

ISSUES RELATED TO MANAGING DISABLED CLAIMANTS

ISSUES RELATED TO DETERMINING OPTIMAL MANAGEMENT
OF PATIENTS IN RECEIPT OF DISABILITY BENEFITS

By

SHANIL EBRAHIM, M.Sc. M.Sc. B.Sc.

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

McMaster University

© Copyright by Shanil Ebrahim, 2013

McMaster University

DOCTOR OF PHILOSOPHY (2013)

Hamilton, Ontario (Health Research Methodology)

TITLE: Issues related to determining optimal management of patients in receipt of disability benefits

AUTHOR: Shanil Ebrahim, M.Sc. M.Sc. B.Sc. (McMaster University)

SUPERVISOR: Dr. Gordon H. Guyatt

NUMBER OF PAGES: 198

ABSTRACT

Approximately 4.2 million Canadian adults suffer from a physical or psychological disability, of whom up to 30% suffer from depression. Those receiving disability benefits versus those not receiving benefits may be at greater risk of unsatisfactory outcomes because their circumstances or psychological status may interfere with successful implementation of standard therapies. This thesis addresses the effectiveness of therapies for depression in patients receiving disability benefits, using an individual patient data meta-analysis of all published randomized controlled trials evaluating Cognitive Behavioural Therapy and a secondary analysis of an administrative database from a large, private, Canadian insurer. Additionally, this thesis addresses an important methodological issue: assessing the impact of missing participant data for continuous outcomes in systematic reviews. Missing participant data may bias results of individual trials or systematic reviews of individual trials if participants with missing data have different expected outcomes from those with available data. No methods have been proposed for investigating the extent to which missing participant data for continuous outcomes might bias the results of systematic reviews, and this thesis addresses that gap.

ACKNOWLEDGMENTS

Dissertation supervisor and mentor

I want to thank Dr. Gordon Guyatt for his outstanding mentorship, his strong commitment to teaching and supervising, and for being such an affable individual. His passion for research has been infectious. You have helped me grow both professionally and personally, and I am forever indebted.

Committee members

Thank you to Dr. Jason Busse for his tremendous support and feedback, and for always bringing his witty humour to our meetings, and Dr. Stephen Walter for his immense support and guidance with statistical issues. Thank you both for providing me exceptional advice.

Co-authors

Thank you to all the wonderful co-authors that I had the privilege to work with in all my projects. I could not have completed this thesis without your help.

Funding agencies

Thank you to the Canadian Institutes of Health Research, MITACS Accelerate, the Government of Ontario QEIIIGSST program, and McMaster University for their

assistance with my salary and research support to undertake my PhD. This provided me a valuable lesson—a “poor grad student” lifestyle is not a prerequisite for completing a PhD.

Health Research Methodology program

Thank you to all the wonderful people in the HRM program, ranging from the students to the faculty to the administrative team. This is an exceptional program because of all of you, and I am glad to have been part of it.

Fiancée, family and friends

Thank you to my fiancée, Sheena, who has supported me through all the milestones in my life; my love for you is immeasurable. Thanks to my family (my dearest mama, papa, Shyna, Zamreen, Rasheed and my two adorable nieces, Naima and Zayna), my best friend Aria, and all my friends for their constant love and encouragement. You have all made me the person I am today, and I am eternally grateful.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF APPENDICES	xiv
LIST OF ABBREVIATIONS	xv
DECLARATION OF ACADEMIC ACHIEVEMENT	xvii
THESIS OBJECTIVES	1
CHAPTER 1: OBJECTIVES.....	2
REFERENCES.....	7
STUDY 1: Effectiveness of Cognitive Behavioral Therapy for depression in patients receiving disability benefits: A systematic review and individual patient data meta-analysis	10
CHAPTER 2: ABSTRACT.....	12
CHAPTER 3: INTRODUCTION.....	14
CHAPTER 4: OBJECTIVES AND QUESTIONS.....	17
4.1 Objectives.....	17
4.2 Questions.....	17
CHAPTER 5: METHODS.....	18
5.1 Protocol and registration.....	18
5.2 Eligibility criteria.....	18
5.3 Information sources.....	19

5.4	Search.....	19
5.5	Study selection.....	19
5.6	Contacting authors of eligible studies.....	20
5.7	Data collection process.....	21
5.8	Risk of bias within studies.....	22
5.9	Synthesis of results.....	23
CHAPTER 6: RESULTS.....		26
6.1	Study selection.....	26
6.2	Study characteristics.....	27
6.3	Risk of bias within studies.....	27
6.4	IPDMA.....	27
6.5	Secondary analyses.....	34
CHAPTER 7: DISCUSSION.....		39
7.1	Summary of evidence.....	39
7.2	Strengths.....	39
7.3	Previous Evidence.....	40
7.4	Limitations.....	41
7.5	Conclusions.....	42
ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS.....		43
REFERENCES.....		45

STUDY 2: Association of psychotherapy with disability benefit claim

closure among patients disabled due to depression.....	53
CHAPTER 8: ABSTRACT.....	54
CHAPTER 9: INTRODUCTION.....	56
9.1 Background.....	56
9.2 Objectives.....	58
CHAPTER 10: METHODS.....	59

10.1	Ethics statement.....	59
10.2	Design.....	59
10.3	Description of patients and eligibility criteria.....	59
10.4	Administrative variables.....	60
10.5	Outcomes.....	63
10.6	Data management and data cleaning.....	63
10.7	Statistical analysis.....	63
CHAPTER 11: RESULTS.....		67
11.1	Description of claims.....	67
11.2	Short-term disability.....	70
	11.2.1 <i>Factors associated with time to STD claim closure</i>	70
	11.2.2 <i>Factors predictive of receipt of psychotherapy among STD claims</i>	72
11.3	Long-term disability.....	74
	11.3.1 <i>Factors associated with time to LTD claim closure</i>	74
	11.3.2 <i>Factors predictive of receipt of psychotherapy among LTD claims</i>	77
CHAPTER 12: DISCUSSION.....		79
12.1	Summary of main results.....	79
12.2	Strengths and limitations.....	79
12.3	Findings in context with previous evidence.....	81
12.4	Future directions.....	83
CHAPTER 13: CONCLUSIONS.....		84
ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS.....		85
REFERENCES.....		87

STUDY 3: Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers.....	92
CHAPTER 14: ABSTRACT.....	93
CHAPTER 15: WHAT THIS STUDY ADDS?	94

CHAPTER 16: INTRODUCTION.....	95
CHAPTER 17: METHODS.....	98
17.1 Development of approach.....	98
17.2 Imputing measure of effect.....	99
17.3 Imputing measure of precision.....	100
17.4 Combining observed and imputed data.....	102
CHAPTER 18: APPLICATION OF APPROACH.....	104
18.1 CBT for depression in patients receiving disability benefits.....	104
18.2 Finasteride for improvement in scalp hair for men with androgenetic alopecia.....	106
CHAPTER 19: DISCUSSION.....	108
19.1 Summary of findings.....	108
19.2 Implications on risk of bias.....	108
19.3 Properties.....	109
19.4 Strengths, limitations and other issues.....	110
19.5 Conclusions.....	114
ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS.....	115
REFERENCES.....	117

STUDY 4: Addressing missing participant data in systematic reviews:

continuous outcomes assessed with different instruments.....	121
CHAPTER 20: ABSTRACT.....	122
CHAPTER 21: INTRODUCTION.....	124
CHAPTER 22: METHODS.....	126
CHAPTER 23: RESULTS.....	128
23.1 Choosing the reference measurement instrument.....	128
23.2 Converting scores from different instruments to the reference instrument.....	129

23.3	Imputing measure of effect and its precision.....	130
23.4	Combining observed and imputed continuous data.....	131
23.5	Enhancing interpretability using MID.....	132
23.6	Judging impact of missing participant data on quality of evidence...	133
23.7	Application of approach.....	135
CHAPTER 24: DISCUSSION.....		141
24.1	Summary of findings.....	141
24.2	Implications on quality of evidence.....	141
24.3	Strengths, limitations and other issues.....	143
24.4	Conclusions.....	146
ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS.....		147
REFERENCES.....		149
THESIS CONCLUSIONS.....		156
CHAPTER 25: CONCLUSIONS.....		157
25.1	Overview.....	157
25.2	Management of patients in receipt of disability benefits.....	157
25.3	Addressing missing participant data for continuous outcomes in systematic reviews.....	162
APPENDICES.....		165

LIST OF TABLES

Table 6.1:	Characteristics of CBT studies.....	29
Table 6.2:	CBT details from studies.....	31
Table 10.1:	Description of variables.....	62
Table 10.2:	Classification of industry.....	64
Table 11.1:	Baseline characteristics.....	67
Table 11.2:	Factors predictive of time to short-term disability claim closure.....	73
Table 11.3:	Factors predictive of time to long-term disability claim closure.....	76
Table 11.4:	Comparison between predictors associated with time to claim closure for short-term disability versus long-term disability claims...	78
Table 17.1:	Matrix of assumptions for participants with missing data for continuous outcomes in intervention and control arms.....	101
Table 23.1:	Assumptions for participants with missing data for continuous outcomes.....	132

LIST OF FIGURES

Figure 6.1:	Flow chart of study eligibility.....	28
Figure 6.2:	Risk of bias within studies.....	33
Figure 6.3:	Effect of CBT in patients not receiving disability benefits in studies including patients receiving benefits versus those that did not.....	35
Figure 6.4:	Effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT plus pharmacotherapy versus pharmacotherapy alone.....	36
Figure 6.5:	Effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT versus TAU/standard care.....	38
Figure 11.1:	Kaplan Meier time to event curve of time to STD claim closure.....	71
Figure 11.2:	Kaplan Meier time to event curve of time to LTD claim closure.....	75
Figure 18.1:	Forest plots of the complete case analysis and sensitivity analyses for handling participants with missing data for continuous outcomes in a systematic review evaluating CBT for depression in patients receiving disability benefits.....	105
Figure 18.2:	Forest plots of the complete case analysis and sensitivity analyses for handling participants with missing data for continuous outcomes in a systematic review evaluating finasteride therapy for androgenetic alopecia.....	107
Figure 23.1:	Complete case analysis and sensitivity analyses using the four strategies for handling participants with missing data for continuous outcomes in a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease.....	137
Figure 23.2:	Interpreting important treatment effects using the four strategies for handling participants with missing data for continuous outcomes in a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease.....	139

Figure 23.2: Interpreting important treatment effects using different thresholds 140
in a systematic review of respiratory rehabilitation in Chronic
Obstructive Pulmonary Disease

LIST OF APPENDICES

Appendix A:	Search strategies of the IPDMA of CBT for depression in patients receiving disability benefits.....	166
Appendix B:	Published Article: IPDMA of CBT for depression in patients receiving disability benefits.....	168
Appendix C:	Description of all variables in administrative database.....	177

LIST OF ABBREVIATIONS

ASO	Administrative Services Only
BDI-II	Beck Depression Inventory (II)
BOCF	Baseline Observation Carried Forward
BSI	Bradford Somatic Inventory
CBT	Cognitive Behavioural Therapy
CI	Confidence Interval
DSM	Diagnostic and Statistical Manual of Mental Disorders
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HADS	Hospital Anxiety and Depression Scale
HR	Hazard Ratio
HRSD	Hamilton Rating Scale for Depression
ICD	International Classification of Diseases
IPD	Individual Patient Data
IPDMA	Individual Patient Data Meta-Analysis
IQR	Interquartile Range
LOCF	Last Observation Carried Forward
LTD	Long-term Disability
MD	Mean Difference
MESH	Medical Subject Headings
MID	Minimally Important Difference
N/A	Not Applicable

NICE	National Institute for Health and Clinical Excellence
NR	Not Reported
NS	Not Significant
OR	Odds Ratio
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomized Controlled Trial
RDC	Research Diagnostic Criteria
RTW	Return to Work
SD	Standard Deviation
SMD	Standardized Mean Difference
STD	Short-term Disability
TAU	Treatment as Usual

DECLARATION OF ACADEMIC ACHIEVEMENT

I was the main contributor and first author for all studies. The details of my and other authors' contributions are provided at the end of each study.

THESIS OBJECTIVES

CHAPTER 1

OBJECTIVES

Fourteen percent (4.2 million) of Canadian adults reported that they suffered from a physical or psychological disability in 2006 [1], of whom approximately 30% suffer from depression [2]. Depression accounts for an estimated productivity loss of \$14.4 billion annually in Canada [3,4,5,6], and the World Health Organization (WHO) estimates that depression will become the second leading cause of disease burden worldwide by the year 2020 [7,8,9].

Based on the level of impairment, individuals suffering from depression may be eligible to receive financial compensation. In Canada and most industrialized countries, individuals suffering from depression are typically insured for wages lost as a result of being off work. Most of the individuals who are accepted for wage replacement benefits recover and return to work in a timely manner. However, approximately 10% of individuals receiving benefits do not return to work, and they account for 65% to 75% of resources spent on disability claims [10,11,12,13]. Compared to other disorders, depression incurs more costs for long-term disability (LTD) [14], requires more complex treatment [15], and individuals have more difficulty returning to work [15]. Despite the magnitude of the problem of depression

and disability, the effectiveness of treatments for depression in patients receiving disability benefits has received little attention.

Cognitive Behavioural Therapy (CBT) is one of the most common non-pharmacological methods of treating depression [16,17]. CBT may be less effective in subpopulations of patients whose circumstances or psychological status interferes with the successful implementation of CBT [18]. There has been limited attention on the effectiveness of CBT in patients with depression who are in receipt of wage replacement benefits. Given that depression is the most frequent and costly mental illness for insurers [19], and CBT is one of the most frequently reimbursed therapies for depression by private insurances [20], it is important to ascertain if CBT represents a worthwhile expenditure of time and energy for depressed patients, and a good investment for the insurance industry and society in general.

This thesis is based on a series of papers, published or submitted for publication, addressing issues related to determining optimal management of patients in receipt of disability benefits. The first two papers focus on the effectiveness of CBT and psychotherapy in general in managing patients in receipt of disability benefits secondary to depression.

Chapters 2 to 7 correspond to different sections of the first paper (published in PLoS ONE) [21], which presents a systematic review and an individual patient data meta-analysis (IPDMA) of all randomized controlled trials evaluating the effectiveness of CBT for depression in patients receiving disability benefits. Our primary question was: does the effect of CBT vary by disability status? This was a challenging review to conduct, as we first identified no studies that addressed this question. We contacted all the authors of the potentially eligible trials evaluating CBT to inquire if any of them captured disability benefit status but did not report it in their published report. A few did. We were able to successfully obtain their data, which helped us conduct a more powerful individual patient data meta-analysis to inform our question.

Chapters 8 to 13 correspond to different sections of the second paper, which presents a secondary analysis of an administrative database from a large Canadian private disability insurer to evaluate the association between the provision of psychotherapy and other potential predictors, and short-term disability (STD) and long-term disability (LTD) claim closure. We analyzed 4 years of data consisting of 20,846 claimants with depression. This was the largest study of its kind to explore this question.

This thesis also provided an opportunity to address important methodological issues: addressing the impact of missing participant data for continuous outcomes in

systematic reviews. Randomized controlled trials often exclude populations receiving disability benefits as patients' disability largely prevent participation in studies that require frequent clinic visit for therapy or treatment. When these patients are included, a significant number may drop out of studies due to various reasons. For example, in my first paper, the systematic review and IPDMA evaluating effectiveness of CBT for depression in patients receiving disability benefits, the median missing participant data rate was 21%, with one study having 41% missing participant data. If participants with missing data have different expected outcomes from those with available data, it may introduce bias in the results of the individual trials and of systematic reviews using those results.

There are currently no methods [proposed] for investigating the extent to which missing participant data for continuous outcomes may bias the results of systematic reviews. The third and fourth papers in this thesis address this gap. Although this issue arose when pooling depression outcomes (Beck Depression Inventory II) for my IPDMA, our approach is generalizable to systematic reviews pooling positive trials (i.e., those with a significant treatment effect) with continuous data across all clinical disciplines.

Chapters 14 to 19 correspond to the different sections of the third paper, a guide for systematic reviewers for addressing continuous data for participants excluded from

trial analysis. This guidance involves an initial complete case analysis (analyzing those with available data) and a series of sensitivity analyses making progressively more stringent assumptions about results of patients with missing data. Our approach also addresses the implications of findings to inferences regarding risk of bias, and is compatible with the GRADE/Cochrane handbook guidance for addressing confidence in estimates [22,23].

Chapters 20 to 23 correspond to the different sections of the fourth paper, which addresses an important limitation of our proposed approach in the third paper: restriction to systematic reviews in which each trial used the same instrument to measure continuous outcome data. We extend our approach to systematic reviews pooling trials using different instruments measuring the same construct. We provide detailed guidance on how systematic review authors can first convert scores from all instruments to units on the most familiar and responsive instrument, and then apply the imputation strategies to address missing participant data. As well, we have provided guidance on how to enhance interpretability by calculating important treatment effects from the derived estimates using the minimally important difference (MID), the smallest difference that patients experience as important.

Chapter 24 summarizes the most important findings, addresses limitations, and discusses future directions from this work.

REFERENCES

1. Statistics Canada (2006) Participation and Activity Limitation Survey: Analytical Report.
2. Ebrahim S, Guyatt GH, Walter SD, Heels-Ansdell D, Bellman M, Hanna SE, Busse JW (2012) Association of psychotherapy with disability benefit claim closure among patients disabled due to depression (under review by PLoS ONE). Hamilton: McMaster University.
3. Hirschfeld R, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, Moller HJ, Healy D, Baldwin D, Humble M, Versiani M, Montenegro R, Bourgeois M (2000) Social functioning in depression: a review. *J Clin Psychiatry* 61: 268-275.
4. Lim K, Jacobs P, Ohinmaa A, Schopflocher D, Dewa CS (2008) A new population-based measure of the economic burden of mental illness in Canada. *Chronic Diseases in Canada* 28: 92-98.
5. Moore R, Mao Y, Zhang J, Clarke K, Laboratory Centre for Disease Control (1997) *Economic Burden of Illness in Canada*. Ottawa: Health Canada.
6. Stephens T, Joubert N (2001) The economic burden of mental problems in Canada. *Chronic Diseases in Canada* 22: 1-10.
7. Brown P (2001) Effective treatments for mental illness not being used, WHO says. *BMJ* 323: 769.
8. Mor N, Haran D (2009) Cognitive-behavioral therapy for depression. *The Israel Journal of Psychiatry and Related Sciences* 46: 269-273.

9. World Health Organization (WHO) (2002) *Mental Health: New understanding, New hope*. Geneva: WHO.
10. Abenhaim L, Suissa S (1997) Importance and economic burden of occupational back pain: a study of 2500 cases representative of Quebec. *J Occup Med* 29: 670-674.
11. Frymoyer JW, Cats-Baril WL (1991) An overview of the incidences and costs of low back pain. *Orthop Clin North Am* 22: 263-271.
12. Hashemi L, Webster BS, Clancy EA (1998) Trends in disability duration and cost of workers' compensation low back pain claims (1988-1996). *J Occup Environ Med* 40: 1110-1119.
13. NACHEMSON A (1994) Chronic pain. The end of the welfare state. *Quality of Life Research* 3: 1-3.
14. Salkever DS, Goldman H, Purushothaman M, Shinogle J (2000) Disability management, employee health and fringe benefits, and long-term-disability claims for mental disorders: an empirical exploration. *The Milbank Quarterly* 78: 79-113.
15. Dewa C, Hoch JS, Goering P, et al (2003) Use of antidepressants among Canadian workers receiving depression-related short-term disability benefits. *Psychiatr Serv* 54: 724-729.
16. Dobson KS (2001) *Handbook of cognitive-behavioral therapies*. 2nd Edition. New York; London: Guilford Press.

