral electroconvulsive therapy in depression. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of*

Biological Psychiatry, *13*(4), 248–258. http://doi.org/10.3109/15622975.2011.615863

Elm, E. von, Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*, *335*(7624), 806–808. http://doi.org/10.1136/bmj.39335.541782.AD

Emsley, R., & Hawkridge, S. (2009). The quest for a meaningful evidence base in psychiatry. *World Psychiatry*, *8*(1), 33–34.

Fink, M. (2007). Complaints of loss of personal memories after electroconvulsive therapy: evidence of a somatoform disorder? *Psychosomatics*, *48*(4), 290–293. http://doi.org/10.1176/appi.psy.48.4.290

Fleminger, J. J., de Horne, D. J., & Nott, P. N. (1970). Unilateral electroconvulsive therapy and cerebral dominance: effect of right- and left-sided electrode placement on verbal memory. *Journal of Neurology, Neurosurgery, and Psychiatry*, *33*(3), 408–411.

Flory, J. D., & Yehuda, R. (2015). Comorbidity between post-traumatic stress disorder and major depressive disorder: alternative explanations and treatment considerations. *Dialogues in Clinical Neuroscience*, *17*(2), 141–150.

Gagnier, J. J., Moher, D., Boon, H., Beyene, J., & Bombardier, C. (2012). Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in

the literature. *BMC Medical Research Methodology*, *12*, 111. http://doi.org/10.1186/1471-2288-12-111

Ganguli, A., Clewell, J., & Shillington, A. C. (2016). The impact of patient support programs on adherence, clinical, humanistic, and economic patient outcomes: a targeted systematic review. *Patient Preference and Adherence*, *10*, 711–725. http://doi.org/10.2147/PPA.S101175

Gartlehner, G., Gaynes, B. N., Hansen, R. A., Thieda, P., DeVeaugh-Geiss, A., Krebs, E. E., … Lohr, K. N. (2008). Comparative benefits and harms of second-generation antidepressants: background paper for the American College of Physicians. *Annals of Internal Medicine*, *149*(10), 734–750.

Gaynes, B. N., Dusetzina, S. B., Ellis, A. R., Hansen, R. A., Farley, J. F., Miller, W. C., & Stürmer, T. (2012). Treating depression after initial treatment failure: directly comparing switch and augmenting strategies in STAR*D. *Journal of Clinical Psychopharmacology*, *32*(1), 114–119. http://doi.org/10.1097/JCP.0b013e31823f705d

Gaynes, B. N., Lux, L. J., Lloyd, S. W., Hansen, R. A., Gartlehner, G., Keener, P., … Lohr, K. N. (2011). *Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults*. Rockville (MD): Agency for Healthcare Research and Quality (US). Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK65315/

Gildengers, A. G., Butters, M. A., Chisholm, D., Rogers, J. C., Holm, M. B., Bhalla, R. K., … Mulsant, B. H. (2007). Cognitive functioning and instrumental activities of

daily living in late-life bipolar disorder. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *15*(2), 174–179. http://doi.org/10.1097/JGP.0b013e31802dd367

Government of Canada, S. C. (2016, March 7). Mood disorders, by age group and sex (Number). Retrieved July 11, 2016, from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/health113a-eng.htm

GRADE Working Group. (2004). Grading quality of evidence and strength of recommendations. *BMJ : British Medical Journal*, *328*(7454), 1490.

Gray, G. E., & Pinson, L. A. (2003). Evidence-based medicine and psychiatric practice. *The Psychiatric Quarterly*, *74*(4), 387–399.

Greenhalgh, T., Howick, J., & Maskrey, N. (2014). Evidence based medicine: a movement in crisis? *BMJ*, *348*, g3725. http://doi.org/10.1136/bmj.g3725

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59–82. http://doi.org/10.1177/1525822X05279903

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62.

Hartling, L., Bond, K., Santaguida, P. L., Viswanathan, M., & Dryden, D. M. (2011). Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy.

Journal of Clinical Epidemiology, *64*(8), 861–871. http://doi.org/10.1016/j.jclinepi.2011.01.010

Hartling, L., Milne, A., Hamm, M. P., Vandermeer, B., Ansari, M., Tsertsvadze, A., & Dryden, D. M. (2013). Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *Journal of Clinical Epidemiology*, *66*(9), 982–993. http://doi.org/10.1016/j.jclinepi.2013.03.003

Hasselbalch, B. J., Knorr, U., Hasselbalch, S. G., Gade, A., & Kessing, L. V. (2012). Cognitive deficits in the remitted state of unipolar depressive disorder. *Neuropsychology*, *26*(5), 642–651. http://doi.org/10.1037/a0029301

Health Canada. (2002). *A Report on Mental Illnesses in Canada*. Ottawa: Health Canada.

Higgins, J. P., & Green, S. (eds). (2016). *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. Available at: http://handbook.cochrane.org/ Accessed on: July 3.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., … Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, *343*, d5928. http://doi.org/10.1136/bmj.d5928

Horne, R. L., Pettinati, H. M., Sugerman, A. A., & Varga, E. (1985). Comparing bilateral to unilateral electroconvulsive therapy in a randomized study with EEG monitoring. *Archives of General Psychiatry*, *42*(11), 1087–1092.

Horne, R., Weinman, J., Barber, N., Elliott, R., & Morgan, M. (2006). *Concordance, Adherence and Compliance in Medicine Taking: A conceptual map and research*

priorities. London: National Institute for Health Research (NIHR) Service Delivery and Organisation (SDO) Programme.

Iorfino, F., Hickie, I. B., Lee, R. S. C., Lagopoulos, J., & Hermens, D. F. (2016). The underlying neurobiology of key functional domains in young people with mood and anxiety disorders: a systematic review. *BMC Psychiatry*, *16*, 156. http://doi.org/10.1186/s12888-016-0852-3

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*, *17*(1), 1–12.

Jaeger, J., & Vieta, E. (2007). Functional outcome and disability in bipolar disorders: ongoing research and future directions. *Bipolar Disorders*, *9*(1-2), 1–2. http://doi.org/10.1111/j.1399-5618.2007.00441.x

Kellner, C. H., Knapp, R., Husain, M. M., Rasmussen, K., Sampson, S., Cullum, M., … Petrides, G. (2010). Bifrontal, bitemporal and right unilateral electrode placement in ECT: randomised trial. *The British Journal of Psychiatry : The Journal of Mental Science*, *196*(Journal Article), 226–234. http://doi.org/10.1192/bjp.bp.109.066183

Kellner, C. H., Tobias, K. G., & Wiegand, J. (2010). Electrode placement in electroconvulsive therapy (ECT): A review of the literature. *The Journal of ECT*, *26*(3), 175–180. http://doi.org/10.1097/YCT.0b013e3181e48154

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Chatterji, S., Lee, S., Ormel, J., … Wang, P. S. (2009). The global burden of mental disorders: An update from the WHO World Mental Health (WMH) Surveys. *Epidemiologia E Psichiatria Sociale*, *18*(1), 23–33.

Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications, Inc.

Lam, R. W., Bartley, S., Yatham, L. N., Tam, E. M., & Zis, A. P. (1999). Clinical predictors of short-term outcome in electroconvulsive therapy. *Canadian Journal of psychiatry.Revue Canadienne de Psychiatrie*, *44*(2), 158–163.

Lanius, R. A. (2015). Trauma-related dissociation and altered states of consciousness: a call for clinical, treatment, and neuroscience research. *European Journal of Psychotraumatology*, *6*, 27905.

Lee, J. C., Blumberger, D. M., Fitzgerald, P. B., Daskalakis, Z. J., & Levinson, A. J. (2012). The role of transcranial magnetic stimulation in treatment-resistant depression: a review. *Current Pharmaceutical Design*, *18*(36), 5846–5852.

Lee, R. S. C., Hermens, D. F., Porter, M. A., & Redoblado-hodge, M. A. (2012). A metaanalysis of cognitive deficits in first episode Major Depressive Disorder. *Journal of Affective Disorers, 140*, 113–124

Levy, R. (1968). The clinical evaluation of unilateral electroconvulsive therapy. *The British Journal of Psychiatry : The Journal of Mental Science*, *114*(509), 459–463.

Lo, C. K.-L., Mertz, D., & Loeb, M. (2014). Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Medical Research Methodology*, *14*, 45. http://doi.org/10.1186/1471-2288-14-45

Loo, C. K., Sainsbury, K., Sheehan, P., & Lyndon, B. (2008). A comparison of RUL ultrabrief pulse (0.3 ms) ECT and standard RUL ECT. *The International Journal of Neuropsychopharmacology / Official Scientific Journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, *11*(7), 883–890. http://doi.org/10.1017/S1461145708009292

MacQueen, G. M., Campbell, S., McEwen, B. S., Macdonald, K., Amano, S., Joffe, R. T., … Young, L. T. (2003). Course of illness, hippocampal function, and hippocampal volume in major depression. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(3), 1387–1392. http://doi.org/10.1073/pnas.0337481100

McCall, W. V., Dunn, A., Rosenquist, P. B., & Hughes, D. (2002). Markedly suprathreshold right unilateral ECT versus minimally suprathreshold bilateral ECT: antidepressant and memory effects. *The Journal of ECT*, *18*(3), 126–129.

McCartney, M., Treadwell, J., Maskrey, N., & Lehman, R. (2016). Making evidence based medicine work for individual patients. *BMJ*, *353*, i2452. http://doi.org/10.1136/bmj.i2452

Merikangas, K. R., Jin, R., He, J.-P., Kessler, R. C., Lee, S., Sampson, N. A., … Zarkov, Z. (2011). Prevalence and Correlates of Bipolar Spectrum Disorder in the World

Mental Health Survey Initiative. *Archives of General Psychiatry*, *68*(3), 241–251. http://doi.org/10.1001/archgenpsychiatry.2011.12

Merkl, A., Heuser, I., & Bajbouj, M. (2009). Antidepressant electroconvulsive therapy: mechanism of action, recent advances and limitations. *Experimental Neurology*, *219*(1), 20–26. http://doi.org/10.1016/j.expneurol.2009.04.027

Merkl, A., Schneider, G.-H., Schönecker, T., Aust, S., Kühl, K.-P., Kupsch, A., … Bajbouj, M. (2013). Antidepressant effects after short-term and chronic stimulation of the subgenual cingulate gyrus in treatment-resistant depression. *Experimental Neurology*, *249*, 160–168. http://doi.org/10.1016/j.expneurol.2013.08.017

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., … Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009a). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical Research Ed.)*, *339*, b2535.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009b). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of Clinical Epidemiology*, *62*(10), 1006–1012. http://doi.org/10.1016/j.jclinepi.2009.06.005

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., … PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*, 1. http://doi.org/10.1186/2046-4053-4-1

Monden, R., de Vos, S., Morey, R., Wagenmakers, E.-J., de Jonge, P., & Roest, A. M. (2016). Toward evidence-based medical statistics: a Bayesian analysis of double-blind placebo-controlled antidepressant trials in the treatment of anxiety disorders. *International Journal of Methods in Psychiatric Research*, n/a–n/a. http://doi.org/10.1002/mpr.1507

Muneer, A. (2016). The Neurobiology of Bipolar Disorder: An Integrated Approach. *Chonnam Medical Journal*, *52*(1), 18–37. http://doi.org/10.4068/cmj.2016.52.1.18

Murrough, J. W., Iacoviello, B., Neumeister, A., Charney, D. S., & Iosifescu, D. V. (2011). Cognitive dysfunction in depression: neurocircuitry and new therapeutic strategies. *Neurobiology of Learning and Memory*, *96*(4), 553–563. http://doi.org/10.1016/j.nlm.2011.06.006

National Institute for Health and Care Excellence (NICE). (2016). *Improving health and social care through evidence-based guidance*. London: NICE. Available at: https://www.nice.org.uk/. Accessed on: July 3.

National Institute of Mental Health. (2016). *Post-Traumatic Stress Disorder*. Available at: http://www.nimh.nih.gov/health/topics/post-traumatic-stress-disorder-ptsd/index.shtml: Accessed on: July 20.

O'Connor, M., Lebowitz, B. K., Ly, J., Panizzon, M. S., Elkin-Frankston, S., Dey, S., … Pearlman, C. (2008). A dissociation between anterograde and retrograde amnesia after treatment with electroconvulsive therapy: a naturalistic investigation. *The Journal of ECT*, *24*(2), 146–151. http://doi.org/10.1097/YCT.0b013e318158792f

Oremus, C., Oremus, M., McNeely, H., Losier, B., Parlar, M., King, M., … McKinnon, M. (2015). Effects of electroconvulsive therapy on cognitive functioning in patients with depression: protocol for a systematic review and meta-analysis. *BMJ Open*, *5*(3), e006966. http://doi.org/10.1136/bmjopen-2014-006966

Oremus, M., Oremus, C., Hall, G. B. C., & McKinnon, M. C. (2012). Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales. *BMJ Open*, *2*(4). http://doi.org/10.1136/bmjopen-2012-001368

Parlar, M. (2015). *Neural and Clinical Correlates of cognitve Processes in Major Depressive Disorder and Posttraumatic Stress Disorder*. Hamilton, ON: McMaster University PhD Thesis.