17. NICE (2009) Depression: Treatment management of depression in adults. London: National Institute for Clinical Excellence.
18. (2003) In: Leahy RL, editor. Roadblocks in Cognitive-Behavioral Therapy: Transforming Challenges into Opportunities for Change. New York: The Guilford Press.
19. McGuire TG (1995) Effectiveness, patient matching, and insurance coverage for depression. *Soc Psychiatry Psychiatr Epidemiol* 30: 240-243.
20. Bellman M (2010) Effectiveness of CBT on individuals receiving disability benefits. Personal communication ed. Toronto.
21. Ebrahim S, Montoya L, Truong W, Hsu S, Kamal el Din M, Carrasco-Labra A, Busse JW, Walter S, Heels-Ansdell D, Couban R, Patelis-Siotis I, Bellman M, de Graaf LE, Dozois D, Bieling P, Guyatt GH (2012) Effectiveness of Cognitive Behavioral Therapy for depression in patients receiving disability benefits: A systematic review and individual patient meta-analysis. *PLoS ONE* 7: e50202.
22. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336: 924-926.
23. The Cochrane Collaboration (2011) Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0)*. Oxford, UK: Cochrane Collaboration.

STUDY 1

Effectiveness of Cognitive Behavioral Therapy for depression in patients receiving disability benefits: A systematic review and individual patient data meta-analysis

Shanil Ebrahim^{1§}, Luis Montoya², Wanda Truong³, Sandy Hsu⁴, Mostafa Kamal el Din⁵, Alonso Carrasco-Labra^{1,6}, Jason W. Busse^{1,7}, Stephen D. Walter^{1,8}, Diane Heels-Ansdell¹, Rachel Couban⁷, Irene Patelis-Siotis^{9,10}, Marg Bellman¹¹, L. Esther de Graaf¹², David J. A. Dozois¹³, Peter J. Bieling⁹, Gordon H. Guyatt^{1,14}

¹Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada

²Department of Dentistry, Santo Tomas University, Bogota D.C., Colombia

³McMaster Integrative Neuroscience Discovery & Study program, McMaster University, Hamilton, Canada

⁴Department of Oncology, McMaster University, Hamilton, Canada

⁵Ain Shams University, Faculty of Medicine, Cairo, Egypt

⁶Evidence-Based Dentistry Unit, Faculty of Dentistry, University of Chile, Santiago, Chile

⁷Institute for Work and Health, Toronto, Ontario, Canada

⁸Department of Mathematics and Statistics, McMaster University, Hamilton, Canada

⁹Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, Canada

¹⁰Mood Disorders program, St. Joseph's Healthcare, Hamilton, Canada

¹¹National Disability Services, Policy & Procedure department, Sun Life Financial, Toronto, Canada

¹²Department of Medical Psychology and Psychotherapy, Erasmus MC University Medical Center, Rotterdam, the Netherlands

¹³Department of Psychology, University of Western Ontario, London, Canada

¹⁴Department of Medicine, McMaster University, Hamilton, Canada

§Corresponding author

Correspondence to:

Shanil Ebrahim
Clinical Epidemiology and Biostatistics, McMaster University
1200 Main Street West
Hamilton, Ontario, Canada
L8S 4K1
shanil.ebrahim@utoronto.ca

Keywords

Cognitive Behavioral Therapy, Disability benefits, Depression, Systematic review, Meta-analysis, Individual patient data meta-analysis

Published Article

Ebrahim S, Montoya L, Hsu S, Kamal el Din M, Truong W, Carrasco-Labra A, Busse JW, Walter SD, Heels-Ansdell D, Couban R, Patelis-Siotis I, Bellman M, de Graaf LE, Dozois D, Bieling P, Guyatt GH. Effectiveness of Cognitive Behavioural Therapy for depression in patients receiving disability benefits: A systematic review and individual patient data meta-analysis. PLoS ONE. 2012;7(11):e50202 (Attached in Appendix B).

CHAPTER 2

ABSTRACT

Objectives: To systematically summarize the randomized trial evidence regarding the relative effectiveness of cognitive behavioural therapy (CBT) in patients with depression in receipt of disability benefits in comparison to those not receiving disability benefits.

Data Sources: All relevant RCTs from a database of randomized controlled and comparative studies examining the effects of psychotherapy for adult depression (<http://www.evidencebasedpsychotherapies.org>), electronic databases (MEDLINE, EMBASE, PSYCINFO, AMED, CINAHL and CENTRAL) to June 2011, and bibliographies of all relevant articles.

Study Eligibility Criteria, participants and interventions: Adult patients with major depression, randomly assigned to CBT versus minimal/no treatment or care-as-usual.

Study Appraisal and synthesis methods: Three teams of reviewers, independently and in duplicate, completed title and abstract screening, full text review and data extraction. We performed an individual patient data meta-analysis to summarize data.

Results: Of 92 eligible trials, 70 provided author contact information; of these 56 (80%) were successfully contacted to establish if they captured receipt of benefits as a

baseline characteristic; 8 recorded benefit status, and 3 enrolled some patients in receipt of benefits, of which 2 provided individual patient data. Including both patients receiving and not receiving disability benefits, 2 trials (227 patients) suggested a possible reduction in depression with CBT, as measured by the Beck Depression Inventory-II, mean difference [MD] (95% confidence interval [CI]) = -2.61 (-5.28, 0.07), $p=0.06$; minimally important difference of 5. The effect appeared larger, though not significantly, in those in receipt of benefits (34 patients) versus not receiving benefits (193 patients); MD (95% CI) = -4.46 (-12.21, 3.30), $p=0.26$.

Conclusions: Our data does not support the hypothesis that CBT has smaller effects in depressed patients receiving disability benefits versus other patients. Given that the confidence interval is wide, a decreased effect is still possible, though if the difference exists, it is likely to be small.

CHAPTER 3

INTRODUCTION

Major Depressive Disorder (henceforth referred to as depression) results in immense human suffering and an enormous socioeconomic burden. Depression accounts for 11% of disability worldwide and an estimated productivity loss of \$17 to \$44 billion in the USA [1,2]. Depression is expected to become the second leading cause of disease burden worldwide by the year 2020 [3].

The National Institute for Health and Clinical Excellence (NICE) in the UK recommends that health care professionals provide pharmacological treatments and/or high-intensity psychological interventions for individuals suffering from depression. Pharmacological treatments may accelerate recovery from depression, particularly when symptoms are severe [4] and, over the last few decades, their use has increased dramatically in Western nations [5,6]. NICE guidelines suggest psychological therapies should be offered to individuals suffering from persistent subthreshold symptoms of depression, mild to moderate depression, and those with a high risk of relapse or those declining pharmacological treatment for severe depression [5,6].

Cognitive Behavioral Therapy (CBT) is a common non-pharmacological treatment for depression [5,7]. CBT is based on three fundamental propositions: cognitive activity affects behavior, cognitive activity can be monitored and altered, and desired behavior change may be affected through cognitive change [7]. Twelve systematic reviews evaluating CBT in individuals suffering from depression have demonstrated that CBT reduces depressive symptoms [8,9,10,11,12,13,14,15,16,17,18,19], with the most current and rigorous meta-analysis reporting a pooled standardized mean difference (SMD) of 0.69 (95% confidence interval [CI] of 0.59 to 0.79) [13].

In North America, depression is one of the most frequent reasons for receiving disability benefits [20,21], and disability claims for mental health disorders incur greater costs compared to other disorders [22]. In those receiving disability benefits, individuals suffering from mental health disorders require more treatment and have greater difficulty returning to work than those suffering from other conditions [23]. Although CBT is one of the most frequently reimbursed therapies by insurers, its utilization by insurance companies still remains relatively low at approximately 3% for short-term disability claimants and 15% for long-term disability claimants [24].

CBT may be less effective, or ineffective, in patients receiving disability benefits, because their circumstances or psychological status may interfere with its successful implementation [25]. This may also be associated with the compensation process [26],

secondary gain from financial benefits (benefits of assuming a sick role) [27], or the adversarial nature of litigation [28]. A recent meta-analysis of 129 studies in surgical populations that found a substantially greater risk of an unsatisfactory outcome (functional, quality of life, pain and patient satisfaction) after surgery in compensated patients (odds ratio [95% CI] = 3.79 [3.28 to 4.37]) provides indirect evidence for this hypothesis [29]. The effectiveness of CBT for depression in patients receiving disability benefits has received little attention.

CHAPTER 4

OBJECTIVES AND QUESTIONS

4.1 Objectives

The purpose of our study was to perform a systematic review and an individual patient data meta-analysis of all randomized controlled trials (RCTs) that compared the effectiveness of CBT to minimal/no treatment, or care-as-usual, in patients with depression receiving versus those not receiving disability benefits.

4.2 Questions

In adult patients with depression, is there a difference in the effect of CBT on depression between those receiving disability benefits compared to those not receiving disability benefits?

CHAPTER 5

METHODS

We used the PRISMA guidelines [30] to report our findings.

5.1 Protocol and registration

We developed a protocol prior to conducting the study but did not register it.

5.2 Eligibility criteria

Eligible studies met the following criteria: 1) random allocation of adult patients to CBT or a control arm consisting of minimal/no treatment, treatment as usual (TAU) or pharmacotherapy if it was equally balanced in the treatment groups (e.g. CBT plus pharmacotherapy versus pharmacotherapy alone), and 2) inclusion of patients with depression, classified as Major Depressive Disorder by any edition of the Diagnostic and Statistical Manual (DSM), International Classification of Diseases (ICD), Research Diagnostic Criteria (RDC) or other diagnostic system [31].

5.3 *Information sources*

We identified all relevant RCTs from a database of randomized controlled trials and comparative studies examining the effects of psychotherapy for adult depression (<http://www.evidencebasedpsychotherapies.org>) [32]. This database consisted of 281 trials and was identified from searching the following electronic databases in all languages: PUBMED, EMBASE, PsycINFO and Cochrane Central Register of Controlled Trials, from inception until January 1, 2011 [11]. In addition to the 281 trials, we updated the search with the assistance of an experienced academic librarian (RC) until June 13, 2011 for each electronic database, and also searched AMED and CINAHL. We hand searched the reference lists of all relevant RCTs for additional eligible trials.

5.4 *Search*

Appendix A presents our search strategy including keywords and MESH headings.

5.5 *Study selection*

Three teams of reviewers (SE, SH, LM, WT, MK, ACL) worked in pairs and screened titles and abstracts of identified citations, independently and in duplicate, using a standardized, pilot-tested screening form. The same reviewers independently applied eligibility criteria to the full text of potentially eligible studies. One psychiatrist (IPS) and one psychologist (RM), blinded to study results, independently reviewed and

confirmed eligibility of therapies that were not explicitly described by trial authors as CBT. We measured agreement for the full text review stage, and interpreted the agreement statistics using the guidelines proposed by Landis and Koch [33]. Kappa values of 0 to 0.20 represented slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement, and greater than 0.80 almost perfect agreement.

Reviewers grouped eligible articles into one of the four categories: (i) studies that did not explicitly state if they included or excluded patients receiving disability benefits, (ii) studies that explicitly excluded patients receiving disability benefits, (iii) studies that explicitly included patients receiving disability benefits but did not separately report outcomes based on receipt of disability benefits, and (iv) studies that explicitly included patients receiving disability benefits and reported outcomes separately based on receipt of disability benefits. Disability benefits were defined as wage replacement benefits administered by a third party (e.g. insurer).

5.6 *Contacting authors of eligible studies*

We identified 88 studies in category i, 4 in category iii, and none in either category ii or iv. Contact information was not reported and not available through an Internet search for authors of 22 (24%) trials. We attempted to contact authors of the

remaining 70 trials by email and requested information on whether they had an eligibility stipulation for disability status. If authors included patients on disability benefits, we requested their trial data to facilitate an individual patient data meta-analysis (IPDMA). To maintain patient confidentiality, authors removed any personal identifiers from their dataset prior to transferring it to our center. We clarified uncertainties or discrepancies in the data sets with the study authors and combined individual patient data for variables that were similar across the trials. Based on authors' replies, we classified studies into four groups: (A) those that did have some data specific to patients on disability benefits, (B) those that confirmed that they had no patients on disability benefits, (C) those that did not have an eligibility criterion for disability status and did not collect information on disability status, and (D) unknown or did not respond.

5.7 Data collection process

Using piloted standardized forms and a detailed instruction manual to extract data, the same teams of reviewers extracted data, independently and in duplicate, from studies in groups A and B. We did not abstract data from groups C and D.

Data abstracted included patient characteristics, treatment effect on depression, frequency and timing of follow-up, details of depression (including diagnostic

classification system used, severity of depression, and duration of depression), and CBT intervention details (including the type of CBT administered, expertise of providers administering CBT, and frequency of CBT). Reviewers abstracted data from the following study arms: CBT, TAU and minimal or no treatment. Data comparing CBT only to active comparators were not abstracted, unless the active comparator was equally balanced between both the treatment and control group.

5.8 Risk of Bias within studies

Using a modified Cochrane risk of bias instrument, reviewers assessed risk of bias for each eligible trial on the following domains: sequence generation; allocation concealment; blinding of participants, investigators, data collectors, outcome assessors, and data analysts; incomplete outcome data; selective outcome reporting; and other sources of bias (e.g. bias of study design, trial stopped early, extreme baseline imbalance, and fraudulent trial) [34,35]. Reviewers used response options of “definitely yes”, “probably yes”, “probably no”, and “definitely no” with definitely and probably yes ultimately assigned high risk of bias and probably and definitely no assigned low risk of bias [35]. The reviewers resolved disagreements by discussion, and an arbitrator (JWB) adjudicated any remaining conflicts.

5.9 *Synthesis of results*

For our IPDMA, we compared the effects (mean difference) of CBT on depression, measured by the most commonly reported instrument [Beck Depression Inventory (BDI-II)], in patients receiving disability benefits versus patients not receiving disability benefits. We used a one-stage method [36], and included the following variables in our model: study arm, receipt of disability benefits, interaction term of study arm and receipt of disability benefits, trial as a categorical variable, age and baseline BDI-II score. To guard against multiplicity of data [37], we used the most common follow-up time point of 3 months for our analysis.

Our secondary analyses evaluated whether there were differences in patients not in receipt of disability benefits between trials that included patients in receipt of disability benefits (group A) and trials with aggregate data that did not include patients receiving disability benefits (group B). We compared the following: 1) the effects of CBT between group A and B; 2) the effects of CBT between group A and B that compared CBT plus pharmacotherapy versus pharmacotherapy alone; 3) the effects of CBT between group A and B that compared CBT to TAU.

For our secondary analyses, we used the 2-stage method [38]. In the first stage, we aggregated the IPD data of the patients not receiving disability benefits in group A and

in the second stage, pooled the aggregate data of studies in group A and B using a random-effects model.

We used the means and standard deviations (SDs) of the end of study scores for our secondary pooled analyses. To pool data across trials and to facilitate interpretation for clinicians and other stakeholders, we calculated the mean difference (MD) and its associated 95% confidence interval (CI) of the natural units of the most familiar instrument across trials, the BDI-II. For this calculation, we used the following formulas to convert mean estimates (M) and standard deviations (SD) into the scale of the most familiar instrument: $M_A = (M_B - L_B) (R_A/R_B) + L_A$ and $SD_A = SD_B (R_A/R_B) + L_A$, where A represented the most familiar instrument and B represented the alternative instrument, L_A and L_B represent the lower limits of instrument A and B respectively, and R_A and R_B represented the ranges for instruments A and B respectively [39].

We examined heterogeneity using both a chi-squared test and the I^2 statistic [40]. Heterogeneity defined by an I^2 of 0% to 40% was interpreted as ‘might not be important’, 30% to 60% as ‘moderate heterogeneity’, 50% to 90% as ‘substantial heterogeneity’, and 75% to 100% as ‘considerable heterogeneity’ [40]. We generated the following a priori hypotheses to explain variability between studies in our secondary analyses: studies using in-person CBT will have greater effects than studies

using computer administered-CBT, and studies with high risk of bias will demonstrate larger effects compared to studies with low risk of bias.

We performed analyses using SPSS version 20 and the Cochrane Collaboration Review Manager software (RevMan version 5.1.2).

CHAPTER 6

RESULTS

6.1 *Study selection*

We screened 977 citations and retrieved 421 articles in full text; 329 studies did not meet inclusion criteria and 92 trials were deemed eligible. The kappa (95% CI) chance-corrected agreement on assessing full text eligibility was 0.74 (0.66 to 0.81), representing substantial agreement.

After establishing author contact for 56 of the 70 trials for which we acquired contact information, we found that 45 trials did not have an eligibility criterion based on disability benefit status or collect information on disability status, 6 trials did not enrol any patients in receipt of disability benefits, and 5 trials enrolled some patients in receipt of disability benefits. Authors of 4 of the 5 trials that included patients in receipt of disability benefits agreed to provide individual patient data. Two of these trials combined patients who were disabled with unemployed and retired individuals and information specific to receipt of disability benefits were uncertain; these trials were therefore excluded from our IPDMA. Our primary analysis consisted of the 2 remaining trials that included some patients in receipt of disability benefits [41,42], and our secondary analyses consisted of 8 trials, i.e., 6 trials that did not enrol any

patients in receipt of disability benefits [43,44,45,46,47,48], and 2 trials that included some patients in receipt of disability benefits (Figure 6.1) [41,42].

6.2 *Study characteristics*

Seven studies were parallel group RCTs [41,42,43,44,45,46,47], and one was a cluster RCT [48]. Table 6.1 describes the characteristics of the 8 eligible trials, and Table 6.2 provides details regarding their interventions.

6.3 *Risk of bias within studies*

Protection against bias was generally poor (Figure 6.2). All 8 trials reported loss to follow-up (LTFU), ranging from 4% to 40%. Four trials excluded those LTFU and performed a complete case analysis [41,42,43,48], 2 used the last observation carried forward [44,46], 1 used multiple imputation (56), and 1 did not report an approach [47].

6.4 *IPDMA*

Two trials including data on patients receiving disability benefits enrolled a total of 227 patients; 34 in receipt of disability benefits and 193 not receiving disability

Figure 6.1: Flow chart of study eligibility

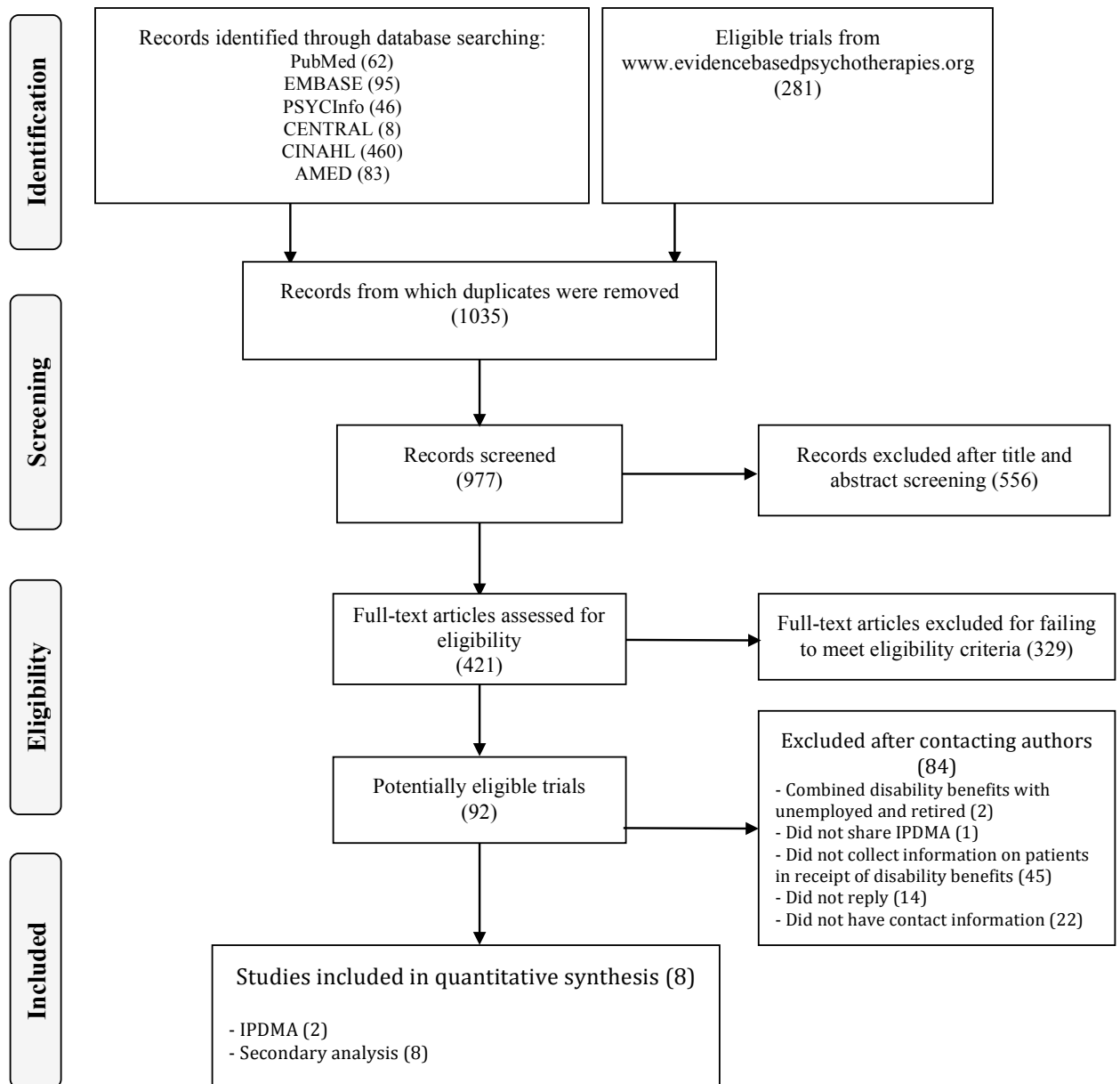


Table 6.1: Characteristics of CBT studies

Study	Country	Sample size		Age (mean ± SD)		Patient Population	Treatment group	Control group	Depression outcomes reported	Included patients on disability benefits
		CBT	Control	CBT	Control					
De Graaf 2009[41]	The Netherlands	100	103	45.2 ± 10.9	45.1 ± 12.2	Depression	CBT + TAU	TAU	BDI-II	Yes
Dozois 2009[42]	Canada	25	23	NR	NR	Depression	CBT + pharmacotherapy	Pharmacotherapy alone	BDI-II; HRSD	Yes
Naeem 2011[47]	Pakistan	17	17	32.35 ± 8.9	33.64 ± 1.0	Depression	CBT + pharmacotherapy	Pharmacotherapy alone	HADS; BSI	No
Faramarzi 2007[43]	Iran	42	40	28.3 ± 3.8	28.4 ± 5.3	Depression in infertile women	CBT	Minimal or no treatment	BDI-II	No
Hollon 1992[44]	USA	25	57	NR	NR	Nonpsychotic, nonbipolar depressed outpatients	CBT + pharmacotherapy	Pharmacotherapy alone	BDI; HRSD	No

Miranda 2003[45]	USA	90	89	29.8 ± 7.9	29.5 ± 9.1	Depression in predominantly low-income young minority women	CBT	TAU	HRSD	No
Misri 2004[46]	Canada	19	16	29.5 ± 3.3	30.8 ± 5.9	Postpartum depression	CBT + pharmaco- therapy	Pharmaco- therapy alone	HRSD	No
Rahman 2008[48]	Pakistan	463	440	26.5 ± 5.2	27.0 ± 5.1	Perinatal depression	CBT	TAU	HRSD	No

CBT –Cognitive Behavioural Therapy; TAU – Treatment As Usual; SD – Standard deviation; NR – Not reported; BDI-II – Beck Depression Inventory-II; HRSD – Hamilton Rating Scale for Depression; HADS – Hospital Anxiety and Depression Scale; BSI – Bradford Somatic Inventory

Table 6.2: CBT details from studies

Study	Mode of administration of CBT	Duration of CBT per visit	Frequency of CBT	Total duration of CBT	Clinical background of the individuals administering CBT	Was there a standardized program or certification process that CBT providers have undergone or had to undergo?
De Graaf 2009[41]	Computer/ internet based CBT	30 minutes	1 per week	9 weeks	Not reported	Not reported
Dozois 2009[42]	In-person individualized CBT	1 hour	1 per week	15 weeks	Master's level therapist	Not reported
Naeem 2011[47]	In-person individualized CBT	Not reported	1 to 2 sessions per week	9 weeks	Psychiatrist; psychology graduates	Not reported
Faramarzi 2007[43]	In-person group CBT	2 hours	1 per week	10 weeks	Psychologist	Not reported
Hollon 1992[44]	In-person individualized CBT	50 minutes	2 in the first 4 weeks, 1 or 2 in the next 4 weeks, and 1 in the last weeks	12 weeks	Psychologist; social worker	Not reported
Miranda 2003[45]	In-person individualized CBT	Not reported	1 per week	8 weeks	Psychologist; psychotherapist	Not reported
Misri 2004[46]	In-person individualized CBT	1 hour	1 per week	12 weeks	Psychologist	Not reported

Rahman 2008[48]	In-person individualized CBT	Not reported	4 in 1st month, 3 in 2nd month, and 1 per month for next 9 months	11 weeks	Lady health workers	Not reported
-----------------	------------------------------	--------------	---	----------	---------------------	--------------

CBT –Cognitive Behavioural Therapy

Figure 6.2: Risk of Bias within studies

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of patients	Blinding of healthcare providers	Blinding of data collectors	Blinding of outcome assessment (detection bias)	Blinding of data analysts	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
DeGraaf 2009	+	+	-	-	-	-	-	-	+	+
Dozois 2009	-	-	-	-	-	-	-	-	+	+
Faramarzi 2007	-	-	-	-	-	-	-	-	+	+
Hollon 1992	+	+	-	-	+	+	-	-	+	+
Miranda 2003	+	-	-	-	-	-	-	-	+	+
Misri 2004	+	-	-	+	+	+	-	-	+	+
Naeem 2011	+	+	-	-	+	+	-	-	+	+
Rahman 2008	+	-	+	-	+	+	-	-	+	+

'+' denotes low risk of bias, and '-' represents high risk of bias.

benefits. The mean (SD) baseline BDI-II score for patients with disability benefits was 32.9 (± 8.55) and for patients not receiving disability benefits 26.9 (± 7.9).

Pooled results from these 2 trials, including both those receiving and not receiving disability benefits, suggested a possible benefit of CBT on depression (MD = -2.61; 95% CI = -5.28 to 0.07; $p=0.06$, minimally important difference [MID] = 5), as did results from both the subgroup of patients in receipt of disability benefits (MD = -6.88; 95% CI = -14.06 to 0.31), and patients not receiving disability benefits (MD = -2.22; 95% CI = -5.07 to 0.63). Results suggested a possible larger effect on reducing depression in those receiving versus not receiving disability benefits, though the confidence interval includes a small reduction in benefit in those receiving benefits (MD = -4.46; 95% CI = -12.21 to 3.30; $p=0.26$; MID = 5).

6.5 *Secondary analyses*

There were no significant differences in the effect of CBT on depression among patients not in receipt of disability benefits across studies that enrolled patients receiving disability benefits and studies that did not ($p=0.26$) (Figure 6.3). There were no significant differences in the effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT plus pharmacotherapy versus pharmacotherapy alone ($p=0.94$) (Figure 6.4). There were no significant differences in

Figure 6.3: Effect of CBT in patients not receiving disability benefits in studies including patients receiving benefits versus those that did not

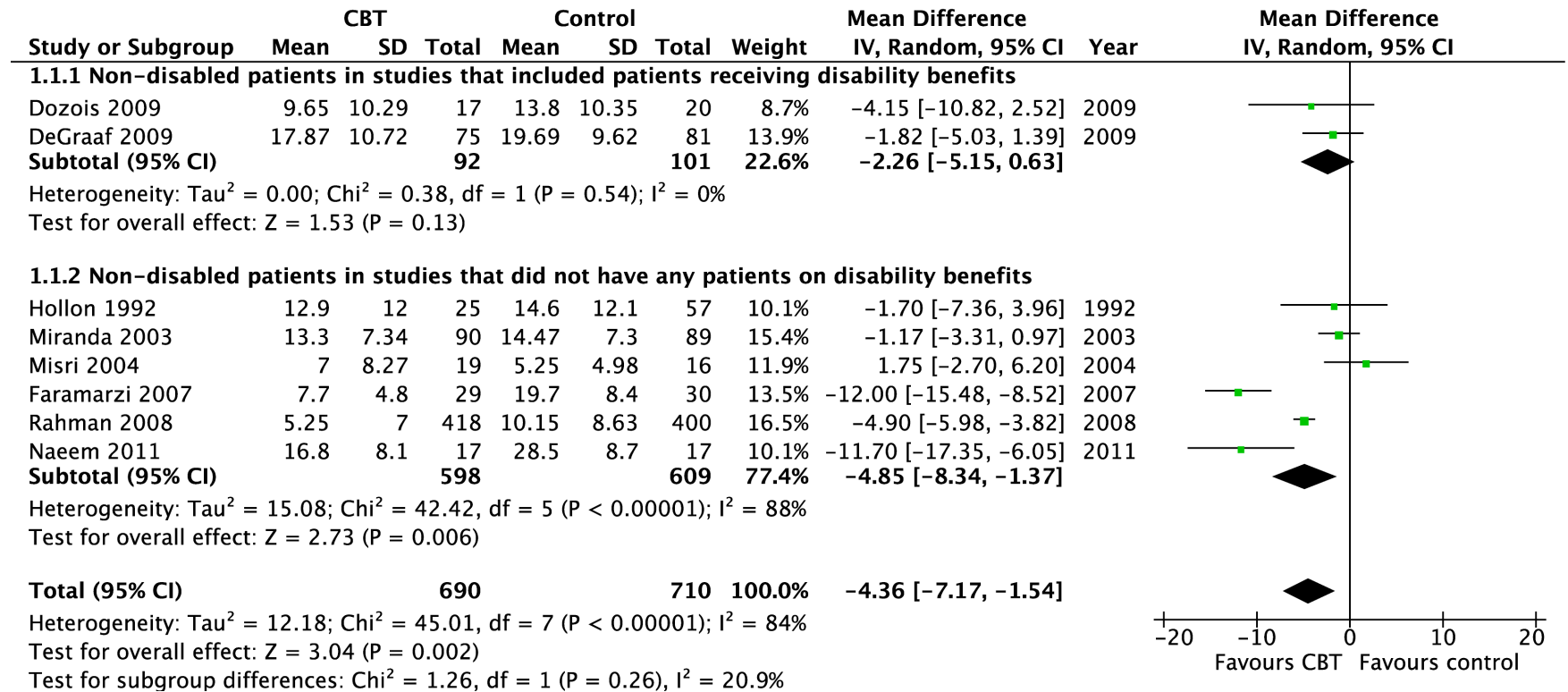
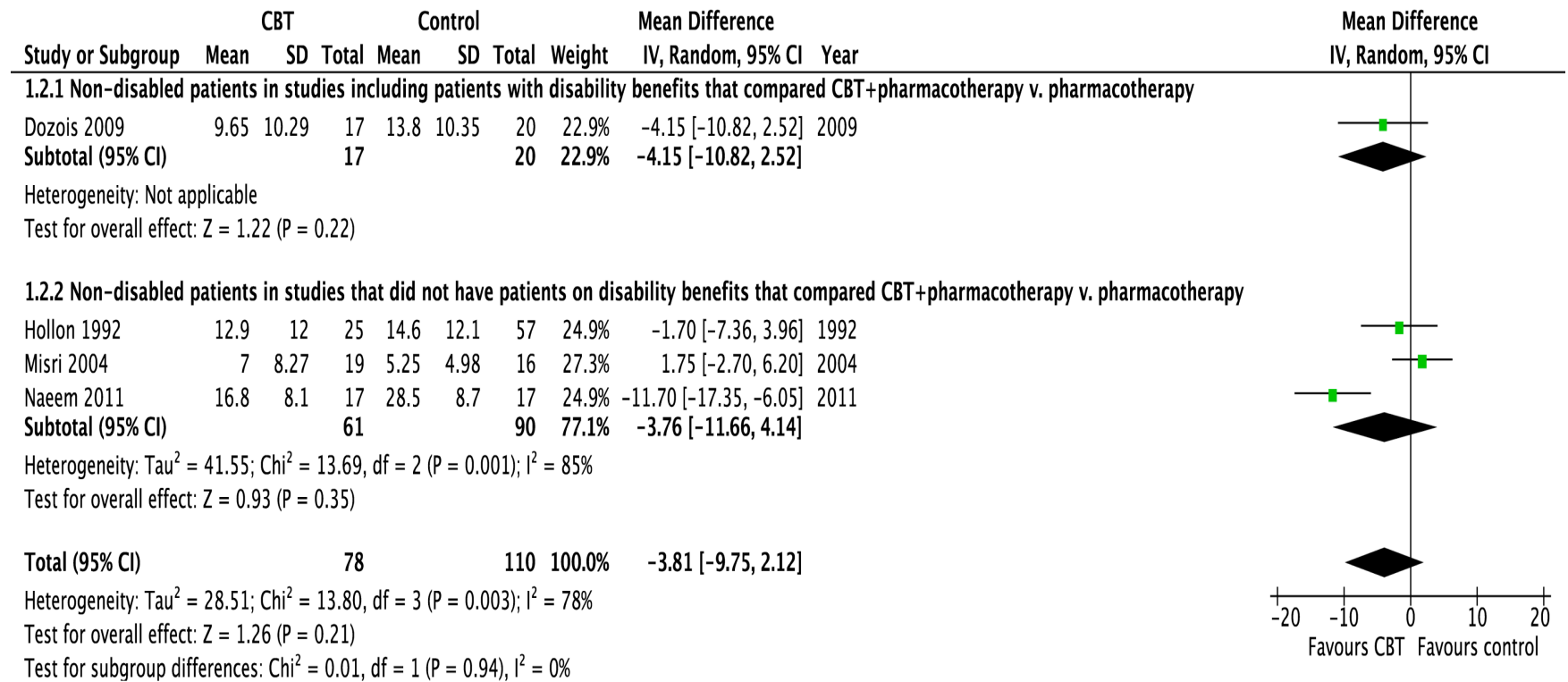
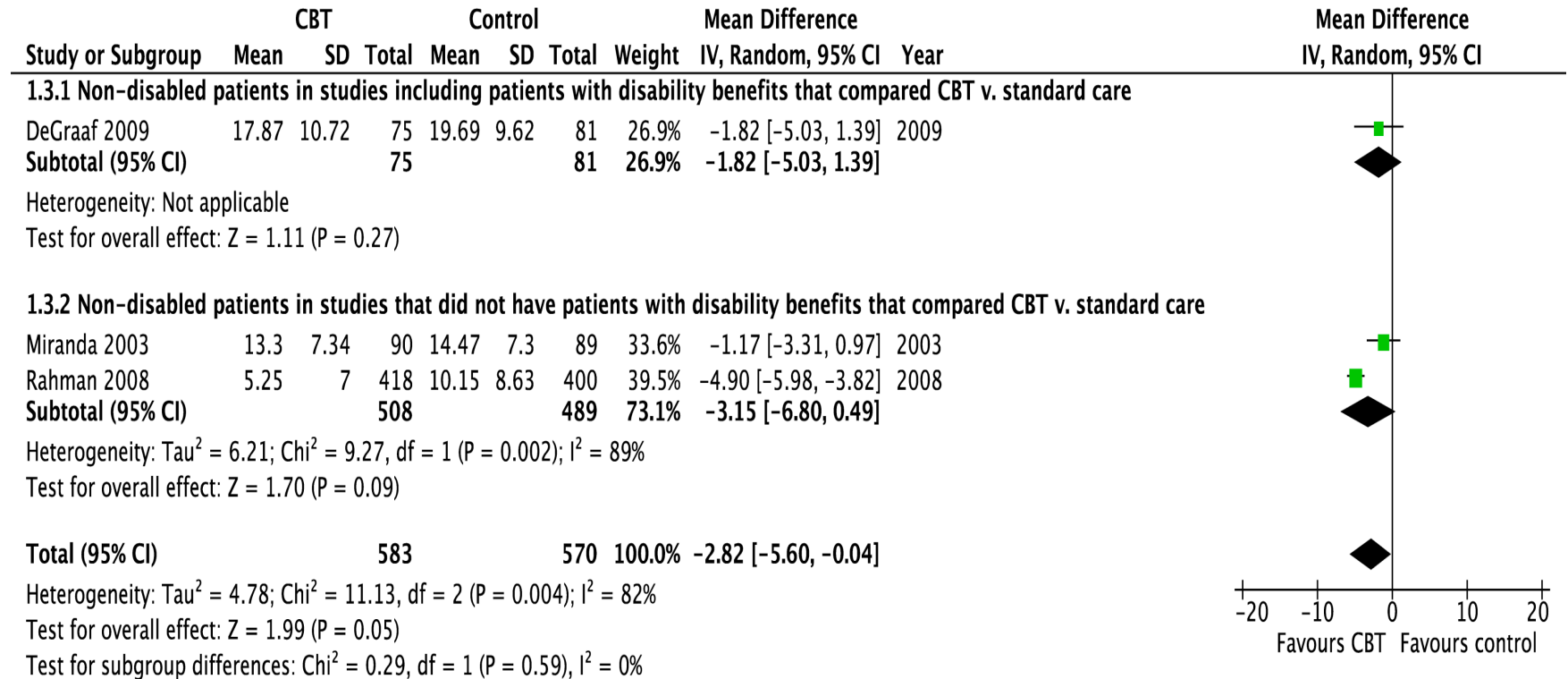


Figure 6.4: Effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT plus pharmacotherapy versus pharmacotherapy alone



the effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT versus TAU/standard care ($p=0.59$) (Figure 6.5). Our a priori subgroup hypotheses failed to explain the heterogeneity observed in our secondary analyses.

Figure 6.5: Effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT versus TAU/standard care



CHAPTER 7

DISCUSSION

7.1 Summary of evidence

This is the first systematic review comparing the effect of receiving disability benefits on depression following treatment with CBT. We failed to find differences in the effect of CBT on depression between patients receiving disability benefits and patients not receiving disability benefits. The results suggest a possible greater effect in those receiving disability benefits (-4.46 BDI units in which the minimally important difference is 5), and the boundaries of the confidence interval suggest that if there is a decrement in benefit, that decrement is small (no greater than 3.30 BDI-II units). Nevertheless, these data come from only 34 patients receiving disability benefits, so that any inferences regarding relative effect in the two populations are very weak.