Parlar, M., Frewen, P. A., Oremus, C., Lanius, R. A., & McKinnon, M. C. (2016). Dissociative symptoms are associated with reduced neuropsychological performance in patients with recurrent depression and a history of trauma exposure. *European Journal of Psychotraumatology*, *7*. http://doi.org/10.3402/ejpt.v7.29061

Polak, A. R., Witteveen, A. B., Reitsma, J. B., & Olff, M. (2012). The role of executive function in posttraumatic stress disorder: a systematic review. *Journal of Affective Disorders*, *141*(1), 11–21. http://doi.org/10.1016/j.jad.2012.01.001

Polyakova, M., Schroeter, M. L., Elzinga, B. M., Holiga, S., Schoenknecht, P., de Kloet, E. R., & Molendijk, M. L. (2015). Brain-Derived Neurotrophic Factor and Antidepressive Effect of Electroconvulsive Therapy: Systematic Review and Meta-Analyses of the Preclinical and Clinical Literature. *PloS One*, *10*(11), e0141564. http://doi.org/10.1371/journal.pone.0141564

Qaseem, A., Snow, V., Cross, J. J. Thomas, Forciea, M. A., Hopkins, J. Robert, Shekelle, P., … Owens, D. K. (2008). Current Pharmacologic Treatment of Dementia: A Clinical Practice Guideline from the American College of Physicians and the American Academy of Family Physicians. *Annals of Internal Medicine*, *148*(5), 370–378. http://doi.org/10.7326/0003-4819-148-5-200803040-00008

Raina, P., Santaguida, P., Ismaila, A., Patterson, C., Cowan, D., Levine, M., … Oremus, M. (2008). Effectiveness of cholinesterase inhibitors and memantine for treating dementia: Evidence review for a clinical practice guideline. *Ann Intern.Med*, *148*(5), 379–397.

Ranjkesh, F., Barekatain, M., & Akuchakian, S. (2005). Bifrontal versus right unilateral and bitemporal electroconvulsive therapy in major depressive disorder. *The Journal of ECT*, *21*(4), 207–210.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis, 2029–2040. *Psychological Medicine, 44* (10). http://dx.doi.org/10.1017/S0033291713002535

Rosenberg, J., & Pettinati, H. M. (1984). Differential memory complaints after bilateral and unilateral ECT. *The American Journal of Psychiatry*, *141*(9), 1071–1074.

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., … Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American Journal of Psychiatry*, *163*(11), 1905–1917. http://doi.org/10.1176/ajp.2006.163.11.1905

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ (Clinical Research Ed.)*, *312*(7023), 71–72.

Sadock, B. J., & Sadock, V. A. (2007). *Synopsis of Psychiatry: Behavioral Sciences/Clinical Psychology* (10th ed.). Philadelphia: Lippincott Williams & Wilkins.

Santarsieri, D., & Schwartz, T. L. (2015). Antidepressant efficacy and side-effect burden: a quick guide for clinicians. *Drugs in Context*, *4*, 212290. http://doi.org/10.7573/dic.212290

Schat, A., van den Broek, W. W., Mulder, P. G., Birkenhager, T. K., van Tuijl, R., & Murre, J. M. (2007). Changes in everyday and semantic memory function after electroconvulsive therapy for unipolar depression. *The Journal of ECT*, *23*(3), 153–157. http://doi.org/10.1097/yct.0b013e318065aa0c

Scott, J., & Pope, M. (2002). Nonadherence with mood stabilizers: prevalence and predictors. *The Journal of Clinical Psychiatry*, *63*(5), 384–390.

Semkovska, M., & McLoughlin, D. M. (2010). Objective cognitive performance associated with electroconvulsive therapy for depression: a systematic review and meta-analysis. *Biological Psychiatry*, *68*(6), 568–577. http://doi.org/10.1016/j.biopsych.2010.06.009

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., … Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*, 10. http://doi.org/10.1186/1471-2288-7-10

Shilyansky, C., Williams, L. M., Gyurak, A., Harris, A., Usherwood, T., & Etkin, A. (2016). Effect of antidepressant treatment on cognitive impairments associated with depression: a randomised longitudinal study. *The Lancet. Psychiatry*, *3*(5), 425–435. http://doi.org/10.1016/S2215-0366(16)00012-2

Sienaert, P., Vansteelandt, K., Demyttenaere, K., & Peuskens, J. (2009). Randomized comparison of ultra-brief bifrontal and unilateral electroconvulsive therapy for

major depression: clinical efficacy. *Journal of Affective Disorders*, *116*(1-2), 106–112. http://doi.org/10.1016/j.jad.2008.11.001

Sienaert, P., Vansteelandt, K., Demyttenaere, K., & Peuskens, J. (2010). Randomized comparison of ultra-brief bifrontal and unilateral electroconvulsive therapy for major depression: cognitive side-effects. *Journal of Affective Disorders*, *122*(1-2), 60–67. http://doi.org/10.1016/j.jad.2009.06.011

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. http://doi.org/10.1037/a0028727

Sobin, C., Sackeim, H. A., Prudic, J., Devanand, D. P., Moody, B. J., & McElhiney, M. C. (1995). Predictors of retrograde amnesia following ECT. *The American Journal of Psychiatry*, *152*(7), 995–1001.

Souery, D., Papakostas, G. I., & Trivedi, M. H. (2006). Treatment-resistant depression. *The Journal of Clinical Psychiatry*, *67 Suppl 6*, 16–22.

Spinhoven, P., Penninx, B. W., van Hemert, A. M., de Rooij, M., & Elzinga, B. M. (2014). Comorbidity of PTSD in anxiety and depressive disorders: prevalence and shared risk factors. *Child Abuse & Neglect*, *38*(8), 1320–1330. http://doi.org/10.1016/j.chiabu.2014.01.017

Squire, L. R., & Chace, P. M. (1975). Memory functions six to nine months after electroconvulsive therapy. *Archives of General Psychiatry*, *32*(12), 1557–1564.

Squire, L. R., & Slater, P. C. (1983). Electroconvulsive therapy and complaints of memory dysfunction: a prospective three-year follow-up study. *The British Journal of Psychiatry : The Journal of Mental Science*, *142*(Journal Article), 1–8.

Stewart, L., Moher, D., & Shekelle, P. (2012). Why prospective registration of systematic reviews makes sense. *Systematic Reviews*, *1*, 7. http://doi.org/10.1186/2046-4053-1-7

Stoppe, A., Louza, M., Rosa, M., Gil, G., & Rigonatti, S. (2006). Fixed high-dose electroconvulsive therapy in the elderly with depression: a double-blind, randomized comparison of efficacy and tolerability between unilateral and bilateral electrode placement. *The Journal of ECT*, *22*(2), 92–99.