7.2 Strengths

The strengths of our review include a comprehensive and transparent search strategy, independent and duplicate eligibility assessment, use of the most commonly reported instrument with the most established reliability and validity (BDI-II) for our pooled analysis, and use of individual patient data from eligible trials, allowing adjustment for potential confounding predictors. We also ensured rigorous data abstraction by using

detailed written instructions, conducting formal calibration exercises, conducting in duplicate, and implementing a consensus approach to resolve disagreement. We contacted authors to verify whether they enrolled patients in receipt of disability benefits and achieved an 80% response rate among trials for which we were able to acquire author contact information.

7.3 Previous Evidence

Although no prior reviews have explored the effect of CBT in patients receiving disability benefits, reviews have explored the effect of compensation in other patient populations. A 2005 systematic review found that the presence of compensation was associated with worse outcome (combination of functional, quality of life, pain and patient satisfaction outcome that was rated as satisfactory or unsatisfactory by review investigators) after surgery [29]. This was consistent with findings from systematic reviews regarding chronic pain and closed-head injuries [49,50], which reported a significant effect between compensation and poor outcome. This indirect evidence, however, does not address the relative effect of interventions in the populations (one may have poorer outcomes, but still have larger treatment effects if results without treatment are very poor). In the two trials we examined, patients in receipt of disability benefits had a greater severity of depression than those who were not receiving disability benefits (baseline BDI-II of 32.9 versus 26.9). Although a prior review reported that the effectiveness of CBT was reduced in patients with severe

depression compared to those with mild to moderate depression [51], we found no suggestion of a smaller effect of CBT in patients receiving disability benefits.

7.4 *Limitations*

Our study has limitations. First, our IPDMA is based on only 34 patients in receipt of disability benefits and 193 patients not receiving disability benefits. The extent to which findings from this small sample will generalize to a wide population of individuals in receipt of benefits is uncertain. Second, our secondary analyses showed substantial heterogeneity within subgroups of patients not receiving disability benefits, which could not be explained by our *a priori* hypotheses. Possible explanatory factors that we were unable to explore due to limitations in the reporting of trials include baseline severity of depression, duration of depression, frequency of CBT, and experience of CBT providers. Third, none of the trials evaluated the effect of CBT on return to work (RTW), a critical outcome for patients receiving disability benefits and for insurers providing benefits. It remains possible that CBT may improve BDI-II scores, but may not have any effect on claim resolution or RTW. Future trials should include these outcomes in order to ascertain a BDI-II threshold that is associated with RTW and claim resolution.

7.5 *Conclusions*

If the use of CBT to manage depression among patients receiving disability benefits was less effective than in patients not receiving disability benefits, clinicians and payers might reasonably choose alternative treatment strategies (e.g. pharmacotherapy, other psychotherapies or a combination of both). The limited evidence available, however, provides no support for this hypothesis and suggests that, for the time being, CBT should continue as a recommended approach for addressing depression in patients receiving disability benefits. Secure inference will, however, only be possible after the conduct of much larger comparative trials, conducted with low risk of bias and in collaboration with insurers.

ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS

Acknowledgements

We thank Dr. Randi McCabe for her assistance in reviewing the eligibility of the treatment in trials.

Funding

There were no funds received to complete this study.

Conflicts of Interest

Shanil Ebrahim was supported by a MITACS scholarship that was partnered with Sun Life Financial. Sun Life did not have a role in the study design, data collection and analysis, or preparation of the manuscript. Jason W. Busse acts as a consultant to Prisma Health Canada, a private incorporated company funded by employers and insurers that consults on and manages long-term disability claims. Marg Bellman is the assistant vice president of the National Disability Services, Policy & Procedure department at Sun Life Financial Assurance Company of Canada.

Authors' contributions

Shanil Ebrahim was responsible for preparation of the first draft of the manuscript.

Shanil Ebrahim, Luis Montoya, Sandy Hsu, Mostafa Kamal el Din, Wanda Truong and

Alonso Carrasco-Labra completed screening, reviewing and data abstraction of the

articles. L. Esther de Graaf, David Dozois and Peter Bieling contributed their

randomized controlled trial patient data. Shanil Ebrahim completed the statistical

analyses with assistance from Stephen Walter, Diane Heels-Ansdell, Jason Busse and

Gordon Guyatt. Shanil Ebrahim, Jason Busse, Stephen Walter and Gordon Guyatt were

responsible for the conception and idea of the manuscript. Irene Patelis-Siotis and

Marg Bellman were engaged as knowledge users in the study to ensure that our

findings are integrated in practice. Shanil Ebrahim is the first author and Gordon

Guyatt is the senior author. All authors critically reviewed several drafts of the

manuscript. All authors read and approved the final manuscript to be published.

REFERENCES

1. Hirschfeld R, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, et al. (2000) Social functioning in depression: a review. *J Clin Psychiatry* 61: 268-275.
2. Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D (2003) Cost of lost productive work time among US workers with depression. *JAMA* 289: 3135-3144.
3. World Health Organization (WHO) (2002) *Mental Health: New understanding, New hope*. Geneva: WHO.
4. Fournier J, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, et al. (2010) Antidepressant drug effects and depression severity: A patient-level meta-analysis. *JAMA* 303: 47-53.
5. NICE (2009) *Depression: Treatment management of depression in adults*. London: National Institute for Clinical Excellence.
6. Olfson M, Marcus SC (2009) National patterns in antidepressant medication treatment. *Arch Gen Psychiatry* 8: 848-856.
7. Dobson KS (2001) *Handbook of cognitive-behavioral therapies*. 2nd Edition. New York; London: Guilford Press.

8. Almeida A, Lotufo Neto F (2003) Cognitive-Behavioral therapy in prevention of depression relapses and recurrences: a review. *Rev Bras Psiquiatr* 25: 239-244.
9. Bortolotti B, Menchetti M, Bellini F, Montaguti MB, Berardi D. (2008) Psychological interventions for major depression in primary care: a meta-analytic review of randomized controlled trials. *General Hospital Psychiatry* 30: 293-302.
10. Churchill R, Hunot V, Corney R, Knapp M, McGuire H, et al. (2001) A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment* 5: 1-173.
11. Cuijpers P, van Straten A, Warmerdam L, Andersson G (2008) Psychological treatment of depression: A meta-analytic database of randomized studies. *BMC Psychiatry* 8.
12. Cuijpers P, van Straten A, van Schaik A, Andersson G (2009) Psychological treatment of depression in primary care: A meta-analysis. *Journal of General Practice* 59: e51-e60.
13. Cuijpers P, Smit F, Bohlmeijer E, Hollon SD, Andersson G (2010) Efficacy of cognitive behavioural therapy and other psychological treatments for adult depression: a meta-analytic study of publication bias. *Br J Psychiatry* 196: 173-178.

14. Gloaguen V, Cottraux J, Cucherat M, Blackburn IM (1998) A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* 49: 59-72.
15. Harrington R, Whittaker J, Shoebridge P, Campbell F (1998) Systematic review of efficacy of cognitive behaviour therapies in childhood and adolescent depressive disorder. *BMJ* 316: 1559-1563.
16. Kaltenthaler E, Parry G, Beverley C, Ferriter M (2008) Computerised cognitive-behavioural therapy for depression: systematic review. *Br J Psychiatry* 193: 181-184.
17. Oei TP, Dingle G. (2008) The effectiveness of group cognitive behaviour therapy for unipolar depressive disorders. *J Affect Disorder* 107: 5-21.
18. Peng XD, Huang CQ, Chen LJ, Lu ZC (2009) Cognitive behavioural therapy and reminiscence techniques for the treatment of depression in the elderly: a systematic review. *J Int Med Res* 37: 975-982.
19. Wampold BE, Minami T, Baskin TW, Callen TS (2002) A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. *Journal of Affective Disorders* 68: 159-165.
20. Kouzis AC, Eaton WW (1997) Psychopathology and the development of disability. *Social and Psychiatry and Psychiatric Epidemiology* 32: 379-386.

21. Kouzis AC (2000) Psychopathology and the initiation of disability payments. *Psychiatr Serv* 51: 908-913.
22. Salkever DS, Goldman H, Purushothaman M, Shinogle J (2000) Disability management, employee health and fringe benefits, and long-term-disability claims for mental disorders: an empirical exploration. *The Milbank Quarterly* 78: 79-113.
23. Dewa C, Hoch JS, Goering P, Lin E, Patterson M (2003) Use of antidepressants among Canadian workers receiving depression-related short-term disability benefits. *Psychiatr Serv* 54: 724-729.
24. Ebrahim S, Guyatt GH, Walter SD, Heels-Ansdell D, Hanna S, et al. (2012) Association of psychotherapy with disability benefit claim closure among patients disabled due to depression (under review). Hamilton: McMaster University.
25. (2003) Roadblocks in Cognitive-Behavioral Therapy: Transforming Challenges into Opportunities for Change; Leahy R, editor. New York: The Guilford Press.
26. Weighill VE (1983) "Compensation neurosis": a review of the literature. *J Psychosom Res* 27: 97-104.
27. Young JN, Shaffrey CI, Laws ER Jr, Lovell LR (1997) Lumbar disc surgery in a fixed compensation population: a model for influence of secondary gain on surgical outcome. *Surg Neurol* 48: 552-558.

28. Olney JR, Quenzer DE, Makowsky M (1999) Contested claims in carpal tunnel surgery: outcome study of worker's compensation factors. *Iowa Orthop J* 19: 111-121.
29. Harris I, Mulford J, Solomon M, van Gelder JM, Young J. (2005) Association between compensation status and outcome after surgery: a meta-analysis. *JAMA* 293: 1644-1652.
30. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 62: 1006-1012.
31. American Psychiatry Association (2010) *Diagnostic and Statistical Manual*. Arlington, VA.
32. Cuijpers P (2011) *Psychotherapy randomized controlled and comparative trials*. Amsterdam: vrije Universiteit.
33. Landis J, Koch GG, Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.
34. The Cochrane Collaboration (2011) Chapter 8: Risk of Bias. In: Higgins JPT, Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0)*. Oxford, UK: Cochrane Collaboration.

35. Akl EA, Sun X, Busse JW, Johnston BC, Briel M, et al. (2011) Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. *J Clin Epidemiol* Dec 24 (Eub ahead of print).
36. (1999) Variable intercepts: fixed or random parameters? In: Snijders T, Bosker RJ, editor. *Multilevel Analysis An introduction to basic and advanced multilevel modeling*. London: SAGE Publications Ltd.
37. Tendal B, Nüesch E, Higgins JP, Jüni P, Gøtzsche PC (2011) Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ* 343.
38. Riley R, Simmonds MC, Look MP (2007) Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J Clin Epi* 60: 431-439.
39. Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH (2012) Pooling health-related quality of life outcomes in meta-analysis—a tutorial and review of methods for enhancing interpretability. *Research Synthesis Methods* In press.
40. The Cochrane Collaboration (2009) 9.5.2 Identifying and measuring heterogeneity. In: Higgins JPT, Thompson SG, editor. *Cochrane Handbook for Systematic Reviews of Interventions (Version 502)*. Oxford, UK: Cochrane Collaboration.

41. de Graaf LE, Gerhards SAH, Arntz A, Riper H, Metsemakers JFM, et al. (2009) Clinical effectiveness of online computerised cognitive behavioural therapy without support for depression in primary care: randomised trial. *The British Journal of Psychiatry* 195: 73-80.
42. Dozois DJA, Bieling PJ, Patelis-Siotis I, Hoar L, Chudzik S, et al. (2009) Changes in Self-Schema Structure in Cognitive Therapy for Major Depressive Disorder: A Randomized Clinical Trial. *Journal of Consulting and Clinical Psychology* 77: 1078-1088.
43. Faramarzi M, Alipor A, Esmaelzadeh S, Kheirkhah F, Poladi K, et al. (2008) Treatment of depression and anxiety in infertile women: Cognitive behavioral therapy versus fluoxetine. *Journal of Affective Disorders* 108: 159-164.
44. Hollon SD, DeRubeis RJ, Evans MD, Wiemer MJ, Garvey MJ, et al. (1992) Cognitive therapy and pharmacotherapy for depression. *Arch Gen Psychiatry* 49: 774-781.
45. Miranda J, Chung JY, Green BL, Krupnick J, Siddique J, et al. (2003) Treating depression in low-income young minority women: A randomized controlled trial. *JAMA* 290: 57-65.
46. Misri S, Reebye P, Corral M, Milis L (2004) The Use of Paroxetine and Cognitive-Behavioral Therapy in Postpartum Depression and Anxiety: A Randomized Controlled Trial. *Journal of Clinical Psychiatry* 65: 1236-1241.

47. Naeem F, Waheed W, Gobbi M, Ayub M, Kingdon D (2011) Preliminary Evaluation of Culturally Sensitive CBT for Depression in Pakistan: Findings from Developing Culturally-Sensitive CBT Project (DCCP). *Behavioural and Cognitive Psychotherapy* 39: 165-173.
48. Rahman A, Malik A, Sikandar S, Roberts C, Creed F (2008) Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *The Lancet* 372: 902-909.
49. Binder LM, Rohling ML (1996) Money matters: a metaanalytic review of the effects of financial incentives on recovery after closed-head injury. *Am J Psychiatry* 153: 7-10.
50. Rohling ML, Binder LM, Langhinrichsen-Rohling J (1995) Money matters: a meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. *Health Psychol* 14: 537-547.
51. Haby MM, Donnelly M, Corry J, Vos T (2006) Cognitive behavioural therapy for depression, panic disorder and generalized anxiety disorder: a meta-regression of factors that may predict outcome. *Aust N Z J Psychiatry* 40: 9-19.

STUDY 2

Association of psychotherapy with disability benefit claim closure among patients disabled due to depression

Shanil Ebrahim PhD(c)^{1§}, Gordon H. Guyatt MD^{1,2}, Stephen D. Walter PhD^{1,3}, Diane Heels-Ansdell MSc¹, Marg Bellman⁴, Steven E. Hanna¹, Irene Patelis-Siotis⁵, Jason W. Busse^{1,6}

¹Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada

²Department of Medicine, McMaster University, Hamilton, Canada

³Department of Mathematics and Statistics, McMaster University, Hamilton, Canada

⁴National Disability Services, Policy & Procedure department, Sun Life Financial, Toronto, Canada

⁵Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, Canada

⁶Institute for Work and Health, Toronto, Ontario, Canada

[§]Corresponding author

Correspondence to:

Shanil Ebrahim, Clinical Epidemiology and Biostatistics, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4K1, shanil.ebrahim@utoronto.ca

Keywords

Psychotherapy, Disability benefits, Depression, Claim closure, Administrative database analysis, Observational cohort study

Accepted for publication by PLoS ONE

CHAPTER 8

ABSTRACT

Background: Depression is the most frequent reason for receiving disability benefits in North America, and treatment with psychotherapy is often funded by private insurers. No studies have explored the association between the provision of psychotherapy for depression and time to claim closure.

Methods: Using administrative data from a Canadian disability insurer, we evaluated the association between the provision of psychotherapy and short-term disability (STD) and long-term disability (LTD) claim closure by performing Cox proportional hazards regression.

Results: We analyzed 10,508 STD and 10,338 LTD claims for depression. In our adjusted analyses, receipt of psychotherapy was associated with longer time to STD closure (HR [99% CI] = 0.81 [0.68 to 0.97]) and faster LTD claim closure (1.42 [1.33 to 1.52]). In both STD and LTD, older age (0.90 [0.88 to 0.92] and 0.83 [0.80 to 0.85]), per decade), a primary diagnosis of recurrent depression versus non-recurrent major depression (0.78 [0.69 to 0.87] and 0.80 [0.72 to 0.89]), a psychological secondary diagnosis (0.90 [0.84 to 0.97] and 0.66 [0.61 to 0.71]), or a non-psychological secondary diagnosis (0.81 [0.73 to 0.90] and 0.77 [0.71 to 0.83]) versus no secondary diagnosis, and an administrative services only policy ([0.94 [0.88 to 1.00] and 0.87

[0.75 to 0.996]) or refund policy (0.86 [0.80 to 0.92] and 0.73 [0.68 to 0.78]) compared to non-refund policy claims were independently associated with longer time to STD claim closure.

Conclusions: We found, paradoxically, that receipt of psychotherapy was independently associated with longer time to STD claim closure and faster LTD claim closure in patients with depression. We also found multiple factors that were predictive of time to both STD and LTD claim closure. Our study has limitations, and well-designed prospective studies are needed to establish the effect of psychotherapy on disabling depression.

CHAPTER 9

INTRODUCTION

9.1 *Background*

Major Depressive Disorder (henceforth referred to as depression) results in immense human suffering and is associated with considerable socioeconomic costs. Depression accounts for 11% of disability worldwide and an estimated productivity loss of \$14.4 billion annually in Canada [1,2,3,4]. The World Health Organization (WHO) estimates suggest that depression will become the second leading cause of disease burden worldwide by the year 2020 [5,6,7].

The National Institute for Health and Clinical Excellence (NICE) in the UK has recommended that health care professionals provide, alone or in combination, pharmacological treatments and high-intensity psychological interventions for individuals suffering from moderate or severe depression. The most frequently prescribed psychotherapy for treating depression is cognitive behavioural therapy (CBT) [8,9].

Depression is a common reason for receiving disability benefits [10,11,12], incurring more costs for long-term disability (LTD) than other disorders [13]. Individuals suffering from psychiatric disorders who are also receiving disability benefits require more complex treatment and have more difficulty returning to work than those suffering from other disabling complaints [14]. Psychological therapy may be less effective, or ineffective, in patients receiving disability benefits, as their circumstances or psychological status may interfere with the successful implementation of therapy [15]. There is indirect evidence for this hypothesis from surgical populations: a recent meta-analysis of 129 studies revealed that the odds of an unsatisfactory outcome in patients receiving disability benefits or engaged in litigation was 3.79 times greater (95% confidence interval [CI]: 3.28 to 4.37) versus similar patients not in receipt of disability benefits or pursuing litigation [16].

Given that psychotherapy is one of the most frequently reimbursed treatment for depression by private insurers [17], it is important to ascertain if psychotherapy represents a worthwhile expenditure of time and energy for depressed patients, and a good investment for insurers. We recently completed a systematic review in which none of 92 randomized controlled trials (RCTs) that explored the effect of CBT on depression reported whether enrolled patients were receiving disability benefits. We successfully contacted 56 trialists and identified 3 trials that captured information on disability benefit status [18]. Our analyses consisting of 2 trials (including 34 patients

on disability benefits) did not find a significant difference in depression between patients receiving disability benefits versus those not receiving disability benefits. However, we were limited by the small number of patients available for analysis.

9.2 *Objectives*

In the present study, we used the administrative data of a large Canadian, private, disability insurer (Sun Life Financial Inc.) to explore the association between the provision of psychotherapy for patients suffering from depression and time to both short-term disability (STD) and LTD claim closure. Additionally, we evaluated what factors were associated with receipt of psychotherapy in patients with depression in receipt of disability benefits.

CHAPTER 10

METHODS

10.1 Ethics statement

The Research Ethics Board at McMaster University approved our study.

10.2 Design

Secondary analysis of an insurance administrative database

10.3 Description of patients and eligibility criteria

Between January 2007 and December 2010, Sun Life Financial had 259,510 claims submitted for approval. Of these, 190,527 were STD claims and 68,983 LTD claims. An STD and LTD claim differ in regards to the potential duration of time that claimants may receive wage replacement benefits. The two most common standard benefit periods for an STD claim is up to 17 weeks and 26 weeks, but some plan benefit periods may be less or more than this duration. An LTD claim pay wage replacement benefits for a longer period, up to age 65.