Strain, J. J., Brunschwig, L., Duffy, J. P., Agle, D. P., Rosenbaum, A. L., & Bidder, T. G. (1968). Comparison of therapeutic effects and memory changes with bilateral and unilateral ECT. *The American Journal of Psychiatry*, *125*(3), 50–60.

Straus, S., & Moher, D. (2010). Registering systematic reviews. *CMAJ : Canadian Medical Association Journal*, *182*(1), 13–14. http://doi.org/10.1503/cmaj.081849

Stromgren, L. S., & Juul-Jensen, P. (1975). EEG in unilateral and bilateral electroconvulsive therapy. *Acta Psychiatrica Scandinavica*, *51*(5), 340–360.

Taylor, M. A., & Abrams, R. (1985). Short-term cognitive effects of unilateral and bilateral ECT. *The British Journal of Psychiatry : The Journal of Mental Science*, *146*(Journal Article), 308–311.

Tew, J. D., Jr, Mulsant, B. H., Haskett, R. F., Dolata, D., Hixson, L., & Mann, J. J. (2002). A randomized comparison of high-charge right unilateral electroconvulsive therapy and bilateral electroconvulsive therapy in older depressed patients who failed to respond to 5 to 8 moderate-charge right unilateral treatments. *The Journal of Clinical Psychiatry*, *63*(12), 1102–1105.

The UK ECT Review Group. (2003). Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis. *The Lancet*, *361*(9360), 799–808. http://doi.org/10.1016/S0140-6736(03)12705-5

Tonelli, M., & Guyatt, G. (2016, July 1). The Role of Experience in an Evidence-Based Practice. Retrieved July 3, 2016, from https://themedicalroundtable.com/article/role-experience-evidence-based-practice

Tonelli, M. R. (2011). Integrating clinical research into clinical decision making. *Annali dell'Istituto Superiore Di Sanità*, *47*(1), 26–30. http://doi.org/10.4415/ANN_11_01_07

Tor, P.-C., Bautovich, A., Wang, M.-J., Martin, D., Harvey, S. B., & Loo, C. (2015). A Systematic Review and Meta-Analysis of Brief Versus Ultrabrief Right Unilateral Electroconvulsive Therapy for Depression. *The Journal of Clinical Psychiatry*, *76*(9), e1092–1098. http://doi.org/10.4088/JCP.14r09145

Torrent, C., Martínez-Arán, A., Daban, C., Sánchez-Moreno, J., Comes, M., Goikolea, J. M., … Vieta, E. (2006). Cognitive impairment in bipolar II disorder. *The British*

Journal of Psychiatry: The Journal of Mental Science, *189*, 254–259. http://doi.org/10.1192/bjp.bp.105.017269

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Nierenberg, A. A., Warden, D., Ritz, L., … STAR*D Study Team. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *The American Journal of Psychiatry*, *163*(1), 28–40. http://doi.org/10.1176/appi.ajp.163.1.28

Tunca, Z., Bayn, M., Alkn, T., Özerdem, A., Resmi, H., & Akan, P. (2015). A Preliminary Observation of Increased Glial Cell Line-Derived Neurotrophic Factor in Manic Switch due to Electroconvulsive Treatment in Depressive Patients. *The Journal of ECT*, *31*(3), 167–172. http://doi.org/10.1097/YCT.0000000000000198

Warden, D., Rush, A. J., Trivedi, M. H., Fava, M., & Wisniewski, S. R. (2007). The STAR*D Project results: a comprehensive review of findings. *Current Psychiatry Reports*, *9*(6), 449–459.

Wagner, S., Doering, B., Lieb, K., & Tadic, A. (2012). A meta-analysis of executive dysfunctions in unipolar major depressive disorder without psychotic symptoms and their changes during antidepressant treatment, 281–292. *Acta Psychiatr Scand 125*: 281–292. http://doi.org/10.1111/j.1600-0447.2011.01762.x

Weeks, D., Freeman, C. P., & Kendell, R. E. (1980). ECT: III: Enduring cognitive deficits? *The British Journal of Psychiatry : The Journal of Mental Science*, *137*(Journal Article), 26–37.

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2016). Global Burden of Mental, Neurological, and Substance Use Disorders: An Analysis from the Global Burden of Disease Study 2010. In V. Patel, D. Chisholm, T. Dua, R. Laxminarayan, & M. E. Medina-Mora (Eds.), *Mental, Neurological, and Substance Use Disorders: Disease Control Priorities, Third Edition (Volume 4)*. Washington (DC): The International Bank for Reconstruction and Development / The World Bank. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK361944/

Williams, L. M., Goldstein-Piekarski, A. N., Chowdhry, N., Grisanzio, K. A., Haug, N. A., Samara, Z., … Yesavage, J. (2016). Developing a clinical translational neuroscience taxonomy for anxiety and mood disorder: protocol for the baseline-follow up Research domain criteria Anxiety and Depression ("RAD") project. *BMC Psychiatry*, *16*, 68. http://doi.org/10.1186/s12888-016-0771-3

World Health Organization. (2008). *Global Burden of Disease: 2004 Update*. Geneva: World Health Organization.

APPENDIX

Hamilton Health Sciences
McMaster University
Inspiring Innovation and Discovery
St. Joseph's Healthcare Hamilton

Hamilton Integrated Research Ethics Board (HIREB)
293 Wellington St. N., Suite 102, Hamilton, ON L8L 8E7
Telephone: 905-521-2100, Ext. 42013
Fax: 905-577-8378

December 8, 2014

PROJECT NUMBER: **14-733**

PROJECT TITLE: Partnership in medication management (PIMM): The effects of one-on-one medication training on medication adherence in patients with mood disorders

PRINCIPAL INVESTIGATOR: **Dr. Carolina Oremus**
LOCAL PI: **Dr. Margaret McKinnon**

This will acknowledge receipt of your letter dated November 20, 2014 which enclosed revised copies of the Information/Consent Form and the Application Form for the above-named study. These issues were raised by the Hamilton Integrated Research Ethics Board at their meeting held on October 21, 2014. Based on this additional information, we wish to advise your study has been given ***final*** approval from the full HIREB.

The following documents have been approved on both ethical and scientific grounds:

- The submission
- Clinical Study Protocol version 1 dated August 14, 2014
- Information/Consent Form version 2 dated November 19, 2014
- PIMM Checklist – Goals and Steps for Study Participants version 1 dated September 23, 2014
- PIMM Checklist for Staff version 1 dated September 23, 2014

The following documents have been acknowledged:

- Clinical Trial Registration # NCT02285608

Please note attached you will find the Information/Consent Form with the HIREB approval affixed; all consent forms used in this study must be copies of the attached materials.

REB #: 14-733 McKinnon/Oremus

We are pleased to issue final approval for the above-named study for a period of 12 months from the date of the HIREB meeting on October 21, 2014. Continuation beyond that date will require further review and renewal of HIREB approval. Any changes or revisions to the original submission must be submitted on an HIREB amendment form for review and approval by the Hamilton Integrated Research Ethics Board.

PLEASE QUOTE THE ABOVE-REFERENCE PROJECT NUMBER ON
ALL FUTURE CORRESPONDENCE

Sincerely,

S. Salama

Suzette Salama, PhD.
Chair, Hamilton Integrated Research Ethics Board

St. Joseph's Healthcare Hamilton

McMaster University

LETTER OF INFORMATION / CONSENT

Partnership in medication management (PIMM): The effects of one-on-one medication training on medication adherence in patients with mood disorders

Principal Investigators:
Carolina Oremus, MD, PhD Candidate, McMaster University
Sharon Simons, RN, BScN, CPMHN(c), St. Joseph's Healthcare Hamilton

Local Principal Investigator:
Margaret McKinnon, PhD, C. Psych., McMaster University
Mood Disorders Program
St. Joseph's Healthcare;
100 West 5th St., Box 585
Hamilton, ON L8N 3K7
mmckinno@stjosham.on.ca
Tel: (905) 522-1155, ext. 35438
Fax: (905) 575-6029

Co-Investigator:
Mark Oremus, PhD, University of Waterloo

Funding Source: Canadian Institutes of Health Research (CIHR) and the Mood Disorders Program, St. Joseph's Healthcare Hamilton

You are invited to take part in a research study looking at whether a new personalized medication training program for patients with mood disorders (major depressive disorder [MDD] and bipolar disorder [BD]) will help them take their medications as prescribed by doctors. This study is being carried out at the inpatient unit of the Mood Disorders Program, St. Joseph's Healthcare Hamilton (SJHH).

Participation in this study is voluntary. To see if you want to take part in this research study, you should understand what is involved in participating. You should also understand the potential risks and benefits of participating. Please take your time to review this letter of information and consent form, which gives detailed information about the research study, and feel free to ask any questions you may have. You may take your time to decide whether to participate. You may discuss your participation with your friends, family, or your doctor before making a decision. If these documents contain words that you do not understand, then please ask the study staff for more details.

Why is this research being done?
Drug treatments like anti-depressants or lithium often do not have the desired results in persons with mood disorders. Sometimes the lack of results happens because people do not take their medications as prescribed. Whether people take drugs as prescribed is called 'adherence'.

The World Health Organization (WHO) defines treatment adherence as "the extent to which a patient follows medical instructions". The WHO also recognizes that the doctor-patient relationship is important in treatment adherence.

What is the purpose of this study?
We want to see if a training program designed to educate people about the purpose, dosage, benefits, and side effects of the medications they are taking for mood disorders will improve adherence and reduce re-hospitalizations. Some people in the study will receive the training program, which is called the Partnership in Medication Management (PIMM) model and Self-Administered Medication (SAM) model. Other people in the study will continue being treated in the same way as they are being treated now. The people who being treated the same will be getting standard prescribing practice (SPP).

PIMM/SAM will include education to improve patients' knowledge regarding their medication's purpose, dosage, benefits, and side effects. The program will also include tools like a checklist or alarm clock to remind patients of when and how to take their medication. Furthermore, the program will contain an interactive listening period where healthcare professionals involved in medication dispensing will listen to patients' concerns, questions and thoughts regarding their medications.

How many people will be in the study?
Approximately 166 people between the ages of 18 and 60 years will take part in the study.

What will my responsibilities be if I decide to take part in the study?

If you agree to participate in the study:

You will be randomly assigned to either the PIMM/SAM group or the SPP group. If you are assigned to the PIMM/SAM group, you will meet with the nurse and your psychiatrist to discuss how and when you take your medication at home. Then, you will be assigned a drawer in the medication cart. An initial education session will be held where the nurse will review your medications, dosage and purpose (e.g., assist with sleep), benefits and side effects. Also, the nurse will provide you with tools to help you remember when and how to take your medication (e.g. alarm clock, notebook with medication information).

In the PIMM model, you will be required to notify the nurse when it is time to take your medications. Then, the nurse and you will go to the medication room. There, with nurse supervision, you will take your medication drawer, take out your medications (choosing correct bottles or blister package), and indicate to the nurse the dosage and purpose of each medication.

If you do not know the dosage or purpose of these medications, the nurse will provide you with this information and answer any questions you may have about the medication. If you forget your medication, the nurse will remind you to take it an hour after you were expected to do so. Re-education will also be provided each time a medication change occurs (e.g. dosage, new medication)

You will transition to SAM once the clinical team feels that no further medication changes are required. In the SAM model, you will be required to notify the nurse when it is time to take your medications. Also, you will be required to tell the nurse where your

medications are, dosage, purpose, and how to take them. SAM is also the model that you will follow after discharge. The nurse and you will discuss where do you keep your medication, how is the medication administered, when should the medication be taken and the use of reminders at home (e.g. alarm clock, journal, notebook with medication information). Once at home, participants are not required to call the nurse to inform her/him about their medications.

If you are assigned to the SPP group, you will receive the same standard care that the patients not participating in this study do. You will not receive the personalized medication training. The nurse will administer your medications and provide you with some information about them. Please feel free to ask the nurse any question that you have regarding your medications. You will not be provided with any tool to help you remember when to take your medications because the nurse will give you your medications as prescribed by your clinical team. Please be assured that your clinical team will provide you with the best of care.

All study participants will be required to complete questionnaires about symptoms of anxiety and depression, medication adherence, quality of life, and beliefs about medications. You will be required to fill out these questionnaires after you are assigned to either PIMM/SAM group or SPP and after you are discharge from the hospital: 1 week, 1, 3, 6 and 12 months.

What are the possible risks and discomforts?
Study participants may experience some emotional discomfort when providing responses to the study questions. You do not have to answer any question that causes you to feel uncomfortable.

What are the possible benefits for me and/or society?

You are not likely to personally benefit from participating in the study. However, the information from this study will tell us whether the PIMM/SAM program can help people take their medications as prescribed. If so, then we can see if better adherence can lower the number of re-hospitalizations from mood disorders.

Confidentiality
All hardcopy material will be stored in a locked cabinet on the premises of Mood Disorders Research Unit at St. Joseph's Healthcare Hamilton. You will be identified using a unique study identification number. This number, not your name, will be shown on all study materials related to you, except for consent form. On the consent form, your name will necessarily appear; however the unique study identification number will not appear on the consent form. A separate file linking your name and identification number will be stored in a different locked cabinet on the premises of Mood Disorders Research Unit at St. Joseph's Healthcare Hamilton. Only the principal investigators will have access to the linking file. Any data that may be published in scientific journals or tests will not reveal your identity.