Of the 259,510 filed claims at Sun Life, 172,425 (90.5%) STD and 55,530 (80.5%) LTD claims were approved. For our analyses, we included all claims (1 claim per claimant) that were approved for STD or LTD benefits with a primary diagnosis of major depressive disorder or recurrent depressive disorder. We excluded all individuals

whose claims were recorded as closed prior to contractual approval, and STD claims with a maximum claim benefit period over 2.5 years as we deemed those to be data entry errors.

10.4 Administrative Variables

The database consisted of demographic, administrative, and clinical information. The case manager(s) responsible for overseeing each claim entered all data. The standard requirement is for data to be entered within 5 days of claim receipt for STD claims, and within 10 days for LTD claims, although this may vary.

Appendix C presents the list and description of all variables in the database. Guided by the results from observational studies evaluating predictors of recovery in patients receiving disability benefits due to depression [19,20,21], and a systematic review evaluating prognostic factors of long term disability due to mental disorders [22], we selected, a priori, 14 variables from the database that we judged may be associated with claim closure and receipt of psychotherapy, and predicted the direction of anticipated effects. In addition to the variables we chose from previous evidence, we included and predicted the direction of anticipated effects of two additional variables (time to claim approval and disability funding policy) in our model, as per recommendations by content experts in our research team and the administrative team at Sun Life. Two psychologists, blinded to study results, provided hypotheses on the anticipated direction of effect of receipt of psychotherapy on STD and LTD claim

closure. They predicted that individuals in the STD group are less likely to benefit from psychotherapy given the time taken for CBT to be successfully implemented, which may take them into the LTD timeframe. They predicted that individuals in the LTD group are more likely to benefit from psychotherapy. Further, they hypothesized that if there was a difference in the anticipated directions between the two groups, it may be due to the increased severity of illness or secondary gains in those receiving psychotherapy in the STD group. Table 10.1 provides a description of all independent variables considered in our models and our predictions on the anticipated direction of effect on disability claim closure.

Disability funding policies can be purchased by employers under three types of financial arrangements: non-refund policies where the insurer approves and funds services and treatments, refund policies in which the insurer and the plan sponsor (e.g., the employer) shares the funding for services and treatments, and administrative services only (ASO) policies in which the plan sponsor approves and pays for all services and treatments. For LTD claims, all types of policies require funds to be put aside (as reserves) that amount to two-thirds of the claimant's pre-disability income that would be earned from the age their long-term benefits began until the age of 65. Under a non-refund policy, the reserves are funded by the insurer, and released back to the insurer if the claim resolves. Under refund or ASO policies, the reserves for LTD claims are funded by the employer, and released back to the employer if the claim

Table 10.1: Description of variables

Variable	Description	Anticipated direction of claim closure
<i>Claimant demographic variables</i>		
Age	Age of claimant at disability	Older age: (-)
Gender	Gender of claimant	Females: (-)
Province	Province the claimant resides in	Ontario, Quebec: (+)
Industry	Type of industry the claimant is working in (blue collar, grey collar, white collar*)	White, grey collar: (-)
Salary	Salary of claimant	Higher salary: (-)
ICD-10 primary diagnosis	Primary diagnosis of claimant (major depression or recurrent depression)	Recurrent depression: (-)
ICD-10 secondary diagnosis	Secondary diagnosis of claimant (none, psychological diagnosis, non-psychological diagnosis)	Secondary diagnosis: (-)
<i>Claim coverage variables</i>		
Time to claim registration	Duration from claimant's disability date to disability claim registration date	Longer time to claim registration: (-)
Time to claim approval	Duration from disability claim registration date to disability claim contractual approval date	Longer time to claim approval: (-)
Elimination period	Duration from claimant's disability date to first payment date	Longer elimination period: (-)
Maximum claim benefit period	Duration from disability claim contractual approval date to maximum claim benefit date	Longer claim benefit period: (-)
Location of claim office	Office where the claim is currently managed (Edmonton, Montreal, Toronto, Vancouver, Waterloo)	Toronto, Montreal: (+)
Funding type	Funding arrangement of claim (non-refund, refund, administrative services only [ASO])	Refund, ASO: (-)
Total reserve amount	Reserves held on claims (LTD only)	Higher reserves: (+)
Receipt of psychotherapy	If claimant has received psychotherapy or not	Receipt of psychotherapy: (0 for STD / + for LTD)

ICD: International classification of diseases; +: associated with faster claim closure; -: associated with slower claim closure; 0: associated with similar resolution; * - classifications of industry in Table 10.2; LTD - long term disability

resolves. For our regression models, we used non-refund policy as the reference group. Two authors (SE and JWB) independently grouped 66 different industries into blue-collar, grey-collar and white-collar industries (94% agreement) and reached consensus through discussion (Table 10.2).

10.5 Outcomes

Our primary outcome was time to claim closure, defined as the duration from disability claim approval until the closure/resolution of the claim. Our secondary outcome was receipt of psychotherapy.

10.6 Data management and data cleaning

We screened all data to identify outliers, inconsistencies and missing data by calculating summary statistics, and exploring distributions graphically. If clear outliers and inconsistencies were identified, we worked with Sun Life Financial to correct the data. If inconsistencies could not be corrected, we treated them as missing data. We excluded variables that were missing for more than 10% of claimants. Of the variables that were not excluded, less than 1% was missing.

10.7 Statistical analysis

We generated frequencies for all collected data. We reported the mean and standard deviation (SD) of continuous variables that were normally distributed, the median and

Table 10.2: Classification of industry

White collar	Grey collar	Blue collar
Banking	Apparel & other finished products of fabrics & similar material	Agricultural production-livestock & animal specialties
Business services	Communication	Auto dealers & gasoline stations
Credit agencies	Eating drinking places	Automotive repair, services & parking
Educational service	Food stores	Bituminous coal & lignite mining
Federal government	Furniture, home furnishings	Building material & farming equipment
Holding related investments	General merchandise	Chemicals, allied products
Insurance agent, broker, service	Hotels, rooming houses, camps & other lodging places	Electrical equipment
Insurance carriers	Local passenger train	Electricity, gas, sanitary service
Legal services	Miscellaneous retail stores	Fabricated metal products
Local government	Miscellaneous services	Food and kindred products
Medical, related health	Motion pictures	Forestry
Nonprofit membership organization	Personal services	General building Construction
Provincial government	Printing, publishing	Heavy construction
Real estate	Retail-apparel & accessory stores	Instruments and related products
Social services	Security & commodity brokers, dealers, exchanges & services	Lumber, wood products
	Services-amusement & recreation services	Metal mining
	Transportation by air	Mining & quarrying of nonmetallic minerals (no fuels)
	Transportation service	Miscellaneous manufacturing
	Water transportation	Miscellaneous repair service
	Wholesale durables	Oil gas extraction
	Wholesale nondurables	Paper allied products
		Primary metals
		Railroad transportation
		Rubber & plastics
		Special trade contractors
		Stone Clay, glass products
		Textile mill products
		Tobacco manufacturers
		Transportation equipment
		Trucking, warehousing

interquartile range (IQR) for continuous variables that were not normally distributed (assessed through residual analysis and computing kurtosis and skewness measures [skewness and kurtosis measures of +2 to -2 considered normal]), and the number of occurrences with proportions represented as percentages for categorical variables.

We tested for collinearity to assess if a predictor was highly correlated with another (correlation coefficient $r > 0.5$) using a correlation matrix. If two variables were highly correlated, we removed the variable that was considered to be of lesser importance, as guided by the administrators at Sun Life Financial and content experts on our team.

We performed a time-to-event analysis using Cox proportional hazards regression to assess the association between time to claim closure and the independent variables. Receipt of psychotherapy was treated as a time-dependent covariate to account for when it was initiated during the course of the disability claim. For STD claims that were unresolved 26 weeks after claim approval, we used 181 days (26 weeks [the more common STD benefit duration] minus 1 day) as our censoring point to maintain the proportionality assumption. For LTD claims that were unresolved when the data was extracted, we used the date of data extraction as our censoring point. For our secondary analysis, we performed an adjusted logistic regression to assess the association between receipt of psychotherapy and potentially predictive factors. To avoid overfitting our models, we required at least 10 events per variable for our Cox regression model and 10 events of the least common outcome—receipt of

psychotherapy—for our logistic regression model [23]. Our regression model excludes independent variables with less than 200 observations unless we were able to collapse them with other related variables to exceed this threshold. We calculated hazard ratios (HRs) for our time-to-event analyses and odds ratios (ORs) for our logistic regression analyses, their associated 99% confidence intervals (CIs), the unstandardized beta coefficients for each variable and the associated p-values. In order to be more stringent and minimize the likelihood of spurious findings, we considered an independent variable as statistically significant if it had a p-value of less than or equal to 0.01 in each final adjusted model.

We performed bootstrapping for our regression models to measure the accuracy of our sample estimates [24], and performed the Hosmer-Lemeshow test to assess the goodness-of-fit in our logistic regression model.

We used SPSS v20.0 to perform all statistical analyses.

CHAPTER 11

RESULTS

11.1 Description of claims

Of 13,758 STD and 11,275 LTD claims received with a primary diagnosis of depression, 3250 (24%) STD and 937 (8%) LTD claims were excluded due to the claim being declined, exceeding a claim benefit period of 2.5 years (STD only), or having their claim paid retroactively. Our final analysis included 10,508 STD and 10,338 LTD claims. Depression management included psychotherapy in 261 STD claims and 1,582 LTD claims. Table 11.1 presents baseline characteristics of all factors for STD and LTD claims.

We excluded the following variables from our regression models: total reserve amount for STD model as there are no reserves put aside for STD claims; claim office due to high correlation with province of residence; time to receipt of disability benefits and elimination period (date of disability to the date the claim was first paid) due to high correlation with time to claim approval; and type of employment industry from our LTD models due to a high frequency (27.4%) of missing data. Due to small numbers of observations from some provinces, we merged Prince Edward Island, Newfoundland, Nova Scotia and New Brunswick into one category, “Maritimes”, and

Table 11.1: Baseline characteristics

Variables	STD n(%)	LTD n(%)
Total claimants	10508	10338
Claim closure reason		
Censored at time of data extraction	173 (1.6%)	3670 (35.5%)
Return to work	3390 (32.3%)	4542 (43.9%)
Anticipated return to work	2900 (27.6%)	0 (0%)
Maximum benefit date reached	3003 (28.6%)	0 (0%)
Offsets exceed benefits	4 (0.04%)	8 (0.1%)
Retirement/terminal age	11 (0.1%)	141 (1.4%)
Own occupation termination	0 (0%)	451 (4.4%)
No longer disabled	305 (2.9%)	1092 (10.6%)
Claim transfer or settlement	0 (0%)	57 (0.6%)
Rehabilitation settlement	0 (0%)	280 (2.7%)
Litigation settlement	0 (0%)	31 (0.3%)
Other (including strike and securing info)	715 (6.8%)	20 (0.2%)
Death	7 (0.1%)	41 (0.4%)
Age: Mean (SD) years	43.3 (10.1)	47.0 (9.3)
Sex		
Male	5859 (55.8%)	3380 (67.3%)
Female	4640 (44.2%)	6958 (32.7%)
Salary per month: Median (IQR)	\$3337 (\$2607 to \$4450)	\$4344 (\$3204 to \$5154)
Industry		
White-collar	2502 (23.8%)	2904 (28.1%)
Grey-collar	3474 (33.1%)	2112 (20.4%)
Blue-collar	4532 (42.1%)	2490 (24.1%)
Unknown	0 (0.0%)	2832 (27.4%)
Primary diagnosis		
Major depression	10004 (95.2%)	9034 (87.4%)
Recurrent depression	504 (4.8%)	1304 (12.6%)
Secondary diagnosis		
None	8043 (76.5%)	6657 (64.4%)
Psychological diagnosis	1746 (16.6%)	2337 (22.6%)
Nonpsychological diagnosis	719 (6.8%)	1344 (13.0%)
Province		
British Columbia	842 (8.0%)	966 (9.3%)
Alberta**	1736 (16.5%)	1112 (10.8%)
Saskatchewan	237 (2.2%)	119 (1.2%)
Manitoba	155 (1.5%)	155 (1.5%)
Ontario	3798 (36.1%)	3532 (34.2%)
Quebec	2956 (28.1%)	3922 (37.9%)
New Brunswick**	105 (1.0%)	189 (1.8%)
Nova Scotia**	537 (5.1%)	224 (2.2%)
Prince Edward Island**	5 (0.05%)	35 (0.3%)
Newfoundland**	139 (1.3%)	83 (0.8%)
Claim office		
Vancouver	507 (4.8%)	798 (7.7%)
Edmonton	1953 (18.6%)	959 (9.3%)
Waterloo	2702 (25.7%)	1534 (14.8%)

Toronto	1707 (16.2%)	1084 (10.5%)
Montreal	3639 (34.6%)	5963 (57.6%)
Funding type		
ASO	4042 (38.5%)	645 (6.2%)
Non-Refund	3376 (32.1%)	4399 (42.6%)
Refund	3090 (29.4%)	5294 (51.2%)
Time to claim receipt months: median (IQR)	0.3 (0.2 to 0.6)	4.0 (2.8 to 5.3)
Time to claim approval months: median (IQR)	0.4 (0.2 to 0.7)	1.5 (0.9 to 2.3)
Reserve amount: median (IQR)	N/A	\$283,678 (\$121,742 to \$629,813)
Receipt of psychotherapy		
No	10247	8756
Yes	261	1582
Time to initiate psychotherapy weeks: median (IQR)	10.1 (6.8 to 14.9)	17.1 (8.7 to 34)

STD- Short term disability; LTD- Long-term disability; SD- Standard deviation; IQR- Interquartile range;; * - Excluded from analyses; ** - merged as 'Maritimes' for our analyses; *** - merged as 'Prairies' for our analyses; ^ - missing for 98.5% of cases; N/A - not applicable due to no reserves for STD

Alberta, Saskatchewan and Manitoba into “Prairies”. We excluded Yukon and Northwest Territories as they had fewer than 200 observations and merging them with other provinces was not considered appropriate.

11.2 Short-term disability

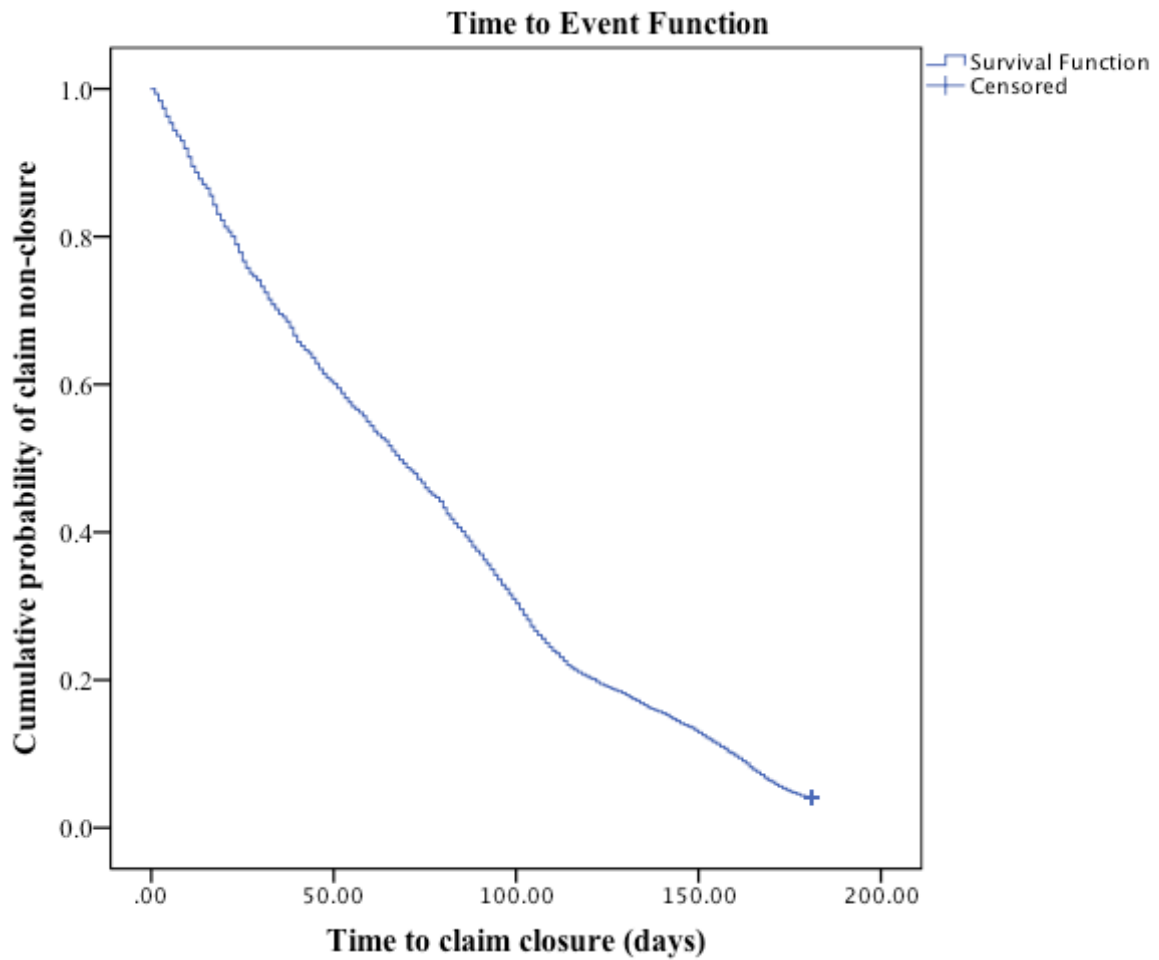
Of 10,508 STD claims due to depression, 10,335 (98.4%) were closed prior to a maximum STD benefit duration of 26 weeks and 173 (1.6%) were censored. Of the 10,335 closed claims, 3390 (32.8%) returned to work and 2900 (28.1%) were expected to return to work (Table 11.1).

Figure 11.1 presents the time to closure survival curve for STD claimants.

11.2.1 Factors associated with time to STD claim closure

Our adjusted regression analysis showed that psychotherapy was associated with longer time to STD claim closure (HR [99% CI] = 0.81 [0.68 to 0.97]). Older age (0.90 [0.88 to 0.92], per decade), female gender (0.92 [0.87 to 0.97]), working in a white-collar industry (0.86 [0.80 to 0.92]), higher salary (0.87 [0.82 to 0.93]), a primary diagnosis of recurrent depression versus non-recurrent major depression (0.78 [0.69

Figure 11.1: Kaplan Meier time to event curve of time to STD claim closure



to 0.87], a psychological secondary diagnosis (0.90 [0.84 to 0.97] or a non-psychological secondary diagnosis (0.81 [0.73 to 0.90]), a longer maximum claim benefit period duration (0.87 [0.86 to 0.88], and an administrative services only (ASO) ([0.94 [0.88 to 1.00]) or refund policy (0.86 [0.80 to 0.92]) compared to non-refund policy claims, and residing in Quebec compared to Ontario (0.73 [0.68 to 0.78] were independently associated with slower STD claim closure (Table 11.2).

Longer time from claim registration to claim approval (1.11 [1.07 to 1.15]), and residing in the Prairies (1.08 [1.00 to 1.16]) or the Maritimes (1.19 [1.07 to 1.32]) compared to Ontario were associated with faster STD claim closure (Table 11.2).

11.2.2 Factors predictive of receipt of psychotherapy among STD claims

Working in a white-collar industry (OR [99% CI]=2.59 [1.64 to 4.08]) or a grey-collar industry (1.99 [1.27 to 3.13]) compared to a blue-collar industry, and a refund policy compared to a non-refund policy claim (1.58 [1.05 to 2.38]) were associated with a higher likelihood of receiving psychotherapy while in receipt of STD benefits.