To promote proper study monitoring, the St. Joseph's Healthcare Hamilton Research Ethics Board may consult your research data or medical records. However, no records that identify you by name or initials will leave the research premises.

If I do not want to take part in the study, are there other choices?
Your decision to take part in this study is voluntary. If you decide to participate, you are free to withdraw your consent and to discontinue your participation at any time without explanation or prejudice to you. Your withdrawal from this study will in no way affect the care or treatment you or your family receive at St. Joseph's Healthcare, Hamilton either now or in the future.

Will I be paid to participate in this study?
You will not be paid to participate in this study. However, you will receive a $5 Tim Horton's gift certificate each time you complete all the questionnaires.

What happens if I have a research-related injury?
If you are injured as a direct result of taking part in the study, all necessary medical treatment will be made available to you at no cost. Financial compensation for such things as lost wages, disability, or discomfort due to this type of injury is not routinely available.

If you sign this consent form, you are not waiving any legal rights you may have under the law, nor are you releasing the investigators, institution, or sponsor from their legal and professional responsibilities.

Conditions of your involvement
You must be able to follow the study instructions in English to participate.

Questions
You are encouraged to ask questions at any time during the study. If you have any questions about the study or if you are dissatisfied with the manner in which this study is being conducted, please contact Dr. Margaret McKinnon at 905-522-1155 ext 35438 or Sharon Simons RN, BScN, CPMHN(c) at 905-522-1155 ext 36738.

If you have any questions regarding your rights as a research participant, you may contact the Office of the Chair of Hamilton Integrated Research Ethics Board at 905-521-2100 ext. 42013.

RESEARCH STUDY CONSENT FORM STATEMENT AND SIGNATURE

Do not sign this consent form unless you have had a chance to ask questions and have received satisfactory answers to all of your questions.

- I have read this consent form about a study being conducted by Drs. Margaret McKinnon, Carolina Oremus and Sharon Simons RN, BScN, CPMHN(c) of St Joseph's Healthcare Hamilton/McMaster University and Dr. Mark Oremus of University of Waterloo.
- I have had the opportunity to ask questions about my involvement in this study and to receive any additional details that I requested.
- The risks and benefits have been explained to me.
- I understand that I will be given a SIGNED copy of this consent.
- I understand that taking part in this study is voluntary and that I may choose to withdraw at any time. I understand that information regarding my personal identity will be kept confidential.
- I understand that my medical records will be obtained and reviewed by Dr. Carolina Oremus, Sharon Simons RN, BScN, CPMHN(c), and Dr. Margaret McKinnon of St Joseph's Healthcare Hamilton/McMaster University, and Dr. Mark Oremus of University of Waterloo for purposes of conducting this study.
- I authorize the investigators of this study to obtain and review my medical records for the purposes of conducting this study.
- I authorize the inspection of any of my records that relate to this study by regulatory authorities such as the Research Review Committee for quality assurance purposes.
- By signing and dating this consent form, I am aware that none of my legal rights are being waived.

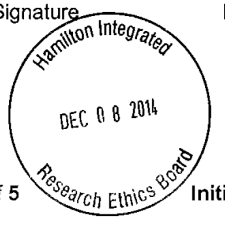
I freely consent to take part in this study.

Name of Participant (Printed)	Signature	Date

Consent form explained in person by:

Name and Role (Printed)	Signature	Date

Investigator's Name	Investigator's Signature	Date



PIMM – Steps to Inform the Patient about the Study

1. Confirm that the attending physician has determined the patient is suitable to take part in the PIMM study.

2. Provide the patient with the following information about the Partnership in Medication Management (PIMM) study:

    - The PIMM study investigates whether a novel personalized medication training (PIMM/SAM) for persons with mood disorders (depression and bipolar disorders) will help them take their medications as prescribed by their doctor.

    - The PIMM/SAM training program will include personalized education to improve persons' knowledge regarding their medication's purpose, dosage, benefits, and side effects.

    - The ultimate purpose of the PIMM/SAM training is to improve medication adherence, quality of life and decrease the number of re-hospitalizations in persons with mood disorders.

3. Ask the patients if s/he is interested in hearing more about the study. If so, provide the patient with the PIMM brochure.

4. Continue following the steps on the PIMM - Steps for Clinical Staff document.

Please write below any comment/suggestion regarding informing the patient about the PIMM study:

Many thanks,

The PIMM Research Team

PIMM - Steps for Clinical Staff

Please follow each step and answer the questions/fill in the blanks. Please add the date when the step was completed.

1. When the patient is admitted in the Mood Disorders Program (MDP) inpatient unit, the nurse most responsible for the patient will administer the MOCA.

   - MOCA completed by:_____ Date:_____

2. The attending physician will review the MOCA and decide if the patient is clinically and cognitively suitable to take part in the PIMM study.

   - Reviewed MOCA & decided patient's suitability for the study_____ Date:_____

3. If the patient is suitable, the nurse and/or attending physician will inform the patient about the Partnership in Medication Management/Self-Administered Medication (PIMM/SAM) study (please see the PIMM - Steps to Inform the Patient about the Study document [attached]) and ask whether s/he is interested in hearing more about the study.

   - Completed by_____ Date:_____

4. The nurse will contact the research team to advise them that the patient would like to hear more about the study. The primary research team contact is: Dr. Carolina Oremus (coremus@stjoes.ca; ext. 36326). Back-up coverage will be provided by Laura Garrick (x 35409).

   - Completed by_____ Date:_____

5. A research staff member will then approach the patient to describe the study and obtain informed consent for participation in the study.

   - Completed by_____ Date:_____

6. The research team member will inform the clinical team that the patient has consented to the study and will place a copy of the consent form in the patient's chart.

   - Completed by_____ Date:_____

7. The patient's nurse will obtain the participant's randomized group assignment (PIMM/SAM or standard prescribing practice [SPP]) by calling Laura Garrick (x 35409), who will be blinded to participant identifiers and diagnosis (Dr. Carolina

Oremus will conduct the data analyses and will be blinded to the participant's randomized group assignment). Laura Garrick will notify Dr. Colleen Merrifield that a new patient has been enrolled in the PIMM study.

- Completed by_____ Date:_____

8. The nurse will write the participant's randomized group assignment on the sheet provided in the chart and put it in the envelope provided in the chart. The randomized group assignment should be kept in the envelope provided to ensure that Dr. Carolina Oremus remains blinded to the patient's allocation group.

   - Completed by_____ Date:_____

9. Baseline testing for the PIMM study will be initiated. Drs. Carolina Oremus or Colleen Merrifield (back-up coverage) will schedule an appointment on the unit to administer the Mini-International Neuropsychiatric Interview (MINI) and provide the patient with copies of the self-report measures. In the event that the patient reports suicidality on the MINI and/or the Beck Depression Inventory, this is reported by research staff to the nurse most responsible for the patient. The nurse is responsible for communicating this information to the attending psychiatrist and the patient's response on the MINI and/or BDI is documented on the chart.

   - Interview completed by_____ Date:_____

10. Dr. Colleen Merrifield will be responsible for administration of the Repeatable Battery for Assessment of Neuropsychological Function (R-BANS). Dr. Colleen Merrifield will be blinded to the participant's randomized group. She will contact the unit to schedule an appointment to administer the measures and will document the results of this assessment on the patient's inpatient chart.

    - RBANS administered by_____ Date:_____

11. If the patient is assigned to the SPP group, the nurse will follow the standard prescribing practice; medication administration will proceed as usual. If your patient is assigned to the PIMM group:

<u>PIMM/SAM: First Day</u>

1. The nurse, physician and patient will discuss when and how the patient takes his/her medications at home (i.e., with a meal), how the patient's pharmacist dispenses his/her medications (i.e., blister pack), and what reminders normally work to assist him/her in remembering to take his/her medications at home.

   - Completed by_____ Date:_____

- Completed by_____ Date:_____

- Medication dispensing method_____

  _____

- Reminders:_____
  _____

2. The nurse will call the patient's community pharmacist to confirm how the patient's medications are dispensed (e.g., blister pack).

   - Completed by_____ Date:_____

3. The nurse will phone the hospital pharmacist and communicate how the patient's medications are dispensed when the patient is home. The nurse and pharmacist will ensure (whenever possible) that the medications are dispensed to the patient in the same manner as the medications would be dispensed when the patient is home.

   - Completed by_____ Date:_____

   - Pharmacist's name_____

4. The nurse will work with the patient to determine what time of day and how s/he will be taking his/her medications while in hospital.

   - Completed by_____ Date:_____

5. The nurse will meet with the patient at the medication education room to teach him/her about his/her medications, including identifying what medication s/he will be taking, the dosage of medication, what each medication is for, the importance of taking the medication as prescribed, benefits and medication side effects.

   - Completed by_____ Date:_____

6. The nurse and the patient will establish reminders (e.g., note on whiteboard in room, alarm clock, notebook with medication information) that will assist him/her in remembering to take his/her medications while in hospital. The PIMM research team will provide the patient with a notebook. The nurse will encourage or remind the patient to take notes during any teaching session.

   - Completed by_____ Date:_____

- Notebook provided by:_____ Date:_____
- Reminders_____
  _____

  _____
  _____

7. The nurse will explain the patient that starting tomorrow, s/he will take responsibility for approaching you at the established time(s) to take his/her medications.
   - Completed by_____ Date:_____

Subsequent Days

1. The patient will go to the nurse station and notify the nurse that it is time to take his/her medications. This responsibility falls to the patient and not to the nurse. If the patient forgets to take a medication, the nurse will need to remind the patient. The nurse and physician will establish the window of time for when the nurse should remind the patient to take his/her medications.

2. The time at which the patient: i) notifies you that it is time to take the medication or ii) the nurse gives a reminder will need to be charted and recorded on the PIMM -Record Sheet.

3. Once the patient notifies the nurse or the nurse reminds the patient that it is time to take the medication, the nurse and the patient will attend the teaching medication education room. The nurse will verify/ask the patient to have his/her notebook with him/her before going to the teaching medication education room.

4. At the teaching medication education room, the nurse will take out the patient's medication drawer and give it to him/her.

5. Once the patient has the medication drawer, s/he will take out his/her medications (choosing correct bottles, etc.), and indicate the dosage, purpose, benefits and side effects of each medication and identify any changes in medication or dosage. Then, the nurse will record the patients' knowledge concerning his/her medications on the patient's chart and recorded on the PIMM - Record Sheet.

6. Also, the nurse will be required to complete the PIMM Checklist for staff each time the patient takes his/her medication.

7. If the patient does not know the dosage or purpose of these medications, then the nurse will be required to provide him/her with this information and to answer any questions the patient may have about his/her medication.

8. The nurse will monitor and document on the chart and the PIMM-Record Sheet, whether or not the patient takes his/ her medication correctly. The nurse will intervene only if the patient makes an error and will then provide re-education.

9. Re-education will also be provided each time a medication change occurs. The next time a supervised medication administration occurs, the nurse will record on the chart and on the PIMM -Record Sheet, whether or not the patient recalls that a medication change has been made, the nature of the change (e.g., increase/ decrease in dosage; change in medication) and why.

SAM

1. Patients will transition to SAM once the clinical team feels that no further medication changes are required.

2. In this model, the patients will be required to notify you when it is time to take their medications, where their medications are, dosage, purpose, and how to take them.

3. SAM is also the model that the participants will follow after discharge. The nurse and the patient will discuss where does s/he keeps his/her medication, the how is the medication administered, when should the medication be taken and the use of reminders at home (e.g. alarm clock, notebook with medication information).

SPP

If the patient is assigned to the SPP group:

1. The patient will receive the same standard care that the patients not participating in this study do.

2. Study participants will not receive a personalized medication training (PIMM/SAM).

3. The nurse will administer the patient's medications. However, patients are encouraged to ask you any questions regarding his/her medications.

4. Patients in the SPP group, will not be provided with any tool to help them to remember when to take their medications. In this model, the nurse will administer the medications as prescribed by the clinical team. The nurse will record the patient's knowledge regarding his/her medications on the patient's chart and on the PIMM - Record Sheet.

Before discharge

1. A week prior to the patient's anticipated discharge, the patient's nurse will contact Dr. Carolina Oremus (coremus@stjoes.ca or x 36326) or Laura Garrick (x 35409), to notify the research team of the patient's upcoming discharge day.

   - Completed by_____ Date:_____

2. Drs. Carolina Oremus or Colleen Merrifield (back-up coverage) will schedule an appointment on the unit. The appointment will take place two days before discharge. During this appointment, Drs. Carolina Oremus or Colleen Merrifield (back-up coverage) will conduct the follow up interview and provide the patient with the PIMM package of scales. In the event that the patient reports suicidality on the Beck Depression Inventory, this is reported by research staff to the nurse most responsible for the patient. The nurse is responsible for communicating this information to the attending psychiatrist and the patient's response on the BDI is documented on the chart.

   - Completed by_____ Date:_____

Broad overview: In addition to measuring patient's ability to recall the time at which their medication should be taken (at approximately the same time each day; time determined by the patient), the PIMM research team will also administer the MINI and a package of scales that will measure cognition, depression and anxiety symptoms, quality of life, medication adherence, beliefs about medication, general self-efficacy, and some aspects of the psychiatrist-patient relationship. The PIMM research team will administer these instruments at baseline, two days before discharge and at the following post-discharge times: one week, one month, three months, six months, and 12 months. The PIMM research team will also keep track of the reminders that patients use to help them remember to take their medications and use this information to build reminders into our program (e.g., alarms, white board, notebooks, etc.).