Table 11.2: Factors predictive of time to short-term disability claim closure

Factor	p-value	HR	99.0% CI for HR	
			Lower	Upper
Receipt of psychotherapy	.002	.813	.684	0.966
Age (per 10 years)	<.001	.902	.878	.926
Gender				
Female	<.001	.915	.865	.968
Male (reference group)		1		
Industry				
White collar	<.001	.848	.790	.911
Grey collar	.013	.941	.884	1.002
Blue collar (reference group)		1		
Salary (per \$1000 per week)	<.001	.872	.815	.934
ICD-10 primary diagnosis				
Major depression (reference group)	<.001	1	.687	.876
Recurrent depression				
ICD-10 secondary diagnosis				
Psychological diagnosis	<.001	.904	.842	.971
Non-Psychological diagnosis	<.001	.814	.733	.904
None (reference group)		1		
Time to approval (months)	<.001	1.105	1.065	1.147
Maximum benefit period (months)	<.001	.870	.859	.881
Province				
British Columbia	.434	1.031	.931	1.142
Prairies*	.010	1.077	1.000	1.160
Quebec	<.001	.713	.666	.764
Maritimes**	<.001	1.188	1.068	1.320
Ontario (reference group)		1		
Funding type				
ASO	.010	.939	.881	1.000
Refund	<.001	.855	.798	.915
Non-refund (reference group)		1		

HR – hazard ratio; CI – confidence interval; ICD-10 – International Classification of Diseases version 10; * - Consists of Alberta, Saskatchewan, and Manitoba; ** - Consists of New Brunswick, Newfoundland, Nova Scotia and Prince Edward Island; ASO – Administrative Services Only

Factors that have a p-value of less than 0.01 are significant predictors of claim closure

An HR of greater than 1 is associated with faster claim closure; an HR of less than 1 is associated with slower claim closure

11.3 Long-term disability

Of 10,338 LTD claims due to depression, 6668 (65%) were closed and 3670 (35%) were censored when our data was captured. Of the 6668 closed claims, 4542 (68.1%) returned to work (Table 11.1).

Figure 11.2 presents the time to closure survival curve for LTD claimants.

11.3.1 Factors predictive of time to claim closure

Our adjusted regression analysis showed that receipt of psychotherapy (HR [99% CI] = 1.42 [1.30 to 1.55]) was independently associated with faster claim closure. Older age (0.83 [0.80 to 0.85]), a primary diagnosis of recurrent depression (0.80 [0.72 to 0.89]), a psychological secondary diagnosis (0.66 [0.61 to 0.71]) or non-psychological secondary diagnosis (0.77 [0.71 to 0.83]), longer time from claim registration to claim approval (0.98 [0.96 to 0.997]), and ASO (0.87 [0.75 to 0.996]) or refund (0.73 [0.68 to 0.78]) policy claims compared to non-refund policy claims were associated with slower claim closure.

Residing in the Prairies (1.46 [1.31 to 1.61]) or Quebec (1.93 [1.78 to 2.09]) versus Ontario were associated with faster claim closure (Table 11.3).

Figure 11.2: Kaplan Meier time to event curve of time to LTD claim closure

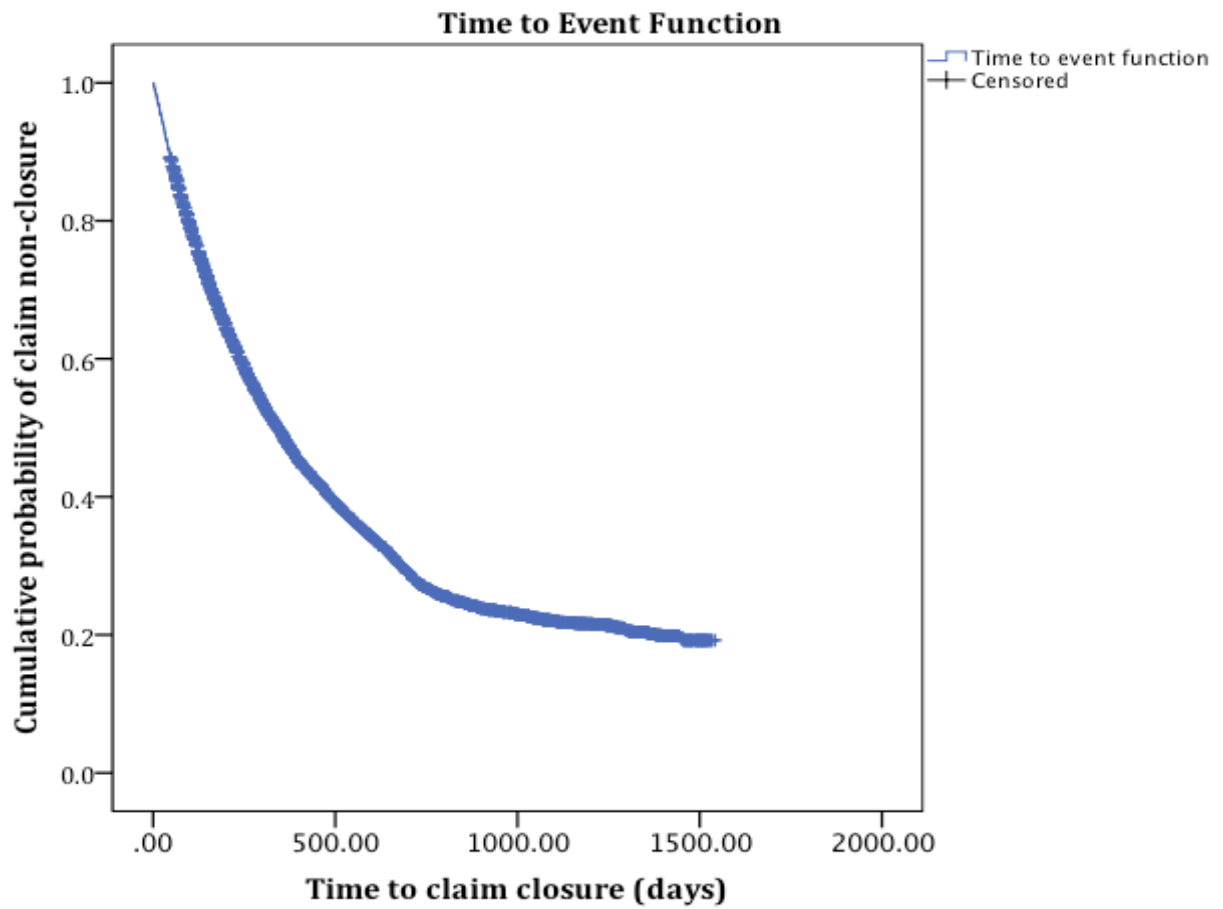


Table 11.3: Factors predictive of time to long-term disability claim closure

Factor	p-value	HR	99.0% CI for HR	
			Lower	Upper
Receipt of psychotherapy	<.001	1.417	1.297	1.547
Age (per 10 years)	<.001	.825	.797	.854
Gender				
Female	.091	.956	.892	1.024
Male (reference group)		1		
Salary (per \$1000 per week)	.278	1.076	.904	1.280
ICD-10 primary diagnosis				
Major depression (reference group)		1		
Recurrent depression	<.001	.803	.724	.891
ICD-10 secondary diagnosis				
Psychological diagnosis	<.001	.765	.705	.830
Non-psychological diagnosis	<.001	.659	.594	.732
None (reference group)		1		
Time to approval (months)	.003	.978	.959	.997
Province				
British Columbia	.197	.937	.823	1.067
Prairies*	<.001	1.455	1.313	1.611
Quebec	<.001	1.928	1.782	2.086
Maritimes**	.613	.968	.820	1.143
Ontario (reference group)		1		
Funding type				
ASO	.008	.866	.754	.996
Refund	<.001	.729	.681	.780
Non-refund (reference group)		1		

HR – hazard ratio; CI – confidence interval; ICD-10 – International Classification of Diseases version 10; * - Consists of Alberta, Saskatchewan, and Manitoba; ** - Consists of New Brunswick, Newfoundland, Nova Scotia and Prince Edward Island; ASO – Administrative Services Only

Factors that have a p-value of less than 0.01 are significant predictors of claim closure

An HR of greater than 1 is associated with faster claim closure; an HR of less than 1 is associated with slower claim closure

11.3.2 Factors predictive of receipt of psychotherapy among LTD claims

Older age (OR [99% CI] = 0.90 [0.83 to 0.97]), a non-psychological secondary diagnosis (0.78 [0.62 to 0.98]), residing in Quebec compared to Ontario (0.53 [0.44 to 0.64]), and an ASO (0.58 [0.42 to 0.80]) or refund (0.70 [0.60 to 0.82]) policy claims compared to non-refund were associated with a lower likelihood of receiving psychotherapy. Females versus males (1.20 [1.02 to 1.41]), and residing in the prairies versus Ontario (1.32 [1.08 to 1.62]) were associated with a higher likelihood of receiving psychotherapy.

Table 11.4 provides a summary of factors that were independently associated with time to STD and LTD claim closure compared to our anticipated direction of effect.

Table 11.4: Comparison between predictors associated with time to claim closure for short-term disability versus long-term disability claims

Predictors associated with time to claim closure	STD	LTD	Anticipated direction
Receipt of psychotherapy	-	+	0/+
Older age	-	-	-
Female (compared to males)	-	NS*	-
White collar industry (compared to blue collar industry)	-	N/A	-
Grey collar industry (compared to blue collar industry)	NS*	N/A	-
Higher salary	-	NS**	-
Recurrent depression (compared to major depression)	-	-	-
Secondary psychological diagnosis (compared to no diagnosis)	-	-	-
Secondary nonpsychological diagnosis (compared to no diagnosis)	-	-	-
Longer maximum benefit period duration	-	N/A	-
Longer time from claim registration to claim approval	+	-	-
Prairies (compared to Ontario)	+	+	-
Quebec (compared to Ontario)	-	+	0
Maritimes (compared to Ontario)	+	NS*	-
ASO funding (compared to non-refund)	-	-	-
Refund funding (compared to non-refund)	-	-	-

Shaded rows represent associations that were in the opposite directions of what we anticipated

STD: short-term disability; **LTD**: long-term disability; +: associated with faster claim closure; -: associated with slower claim closure; 0: associated with similar resolution; **NS***: not significant but consistent with anticipated direction; **NS****: not significant and not consistent with anticipated direction; **N/A**= not included in the model

CHAPTER 12

DISCUSSION

12.1 Summary of main results

Our study, evaluating the effect of psychotherapy on disability benefit claim closure in patients suffering from a primary diagnosis of depression, found that receipt of psychotherapy was associated with longer time to STD closure and faster LTD claim closure. For STD and LTD claims, older age, a primary diagnosis of recurrent depression (compared to major depression), a secondary psychological or non-psychological diagnosis (versus no secondary diagnosis), and an administrative services only or a refund policy (compared to a non-refund policy) were commonly predictive of slower claim closure. We found no common predictors that were independently associated with receipt of psychotherapy for both STD and LTD claims.

12.2 Strengths and limitations

The strengths of our study included *a priori* creation of regression models and the anticipated direction of included independent variables. Other strengths include limited missing data, correction of identifiable data errors and inconsistencies, and validation checks to ensure the accuracy of our sample estimates from our regression models.

Our study has several limitations. First, this was a retrospective cohort study in which the reasons for administering psychotherapy are uncertain. Thus, despite our adjusted models, it remains possible that selection bias affected our findings; STD claimants who received psychotherapy were less likely to resolve their claims and LTD claimants chosen to receive psychotherapy were more likely to resolve their claims irrespective of the intervention. Second, a number of variables that may be important to consider were unavailable (e.g., baseline severity of depression, patient expectations regarding recovery, and the use of antidepressants), and some variables were not optimally collected. For example, all psychotherapies were categorized as an aggregate variable and the specific type of psychotherapy provided was not available, although the insurer felt that majority of psychotherapy administered was CBT. Third, our association between psychotherapy and longer time to STD claim closure may represent a misleading finding: the effects of some common forms of psychotherapy (e.g., CBT) typically take months [i.e., greater than 3 months] to manifest [25], meaning that even an effective therapy may not show an effect during the limited time that an STD claim is paid out. Further, it is possible that those receiving psychotherapy in the STD group had more severe depression or may have been motivated by secondary financial gains (e.g., being approved for LTD benefits). Finally, our primary outcome, claim closure, reflects to only a limited extent the more important outcome of sustained return to work [26].

12.3 Findings in context with previous evidence

The association between receipt of psychotherapy and claim closure or return to work has received limited attention in the published literature. However, a recent article reported the effectiveness of a pilot vocationally oriented CBT in assisting very long-term unemployed individuals return to work [27], and a recent RCT reported that individuals treated with work-focused CBT returned to work on average 65 days earlier than those receiving traditional CBT [28].

Our findings are consistent with previous evidence that shows that older age [22,29,30,31], higher salary [29], and comorbidities (presence of secondary diagnoses) are associated with worse recovery [31,32]. We found inconsistent association of female gender with claim closure between STD and LTD claims. Previous evidence is also inconsistent with a recent review finding no significant association [22] and two other reviews finding female gender as a significant predictor of longer duration of sick leave [29,33]. Differences in the results may be explained by the different conditions studied and/or adjusting for different prognostic factors in the regression models.

We found that longer claim approval time was associated with longer time to LTD claim closure but faster STD claim closure, and ASO or refund policies were associated

with longer time to claim closure in both STD and LTD claims. These associations had not been previously reported. Longer approval times for LTD claims may delay treatment initiation, which potentially delays recovery (and claim closure). Although we found an opposite association for STD claims, STD approval times were found to be relatively short (median (interquartile range [IQR]) of 0.4 [0.2 to 0.7] months) compared to LTD approval times (1.5 [0.9 to 2.3] months), and thus this may not represent an important effect.

The association of ASO or refund policies with delayed recovery may be explained by differences in the amount of rehabilitation services provided. For example, among LTD claims, we found that non-refund policies were significantly more likely to receive psychotherapy than other policy types: 13% of claimants with ASO and refund policies received psychotherapy compared to 18% of claimants with a non-refund policy. In our model, we were only able to adjust for psychotherapy and there could be differences in other services (e.g., work hardening program) provided between the policy types. These findings warrant further exploration.

Although we did not predict that the Prairies region would have faster claim closure in both STD and LTD claims and Quebec would have faster claim closure in LTD claims (compared to Ontario), post-hoc discussions with insurance administrators suggested that these associations may be explained by the economic growth in the Prairies

region. These discussions also suggested that greater number of mental health claims treated in Quebec may have resulted in a more established infrastructure facilitating claim resolution. However, we are uncertain as to why the same effect of faster claim closure in Quebec was not illustrated for STD claims.

12.4 Future directions

Our findings reveal uncertainty about the true effect of psychotherapy on time to claim resolution, among patients with depression. A prospective study with careful measurement of all putative determinants of claim resolution would strengthen the evidence; a randomized controlled trial could settle the issue definitively. Such research is required before payers and clinicians can confidently decide whether they should decrease, continue or expand the use of psychotherapy in managing these patients.

CHAPTER 13

CONCLUSIONS

Our study found, paradoxically, that receipt of psychotherapy is significantly associated with longer time to claim closure in individuals receiving STD, and faster claim closure in patients receiving LTD. We also found evidence to suggest that age, presence and type of diagnoses, type of policy funding, gender, salary, industry, time to claim approval and province claimants reside in are predictive of claim closure. Establishing the causal effect of psychotherapy on claim resolution will require well-designed prospective studies.

ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS

Acknowledgements

We thank Drs. Peter Bieling and Randi McCabe for their assistance in developing hypotheses for our study, and Mr. Cory Williams for helpful discussions regarding the Sun Life administrative database.

Funding

There were no funds received for this study. Shanil Ebrahim was supported by MITACS Accelerate scholarship (partnered with Sun Life Financial Assurance Company of Canada) and the Canadian Institutes of Health Research Doctoral Award during the completion of this study. Jason W. Busse is supported by a New Investigator Award from the Canadian Institutes of Health Research and Canadian Chiropractic Research Foundation.

Conflicts of interest

Jason W. Busse acts as a consultant to Prisma Health Canada, a private incorporated company funded by employers and insurers that consults on and manages long-term

disability claims. Marg Bellman is the assistant vice president of the National Disability Services, Policy & Procedure department at Sun Life Financial Assurance Company of Canada. All other authors report no conflicts of interest.

Contributions

Shanil Ebrahim was responsible for preparation of the first draft of the manuscript. Shanil Ebrahim completed the statistical analyses with assistance and consultation from Stephen Walter, Diane Heels-Ansdell, Jason Busse, Steven Hanna, and Gordon Guyatt. Shanil Ebrahim, Jason Busse, Stephen Walter and Gordon Guyatt were responsible for the conception and idea of the manuscript. Marg Bellman was the knowledge user of this study and contributed to the development of the hypotheses. Shanil Ebrahim is the first author and Jason Busse is the senior author of the manuscript. All authors critically reviewed several drafts of the manuscript. All authors read and approved the final manuscript to be published.

REFERENCES

1. Lim K, Jacobs P, Ohinmaa A, Schopflocher D, Dewa CS (2008) A new population-based measure of the economic burden of mental illness in Canada. *Chronic Diseases in Canada* 28: 92-98.
2. Hirschfeld R, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, Moller HJ, Healy D, Baldwin D, Humble M, Versiani M, Montenegro R, Bourgeois M (2000) Social functioning in depression: a review. *J Clin Psychiatry* 61: 268-275.
3. Moore R, Mao Y, Zhang J, Clarke K, Laboratory Centre for Disease Control (1997) *Economic Burden of Illness in Canada*. Ottawa: Health Canada.
4. Stephens T, Joubert N (2001) The economic burden of mental problems in Canada. *Chronic Diseases in Canada* 22: 1-10.
5. Brown P (2001) Effective treatments for mental illness not being used, WHO says. *BMJ* 323: 769.
6. Mor N, Haran D (2009) Cognitive-behavioral therapy for depression. *The Israel Journal of Psychiatry and Related Sciences* 46: 269-273.
7. World Health Organization (WHO) (2002) *Mental Health: New understanding, New hope*. Geneva: WHO.

8. Dobson KS (2001) Handbook of cognitive-behavioral therapies. 2nd Edition. New York; London: Guilford Press.
9. NICE (2009) Depression: Treatment management of depression in adults. London: National Institute for Clinical Excellence.
10. Kouzis AC (2000) Psychopathology and the initiation of disability payments. *Psychiatr Serv* 51: 908-913.
11. Kouzis AC, Eaton WW (1994) Emotional disability: prevalence and predictors. *American Journal of Public Health* 84: 1304-1307.
12. Kouzis AC, Eaton WW (1997) Psychopathology and the development of disability. *Social and Psychiatry and Psychiatric Epidemiology* 32: 379-386.
13. Salkever DS, Goldman H, Purushothaman M, Shinogle J (2000) Disability management, employee health and fringe benefits, and long-term-disability claims for mental disorders: an empirical exploration. *The Milbank Quarterly* 78: 79-113.
14. Dewa C, Hoch JS, Goering P, et al (2003) Use of antidepressants among Canadian workers receiving depression-related short-term disability benefits. *Psychiatr Serv* 54: 724-729.
15. Leahy RL (2003) Roadblocks in Cognitive-Behavioral Therapy: Transforming Challenges into Opportunities for Change. New York: The Guilford Press.

16. Harris I, Mulford J, Solomon M, van Gelder JM, Young J. (2005) Association between compensation status and outcome after surgery: a meta-analysis. *JAMA* 293: 1644-1652.
17. Bellman M (2010) Effectiveness of CBT on individuals receiving disability benefits. Personal communication ed. Toronto.
18. Ebrahim S, Montoya L, Truong W, Hsu S, Kamal el Din M, Carrasco-Labra A, Busse JW, Walter S, Heels-Ansdell D, Couban R, Patelis-Siotis I, Bellman M, de Graaf LE, Dozois D, Bieling P, Guyatt GH (2012) Effectiveness of Cognitive Behavioral Therapy for depression in patients receiving disability benefits: A systematic review and individual patient meta-analysis. *PLoS ONE* 7: e50202.
19. Kronstrom K, Karlsson H, Nabi H, Oksanen T, Salo P, Sjosten N, Virtanen M, Pentti J, Kivimaki M, Vahtera J (2011) Optimism and pessimism as predictors of work disability with a diagnosis of depression: A prospective cohort study of onset and recovery. *J Affect Disorder* 130: 294-299.
20. Nielson MBD, Madsen IEH, Bultmann U, Christensen U, Diderichsen F, Rugulies R (2010) Predictors of return to work in employees sick-listed with mental health problems: findings from a longitudinal study. *European Journal of Public Health* 21: 806-811.

21. Romero CC, Sierra MC, Gutierrez JCS, Barrenechea Albarran JL, Coneso AM, Bonacho EC (2011) Variables moduladoras de la discapacidad asociada al trastorno depresivo. *Rev Calid Asist* 26: 39-46.
22. Cornelius LR, van der Klink JJ, Groothoff JW, Brouwer S (2011) Prognostic factors of long term disability due to mental disorders: a systematic review. *J Occup Rehabil* 21: 259-274.
23. Norman G, Streiner S (2008) *Biostatistics: The bare essentials*. Shelton, Connecticut: People's Medical Publishing House.
24. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eikemans MJ, Vergouwe Y, Habbema JD (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology* 54: 774-781.
25. Hunter MS, Ussher JM, Browne SJ, Cariss M, Jelley R, Katz M (2002) A randomized comparison of psychological (cognitive behavior therapy), medical (fluoxetine) and combined treatment for women with premenstrual dysphoric disorder. *J Psychosom Obstet Gynaecol* 23: 193-199.
26. Steenstra IA, Lee H, de Vroome EM, Busse JW, Hogg-Johnson SJ (2012) Comparing Current Definitions of Return to Work: A Measurement Approach. *J Occup Rehabil* 22: 394-400.

27. Rose VK, Perz J, Harris E (2012) Vocationally oriented cognitive behavioural training for the very long-term unemployed. *Occup Med* 8 May.
28. Lagerveld SE, Blonk RW, Brenninkmeijer V, Wijngaards-de Meij L, Schaufeli WB (2012) Work-focused treatment of common mental disorders and return to work: a comparative outcome study. *J Occup Health Psychol* 17: 220-234.
29. Steenstra IA, Verbeek JH, Heymans MW, Bongers PM (2005) Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occup Environ Med* 62: 851-860.
30. Dekkers-Sánchez PM, Hoving JL, Sluiter JK, Frings-Dresen MH (2008) Factors associated with long-term sick leave in sick-listed employees: a systematic review. *Occup Environ Med* 65: 153-157.
31. Blank L, Peters J, Pickvance S, Wilford J, Macdonald E (2008) A Systematic Review of the Factors which Predict Return to Work for People Suffering Episodes of Poor Mental Health. *J Occup Rehabil* 18: 27-34.
32. Werner EL, Cote P (2009) Low back pain and determinants of sickness absence. *Eur J Gen Pract* 15: 74-79.
33. Detaille S, Heerkens YF, Engels JA, van der Gulden JW, van Dijk FJ. (2009) Common prognostic factors of work disability among employees with a chronic somatic disease: a systematic review of cohort studies. *Scan J Work Environ Health* 35: 261-281.

STUDY 3

Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers

Shanil Ebrahim¹, Elie A. Akl^{1,2}, Reem A. Mustafa¹, Xin Sun^{1,3}, Stephen D. Walter^{1,4}, Diane Heels-Ansdell¹, Pablo Alonso-Coello⁵, Bradley C. Johnston^{1,6,7}, Gordon H. Guyatt^{1,8}[§]

¹Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada

²Department of Internal Medicine, American University of Beirut, Beirut, Lebanon

³The Centre for Health Research, Kaiser Permanente Northwest, Portland, USA

⁴Department of Mathematics and Statistics, McMaster University, Hamilton, Canada

⁵Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Barcelona, Spain

⁶Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, Canada

⁷Child Health Evaluative Sciences, SickKids Research Institute, Toronto, Canada

⁸Department of Medicine, McMaster University, Hamilton, Canada

[§]Corresponding author

Corresponding Author

Gordon Guyatt, Department of Clinical Epidemiology and Biostatistics, McMaster University, 1280 Main Street West, Rm. 2C12, Hamilton, Ontario, Canada, L8S 4K1.
Email: guyatt@mcmaster.ca

Keywords

Missing participant data, continuous outcomes, risk of bias, systematic reviews, meta-analysis, lost to follow-up

Accepted for publication by Journal of Clinical Epidemiology

CHAPTER 14

ABSTRACT

Background: No methods directly address the impact of missing participant data for continuous outcomes in systematic reviews on risk of bias.

Methods: We conducted a consultative, iterative process to develop a framework for handling missing participant data for continuous outcomes. We considered sources reflecting real observed outcomes in participants followed-up in individual trials included in the systematic review, and developed a range of plausible strategies. We applied our approach to two systematic reviews.

Results: We use 5 sources of data for imputing the means for participants with missing data. To impute SD, we use the median SD from the control arms of all included trials. Using these sources, we developed four progressively more stringent imputation strategies. In the first example review, effect estimates diminished and lost significance as strategies became more stringent, suggesting rating down confidence in estimates of effect for risk of bias. In the second, effect estimates maintained statistical significance using even the most stringent strategy, suggesting missing data does not undermine confidence in results.

Conclusions: Our approach provides a useful, reasonable and relatively simple, quantitative guidance for judging the impact of risk of bias as a result of missing participant data in systematic reviews of continuous outcomes.

CHAPTER 15

What this study adds?

[1] Specific guidance for addressing missing participant data for continuous outcomes in meta-analyses and for judging the associated risk of bias is currently unavailable.

[2] We developed an approach consisting of four increasingly stringent data imputation strategies to address these two issues.

[3] Our approach provides a useful, reasonable and relatively simple, quantitative guidance for judging the impact of risk of bias as a result of missing participant data in systematic reviews of continuous outcomes.

CHAPTER 16

INTRODUCTION

Greater than 80% of randomized controlled trials (RCTs) published in top general medical journals suffer from missing participant data [1]. If participants with missing data have different outcomes from those with available data, it may introduce bias in the results of the individual trials and of systematic reviews and meta-analyses using those results. For example, participants' experience of adverse outcomes or toxic effects from therapy may lead them to withdraw from the trial. Alternatively, participants could do very well, making them less interested in continuing the treatment and leading to their withdrawal from the trial.

The Cochrane Collaboration has proposed strategies for handling missing participant data for dichotomous outcomes in systematic reviews [2]. We proposed additional strategies that use plausible assumptions regarding outcomes of trial participants with missing data [3]. A systematic survey of RCTs published in prestigious general medical journals found that applying these strategies could change the interpretation of results in up to 33% of RCTs [1].

Individual trials with continuous outcomes sometimes analyze only participants with available data—complete case (or available case) analysis. Alternatively individual trials commonly use single imputation techniques such as last observation carried forward (LOCF) and baseline observation carried forward (BOCF), and multiple imputation techniques. LOCF replaces the missing value with the participant’s last known value while BOCF replaces the missing value with the participant’s baseline value. These single imputation techniques make an unlikely assumption that outcomes stay constant, and ignore uncertainty in imputed estimates resulting in spuriously narrow confidence intervals [4, 5]. Multiple imputation incorporates missing data uncertainty by replacing missing values with a set of plausible values [6]. Imputation techniques are not applicable in systematic reviews unless individual participant data is available.

Addressing missing participant data for continuous outcomes in systematic reviews provides additional challenges to those of individual studies: individual patient data is usually unavailable, there are a wide range of possible imputed values [2], and any approach should ideally address the measure of effect (such as the mean, mean difference or standardized mean difference) and the associated measure of precision (such as the standard deviation [SD] or standard error).

Currently, no methods have been proposed for investigating the extent to which missing participant data for continuous outcomes may bias the results of systematic reviews. Other than rating down for loss to follow-up in the Cochrane Risk of Bias Tool [7], the Cochrane Collaboration handbook provides no guidance for systematic reviewers to decide on the degree of concern warranted by missing participant data. To address this problem, we propose an approach for addressing the impact of missing participant data for continuous outcomes in systematic reviews.

CHAPTER 17

METHODS

17.1 Development of approach

We formed a group of nine members consisting of clinical epidemiologists, methodologists, and biostatisticians, with extensive experience in systematic reviews of continuous outcomes. Five members of the group had participated in a study of how to address missing participant data for dichotomous outcomes and its potential impact on the estimates of the effect of treatment in RCTs [1]. We reviewed the available literature on the topic including the Cochrane Handbook [1, 2, 7-9], and then conducted a consultative, iterative process to develop our approach.

We defined missing participant data as data unavailable to the investigator(s) or available to the investigator(s) but not included in published reports. For imputing means and SDs for participants with missing data, we considered a number of possible sources of data. All these sources reflect real observed outcomes in participants followed-up (i.e., those with available outcome data) in individual trials included in the systematic review. We developed strategies to combine imputations for participants with missing data in the intervention and control arms with available results from patients with complete data. Our assumption for this approach is that reasons for missing data, and the participants with missing data, were similar across

studies. It then follows that systematic review authors should determine, as far as possible, reasons for missing data, and characteristics of populations with missing data (versus those with available data) in the intervention and control arms of the primary studies. Our goal was to suggest a range of plausible strategies that would be progressively more stringent in challenging the robustness of the pooled estimates of the intervention effect. In this article, we present our proposed approach, which we apply to two example meta-analyses [10, 11].

17.2 Imputing measure of effect

We selected five sources of data reflecting real observed mean scores in participants followed-up (i.e., with available outcome data) in individual trials included in a meta-analysis:

[A] The best mean score among the intervention arms of included trials.

[B] The best mean score among the control arms of included trials.

[C] The mean score from the control arm of the same trial.

[D] The worst mean score among the intervention arms of included trials.

[E] The worst mean score among the control arms of included trials.

The worst mean score represents the poorest outcome and the best mean score represents the best outcome. These could be higher or lower scores depending on the outcome instrument used.

Using the five sources of data, we developed four progressively more stringent imputation strategies for participants with missing data in the intervention and control arms. Table 17.1 provides a matrix describing the following four strategies:

- Strategy 1 uses source C for those with missing data in both the intervention and control arm.
- Strategy 2 uses source D for those with missing data in the intervention arm, and source B for those with missing data in the control arm.
- Strategy 3 uses source E for those with missing data in the intervention arm, and source B for those with missing data in the control arm.
- Strategy 4 uses source E for those with missing data in the intervention arm, and source A for those with missing data in the control arm.

17.3 Imputing measure of precision

We considered three sources of SDs, in combination with our strategies for imputing means, for both the intervention and control arms: the smallest SD (likely most

Table 17.1: Matrix of assumptions for participants with missing data for continuous outcomes in intervention and control arms

		Assumptions about the means of participants in intervention arm					
		Source A	Source B	Source C	Source D	Source E	
Assumptions about the means of participants in control arm		Source A					Strategy 4
		Source B				Strategy 2	Strategy 3
		Source C			Strategy 1		
		Source D					
		Source E					

Source A – Best mean among intervention arms of included trials; Source B – Best mean among the control arms of included trials; Source C - Mean score from the control arm of the same trial; Source D – Worst mean among the intervention arms of included trials; Source E – Worst mean among control arms of included trials

stringent), the median SD, and the largest SD (likely the least stringent) among the control arms of all included trials. We tested these different sources and found only small differences in the impact on the pooled estimate and its confidence interval.

Given the median SD is the most plausible, we propose its use among control arms of all included trials.

17.4 Combining observed and imputed data

We propose a three-step method for each imputation strategy. In the first step, for each arm in each trial, we combined the observed means and SDs of the participants with available data with the imputed means and SDs for participants with missing data using the following formulas:

$$(1) M_{XTi} = \frac{(M_{FTi} \times n_{FTi}) + (M_{LTi} \times n_{LTi})}{n_{FTi} + n_{LTi}}$$

$$(2) M_{XCi} = \frac{(M_{FCi} \times n_{FCi}) + (M_{LCi} \times n_{LCi})}{n_{FCi} + n_{LCi}}$$

$$(3) SD_{XTi} = \sqrt{\frac{(n_{FTi} - 1)SD_{FTi}^2 + (n_{LTi} - 1)SD_{LTi}^2}{n_{FTi} + n_{LTi} - 2}}$$

$$(4) SD_{XCi} = \sqrt{\frac{(n_{FCi} - 1)SD_{FCi}^2 + (n_{LCi} - 1)SD_{LCi}^2}{n_{FCi} + n_{LCi} - 2}}$$

$$(5) n_{XTi} = n_{FTi} + n_{LTi}$$

$$(6) n_{XCi} = n_{FCi} + n_{LCi}$$

where ‘M’ represents the mean, ‘SD’ the standard deviation, ‘n’ the sample size, ‘X’ the combined estimates, ‘F’ the followed-up group, ‘L’ the imputed lost to follow-up group, ‘T’ the treatment group, ‘C’ the control group and ‘i’ the trial.

In the second step, we used the combined mean and standard deviation estimates from each arm to calculate a treatment effect (mean difference) for each study. In the third step, we performed a standard random-effects meta-analysis, where we assumed that the studies included are a random sample of a population of studies [12], to pool the newly calculated mean difference of all included studies.

We used Review Manager Version 5.1.6 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011) and Microsoft Excel 2011 to complete our analyses.

CHAPTER 18

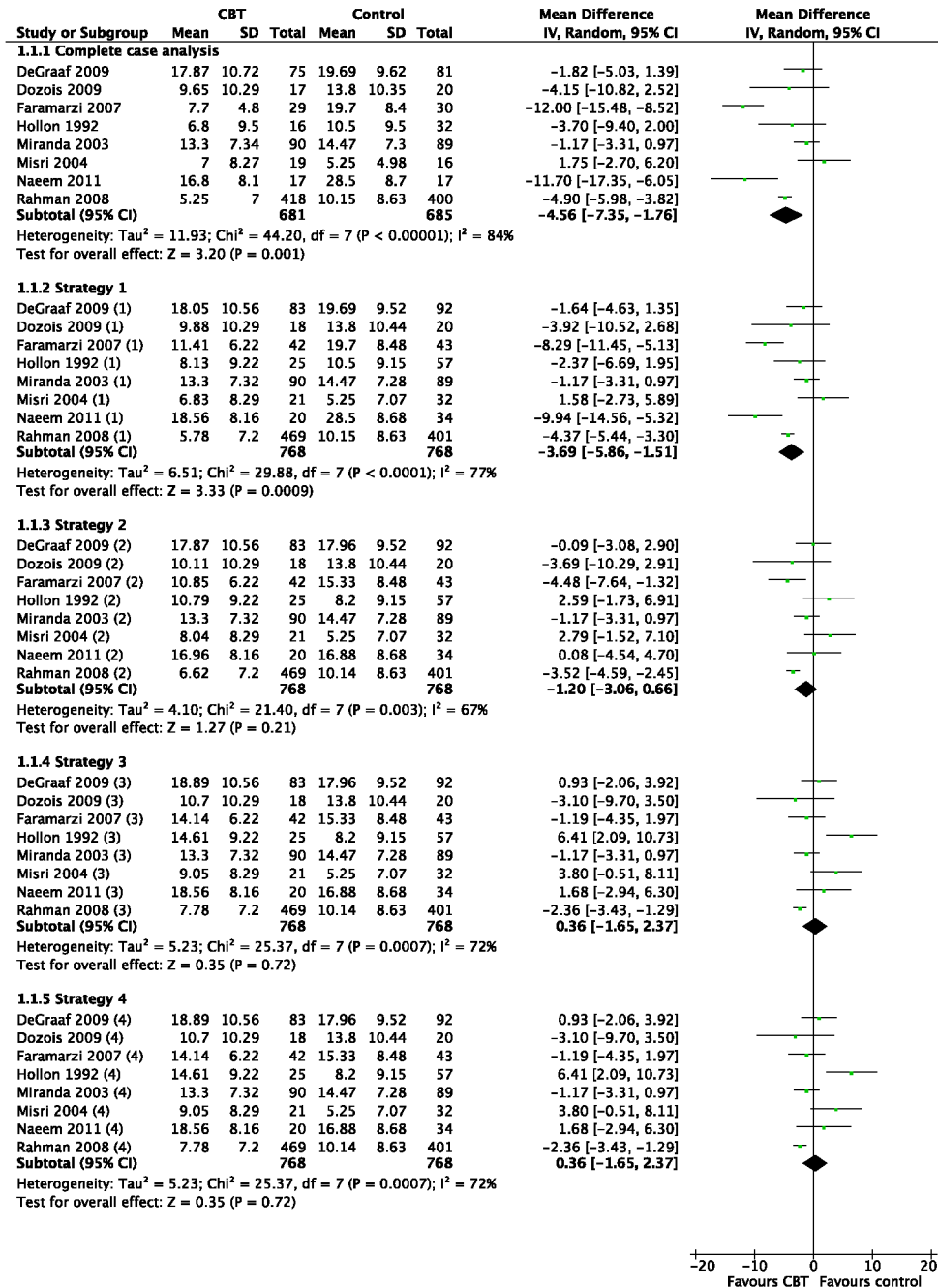
APPLICATION OF APPROACH

We applied our approach to two systematic reviews.