"The ultimate purpose of the PIMM/SAM training is to improve medication adherence in persons with mood disorders."

What are the possible benefits for me and/or society?

The information from this study will tell us whether the PIMM/SAM training can help persons with mood disorders take their medications as prescribed. If so, then we can see if better adherence improves persons' quality of life and lowers the number of re-hospitalizations from mood disorders.

"In the PIMM/SAM training, participants, physicians and nurses work together to establish how and when participants take their medications and the reminders that will help participants improve medication adherence."

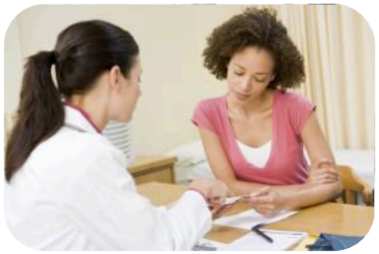
Partnership in Medication Management

PIMM

McMaster University
Inspiring Innovation and Discovery

MDP Mood Disorders Program

St. Joseph's Healthcare Hamilton



PIMM: Partnership in Medication Management

The PIMM study investigates if a novel one-on-one medication training for persons with mood disorders (depression and bipolar disorders) will help them take their medications as prescribed by their doctor. People participating in the study will be placed in either the one-on-one medication training group called PIMM/SAM or the usual care group called standard prescribing practice (SPP). The computer will randomly choose who gets PIMM/SAM or SPP.

"The PIMM/SAM training contains an interactive listening period where healthcare providers involved in medication dispensing listen to participants' concerns, questions and thoughts regarding their medications".

Why is this research being done?

Drug treatments like anti-depressants or lithium often do not have the desired results in persons with mood disorders. Sometimes the lack of results happens because people do not take their medications as prescribed. When people take their medications as prescribed is called 'adherence'.

PIMM/SAM

The nurse and the attending physician will meet with the participant and ask how s/he takes his/her medication at home (i.e.., blister pack), at what time (e.g., breakfast, dinner) and what reminders (i.e., alarm clock) usually work to help him/her remember to take his/her medications at home. An initial education session will be held where the nurse will review the participant's medications, dosage and purpose (e.g., assist with sleep) such that participants will be aware of what medications they are taking and why.
The PIMM/SAM training includes education to improve the participants' knowledge regarding their medication's purpose, dosage, benefits, and side effects. The PIMM/SAM training also includes tools like a checklist or notebook to remind participants of when and how to take their medication, the purpose of the medication and the benefits of taking their medication as prescribed.

SPP

In this group, participants will receive the same standard care that the patients not participating in the PIMM study do. Participants will not receive the PIMM/SAM training.

Version #1, Nov. 17, 2014.

PIMM Checklist – Goals for Study Participants

1. Working with my clinical team to discuss when and how I take my medications at home (e.g., blister pack) and what reminders normally work to assist me in remembering to take my medications at home.

2. Meet with my clinical team to learn about my medications, including identifying what medication I will be taking, the dosage of medication, what each medication is for, and medication side effects.

3. Working with my clinical team to determine what time of day and how I will be taking my medications while in hospital

4. Working with my clinical team to establish reminders (e.g., note on whiteboard in room, alarm clock, notebook with medication information) that will assist me in remembering to take my medications while in hospital.

5. Follow the steps of the PIMM Checklist everyday

6. Taking personal responsibility for approaching my nurse at the established time(s) to take my medications

7. Taking personal responsibility for notifying to my nurse, each time I take my medication:

    i) What medications I am taking

    ii) Why I am taking each medication

    iii) How and when I should take my medication

    iv) If any, ask the nurse questions regarding my medication

PIMM Checklist – Steps for Study Participants

1. Go to the nurse station and notify the nurse that it is time to take your medication
2. With the nurse, go to the teaching medication education room
3. Take your medication drawer and identify the medication you need to take
4. Discuss how and when to take this medication
5. Discuss the needed medications dosage and purpose
6. Discuss the medications side effects
7. If any, discuss any changes in your medication
8. Discuss any further concerns with your nurse
9. Take medication
10. Use your notebook, for any concerns, questions or reminders about your medication
11. Prepare your reminders (e.g., note, alarm clock) for your next medication time

PIMM Document	Administered by:	Date:
1. PIMM - Steps for Clinical Staff		
1. PIMM – Steps to Inform the Patient about the Study		
2. PIMM Checklist – Goals and Steps for Study Participants		
3. PIMM-Record Sheet		
Scale	**Administered by:**	**Date:**
1. Montreal Cognitive Assessment (MOCA)		
2. Mini-International Neuropsychiatric Interview (MINI)		
3. Beck Depression Inventory-II (BDI-II)		
4. Beck Anxiety Inventory (BAI)		
5. Medication Adherence Rating Scale (MARS)		
6. Beliefs about Medicines Questionnaire (BMQ)		
7. Short Form 36-Health Survey (SF-36)		
8. Revised Helping Alliance Questionnaire for Treatment with Psychiatrists (HAQ-PC)		
9. General Self-Efficacy Scale (GSE)		
10. Multiscale Dissociation Inventory (MDI)		
11. Repeatable Battery for the Assessment of Neuropsychological Functioning (RBANS)		
Extras	**Administered by:**	**Date:**
PIMM-Brochure		
Group Allocation		
Envelope		

Date & Time; Completed by:	Notified at Time	Nurse Reminded Participant	Medication Knowledge	Medication Changes	Comments
	☐ Yes ☐ Late _______ mins. ☐ No Comments:	☐ Yes ☐ No Comments:	☐ Identified medication ☐ Knows dosage ☐ Knows purpose ☐ Importance of taking medication as prescribed ☐ Benefits ☐ Side effects Comments:	☐ Yes (explain below) ☐ No Comments:	
	☐ Yes ☐ Late _______ mins. ☐ No Comments:	☐ Yes ☐ No Comments:	☐ Identified medication ☐ Knows dosage ☐ Knows purpose ☐ Importance of taking medication as prescribed ☐ Benefits ☐ Side effects Comments:	☐ Yes (explain below) ☐ No Comments:	
	☐ Yes ☐ Late _______ mins. ☐ No Comments:	☐ Yes ☐ No Comments:	☐ Identified medication ☐ Knows dosage ☐ Knows purpose ☐ Importance of taking medication as prescribed ☐ Benefits ☐ Side effects Comments:	☐ Yes (explain below) ☐ No Comments:	

END OF THESIS