18.1 CBT for depression in patients receiving disability benefits

The first systematic review evaluated cognitive behavioural therapy (CBT) versus minimal to no treatment or usual care for depression in patients receiving disability benefits. Depressive symptoms were measured using the Beck Depression Inventory-II, in which higher scores represented a worse outcome and the minimal important difference is five units (Figure 18.1). The systematic review included eight randomized trials and had a median missing participant data rate of 21% (range 0 to 41%) [10]. The reasons for missing data were provided in 3 of 8 studies [13-15]. The reasons included personal circumstances, medical illness, too time consuming, non-compliance, lack of treatment response, moved residence, refusal and absent. We did not find important differences in reasons for missing data across study arms or studies, and no studies reported any significant differences in the characteristics of those followed versus those not followed. The complete case analysis showed a mean difference of -4.56 (95% confidence interval [CI] of -7.35 to -1.76). Strategy 1 resulted in some loss of effect but maintained statistical significance. Strategy 2 resulted in further loss of effect and a loss of statistical significance. Strategies 3 and 4 resulted in

Figure 18.1: Forest plots of the complete case analysis and sensitivity analyses for handling participants with missing data for continuous outcomes in a systematic review evaluating CBT for depression in patients receiving disability benefits

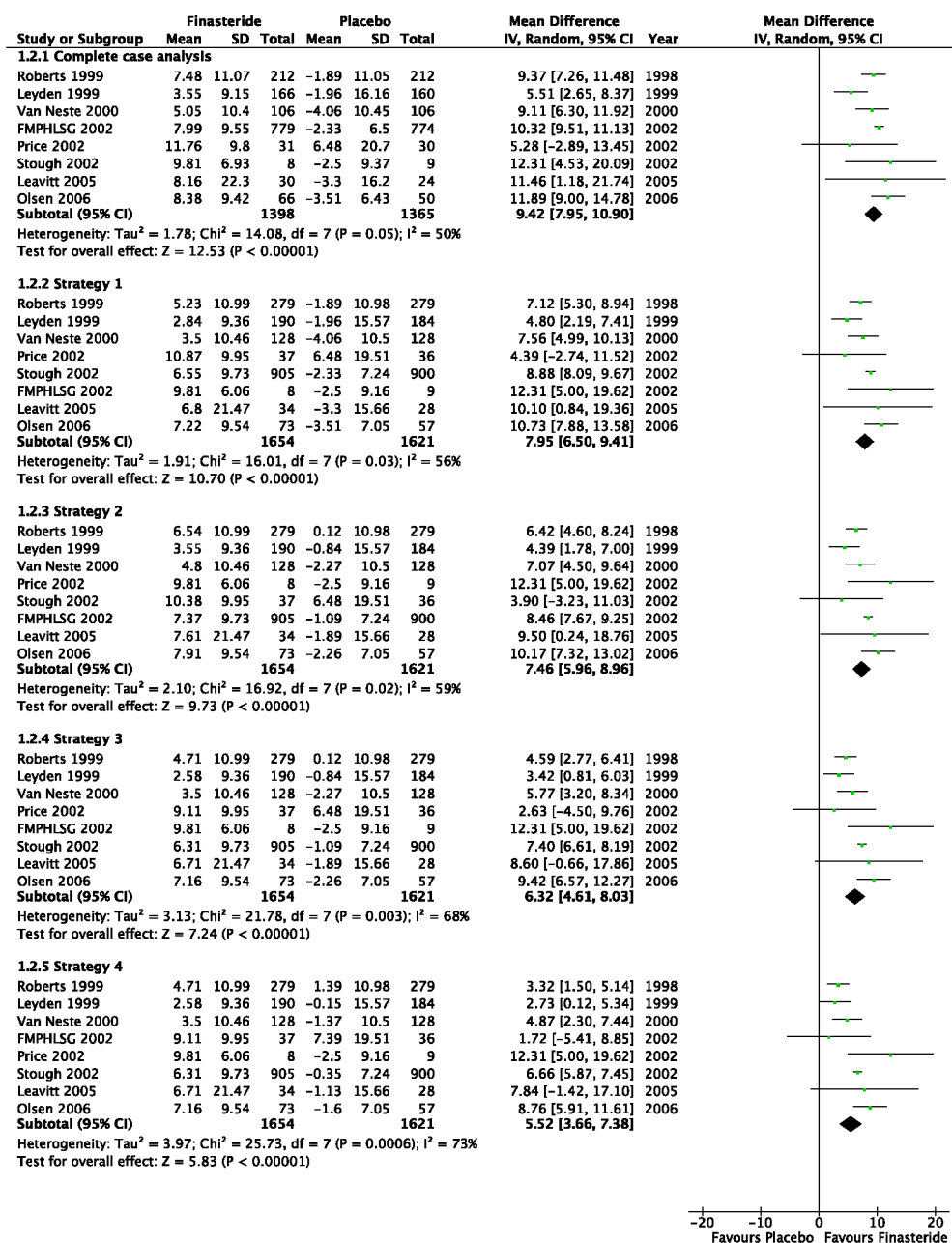


even further loss of effect and much larger p-values (Figure 18.1).

18.2 Finasteride for improvement in scalp hair for men with androgenetic alopecia

The second systematic review evaluated the effectiveness of finasteride therapy versus placebo on improvement in scalp hair for men with androgenetic alopecia. The review included eight randomized trials and had a median missing participant data rate of 14% (range 0% to 24%) [11]. The reasons for missing data were provided in 5 of 8 studies [16-20]. The reasons included lack of efficacy, relocated residence, adverse events, protocol violations, withdrawal of consent, and other. Again, there were no important differences across studies, or in the characteristics of those followed versus those not followed. The complete case analysis showed a mean hair count increase of 9.42% (95% CI of 7.95% to 10.90%), where greater hair count percentage represented a better outcome (Figure 18.2). All four strategies resulted in some loss of effect (in descending order of approach) but maintained statistical significance at $p < 0.001$ (Figure 18.2).

Figure 18.2: Forest plots of the complete case analysis and sensitivity analyses for handling participants with missing data for continuous outcomes in a systematic review evaluating finasteride therapy for androgenetic alopecia



CHAPTER 19

DISCUSSION

19.1 Summary of findings

We developed four increasingly stringent imputation strategies to take into account missing participant data from continuous outcomes in meta-analyses and assessed the associated risk of bias. Our approach demonstrated varying impact on effects in the example systematic reviews. In the first review, effect estimates were diminished and lost significance as the strategies became more stringent. In the second review, effect estimates maintained statistical significance even using the most stringent strategy.

19.2 Implications on risk of bias

What are the implications of the results of our two examples on risk of bias associated with missing data? In the first example (Figure 18.1), the results withstand only the least stringent assumptions regarding missing data. This suggests that the results are vulnerable to risk of bias, and applying the GRADE/Cochrane handbook criteria for confidence in estimates of effect (quality of evidence) [21, 22], one would rate down for risk of bias, even if the studies did not have other risk of bias issues (e.g. concealment, blinding).

In the second example (Figure 18.2), the effect is maintained (though diminished) with even the most stringent assumption, and the confidence interval remains relatively narrow with a low p-value (<0.001). The appropriate conclusion may be that the results are robust with respect to missing data. One might, however, want to consider the boundaries of what one could consider an important increase in percentage of hair count. For instance, if a guideline panel considered an increase of less than 10% unimportant, the panel might conclude that the missing data threatens the inference of an important treatment effect. Were that so, the panel might consider rating down their confidence in effect on the basis of possible risk of bias resulting from the missing data (i.e., a possibly important effect becomes unimportant under stringent assumptions regarding missing data).

19.3 Properties

Our examples appear to illustrate three properties, which likely explain the differences in results between the two examples. The first is by the size of the effect and its precision; small effects with borderline significance are more likely to lose significance than larger, precise effects. The second is the proportion of missing participant data and the distribution between the treatment and control groups; the greater the percentage of participants with missing data, the greater the risk of losing significance. The third is that the more extreme the scores of imputation in the proposed strategies, the more likely the newly calculated effects will deviate from the

complete case analysis results. Although these properties follow logically from first principles, further theoretical or empirical exploration would be desirable.

19.4 Strengths, limitations and other issues

The strengths of our approach include the use of real observed data from individual trials and meta-analyses, varying degrees of plausible scenarios to test the robustness of pooled estimates, and the application to completed systematic reviews with different magnitudes of effects and missing participant data rates. Our approach also addresses the implications of findings to inferences regarding risk of bias. The approach is compatible with the GRADE/Cochrane handbook guidance for addressing confidence in estimates [21, 22]. Another strength is that the approach we offer is the first systematic, quantitative approach addressing a problem in which guidance, up to now, has been very limited.

We applied our approach to two particular examples; testing on a larger number would be desirable, and might uncover problems or difficulties with our approach. Other limitations of our approach, relate to some inevitable arbitrariness in the choices we have made. We suggest using the best and worst means of the control and intervention arms from all included trials for imputing means for participants with missing data. One could make the argument that using the best or worst means may

not be plausible, particularly if these results come from very small trials. On the other hand, we are using real data that has actually been observed in clinical trials.

There is also some arbitrariness in the particular four strategies we made from the combinations of sources of mean values. At the extreme, one could test each of the potential 25 combinations. Further, depending on the reasons for missing data, some assumptions may be more reasonable than others. Indeed, we suggest that if studies provide plausible reasons for missing data, investigators consider these in deciding on plausible imputation strategies. We anticipate, however, that for an approach to be widely adopted, it requires relative simplicity. If mechanisms underlying missingness are provided and are idiosyncratic, or there are important differences between those with available data and those with missing data that vary across studies, then confidence in the estimates derived from our approach decreases, and more sophisticated imputation strategies may be necessary. In our two example systematic reviews, 8 of 16 studies reported reasons for missing data, no study reported idiosyncratic or unique reasons for loss to follow-up that would suggest this necessity, and no study reported important prognostic differences between those followed and those not followed. The reasons for missingness reported were typical of what is generally seen in RCTs. We suspect that the results in our two examples, which do not provide compelling reasons for modifying our approach as a result of different

reasons for missingness or different prognostic characteristics of those missing across studies, will be typical of what systematic review authors will face.

Regarding the choice of SD one imputes for participants with missing data, reasonable options might include the median SD (our choice) or the smallest and largest SDs, or sensitivity analyses testing all three (which would require 12 sensitivity analyses simultaneously varying mean values and SDs). The median SD is the most plausible and offers desirable simplicity. Moreover, in the two examples to which we have applied our approach, the different SD options made very little difference to our results.

Systematic review authors applying our approach will face challenges when they confront trials that fail to adequately report missing participant data or do not report them at all. This failure can be classified into the following categories: 1) trialists do not report the frequency of missing participant data; 2) trialists report total missing participant data but not by trial arm; 3) trialists report missing participant data to which they have applied imputation techniques, and report only the analysis based on the imputed values.

For any trial in which authors do not report missing participant data rates, we suggest using the median missing participant data rate from the systematic review. Some may feel this assumption is too stringent. We therefore propose performing a sensitivity analysis using a missing participant data rate of zero in both arms. If authors reported total missing participant data only, we suggest assuming that missing data was equally distributed in both arms.

If the individual trial handled missing participant data and reported imputed analyses only, we suggest using the imputed results for the meta-analysis. If the authors reported both the imputed analysis and the complete case analysis, we suggest applying our approach to the trial's complete case analysis. All of these suggestions assume that systematic review authors have attempted, and failed, to obtain the missing information directly from trialists.

A final consideration is that our approach is thus far restricted to systematic reviews in which each study has used the same measurement instrument producing continuous outcome data. Systematic review authors often face a group of trials that have used different instruments, to measure a similar construct. Should others find our approach compelling, development of methods to extend the approach to systematic reviews pooling trials using multiple instruments would be important.

19.5 Conclusions

In summary, detailed guidance on how to determine the extent to which missing data in primary studies increases risk of bias in systematic reviews of continuous outcomes has thus far been unavailable. We suggest an approach that involves an initial complete case analysis with subsequent sensitivity analyses making progressively more stringent assumptions about results in patients with missing data; we have provided detailed guidance for conducting such analyses. To the extent that results change with the sensitivity analyses, risk of bias as a result of missing data increases. The explicit guidance we have provided represents a potentially important step in addressing missing continuous outcome data in systematic reviews.

ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS

Acknowledgements

We thank Dr. Steven Hanna for his intellectual input.

Funding

There were no sources of funding for this study. SE is supported by a Canadian Institutes of Health Research (CIHR) Doctoral Award. PA-C is funded by a Miguel Servet investigator contract from the Instituto de Salud Carlos III (CP09/00137).

Authors' contributions

SE and GHG conceived the study. SE, EAA, RAM, XS, SDW, DHA, PAO, BCJ and GHG contributed to the study design and developed the approach for handling missing participant data. SE applied the approach to systematic reviews. SE completed the first draft of the manuscript. SE is the first author and GG is the senior author of the study. All authors read, critically revised and approved the final manuscript to be published.

Authors' information

SE is a doctoral candidate in the Department of Clinical Epidemiology and Biostatistics at McMaster University. EAA is an Associate Professor in the Department of Internal Medicine at the American University of Beirut, and a part-time Associate Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. RAM is a doctoral candidate in the Department of Clinical Epidemiology and Biostatistics at McMaster University and a nephrologist. XS is an investigator in the The Centre for Health Research at Kaiser Permanente Northwest and a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. SDW is a Professor Emeritus, in the Department of Clinical Epidemiology & Biostatistics at McMaster University. DHA is a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. PA-C is a researcher in the Iberoamerican Cochrane Centre at the CIBERESP-IIB Sant Pau in Barcelona, Spain. BCJ is an epidemiologist in the Department of Anaesthesia & Pain Medicine at The Hospital for Sick Children, an Associate Scientist in the Child Health Evaluative Sciences at the SickKids Research Institute, and a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. GHG is a Distinguished Professor in the Departments of Medicine and Clinical Epidemiology and Biostatistics at McMaster University.

REFERENCES

1. Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, Mulla S, Lamontagne F, Bassler D, Vera C, Alshurafa M, Katsios CM, Zhou Q, Cukierman-Yaffe T, Gangji A, Mills EJ, Walter SD, Cook DJ, Schunemann HJ, Altman DG, Guyatt GH: **LOST to follow-up Information in Trials (LOST-IT): Potential impact on estimated treatment effects.** *BMJ* 2012, **344**:e2809.
2. **Further issues in meta-analysis: Intention to treat issues**
[<http://www.cochrane-net.org/openlearning/html/mod14-4.htm>]
3. Akl EA, Johnston BC, Alonso-Coello A, Neumann I, Ebrahim S, Briel M, Cook DJ, Guyatt GH: **Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers [Under review at PLoS ONE].** In.; 2012.
4. Molnar FJ, Hutton B, Fergusson D: **Does analysis using "last observation carried forward" introduce bias in dementia research?** *CMAJ* 2008, **179**(8):751-753.
5. Ware JH: **Interpreting incomplete data in studies of diet and weight loss.** *N Engl J Med* 2003, **348**(21):2136-2137.
6. Rubin DB: **Multiple imputation for nonresponse in surveys.** New York: John Wiley & Sons, Inc.; 1987.

7. The Cochrane Collaboration: **Chapter 8: Risk of Bias**. In: *Cochrane Handbook for Systematic Reviews of Interventions (Version 510)*. edn. Edited by Higgins JPT, Green S. Oxford, UK: Cochrane Collaboration; 2011.
8. Groenwold RHH, Donders RT, Roes KCB, Harrell FE Jr, Moons KGM: **Dealing With Missing Outcome Data in Randomized Trials and Observational Studies**. *American Journal of Epidemiology* 2011, **175**(3):210-217.
9. Higgins JPT, White IR, Wood AM: **Imputation methods for missing outcome data in meta-analysis of clinical trials**. *Clinical Trials* 2008, **5**:225-239.
10. Ebrahim S, Montoya L, Truong W, Hsu S, Kamal el Din M, Carrasco-Labra A, Busse JW, Walter S, Heels-Ansdell D, Couban R, Patelis-Siotis I, Bellman M, de Graaf LE, Dozois D, Bieling P, Guyatt GH: **Effectiveness of Cognitive Behavioral Therapy for depression in patients receiving disability benefits: A systematic review and individual patient meta-analysis**. *PLoS ONE* 2012, **In press**.
11. Mella JM, Perret MC, Manzotti M, Catalano HN, Guyatt GH: **Efficacy and safety of finasteride therapy for androgenetic alopecia: a systematic review**. *Arch Dermatol* 2010, **146**(10):1141-1150.
12. Higgins J, Thompson S, Deeks J, Altman D: **Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice**. *J Policy Health Serv Res Policy* 2002, **7**(1):51-61.
13. de Graaf LE, Gerhards SAH, Arntz A, Riper H, Metsemakers JFM, Evers SMAA, Severens JL, Widdershoven G, Huibers MJH: **Clinical effectiveness of online**

- computerised cognitive behavioural therapy without support for depression in primary care: randomised trial.** *The British Journal of Psychiatry* 2009, **195**:73-80.
14. Misri S, Reebye P, Corral M, Milis L: **The Use of Paroxetine and Cognitive-Behavioral Therapy in Postpartum Depression and Anxiety: A Randomized Controlled Trial.** *Journal of Clinical Psychiatry* 2004, **65**:1236-1241.
15. Rahman A, Malik A, Sikandar S, Roberts C, Creed F: **Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial.** *The Lancet* 2008, **372**:902-909.
16. Finasteride Male Pattern Hair Loss Study Group: **Long-term (5-year) multinational experience with finasteride 1 mg in the treatment of men with androgenetic alopecia.** *Eur J Dermatol* 2002, **12**(1):38-49.
17. Leavitt M, Perez-Meza D, Rao NA, Barusco M, Kaufman KD, Ziering C: **Effects of finasteride (1 mg) on hair transplant.** *Dermatol Surg* 2005, **31**(10):1268-1276.
18. Leyden J, Dunlap F, Miller B, et al: **Finasteride in the treatment of men with frontal male pattern hair loss.** *J Am Acad Dermatol* 1999, **40**(6 pt 1):930-937.
19. Olsen EA, Hordinsky M, Whiting D, et al. Dutasteride Alopecia Research Team: **The importance of dual 5-reductase inhibition in the treatment of male pattern hair loss: results of a randomized placebo-controlled study of**

- dutasteride versus finasteride. *J Am Acad Dermatol*. *J Am Acad Dermatol* 2006, **55**(6):1014-1023.**
20. Roberts JL, Fiedler V, Imperato-McGinley J, et al: **Clinical dose ranging studies with finasteride, a type 2 5alpha-reductase inhibitor, in men with male pattern hair loss.** *J Am Acad Dermatol* 1999, **41**(4):555-563.
21. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group: **GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.** *BMJ* 2008, **336**(7650):924-926.
22. The Cochrane Collaboration: **Chapter 12: Interpreting results and drawing conclusions.** In: *Cochrane Handbook for Systematic Reviews of Interventions (Version 510)*. edn. Edited by Higgins JPT, Green S. Oxford, UK: Cochrane Collaboration; 2011.

STUDY 4

Addressing missing participant data in systematic reviews: continuous outcomes assessed with different instruments

Shanil Ebrahim¹, Bradley C. Johnston^{1,2,3}, Elie A. Akl^{1,4}, Reem A. Mustafa^{1,5}, Xin Sun^{1,6}, Stephen D. Walter^{1,7}, Diane Heels-Ansdell¹, Pablo Alonso-Coello⁸, Gordon H. Guyatt^{1,9}

¹Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada

²Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, Canada

³Child Health Evaluative Sciences, SickKids Research Institute, Toronto, Canada

⁴Department of Internal Medicine, American University of Beirut, Beirut, Lebanon

⁵Department of Medicine, State University of New York at Buffalo, Buffalo, USA

⁶Department of Medicine, University of Missouri-Kansas City, Kansas City, USA

⁷Center for Clinical Epidemiology and Evidence-based Medicine, Xinqiao Hospital, Chongqing, China

⁸Iberoamerican Cochrane Centre, CIBERESP-IIB Sant Pau, Barcelona, Spain

⁹Department of Medicine, McMaster University, Hamilton, Canada

§Corresponding author

Corresponding Author

Shanil Ebrahim, Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main Street West, HHS Rm. 2C, Hamilton, Ontario, Canada, L8S 4K1.
Email: Shanil.ebrahim@utoronto.ca

Keywords

Missing participant data, continuous outcomes, risk of bias, systematic reviews, meta-analysis, lost to follow-up

Submitted to Journal of Clinical Epidemiology

CHAPTER 20

ABSTRACT

Background: We previously developed an approach to address the impact of missing participant data in meta-analyses of continuous variables in trials that used the same measurement instrument. We extend this approach to meta-analyses including trials that use different instruments to measure the same construct.

Methods: We reviewed the available literature, conducted an iterative consultative process, and developed an approach involving conducting a complete case analysis complemented by sensitivity analyses that apply a series of increasingly stringent assumptions about results in patients with missing continuous outcome data.

Results: Our approach involves choosing the reference measurement instrument; converting scores from different instruments to the units of the reference instrument; developing four successively more stringent imputation strategies for addressing missing participant data; calculating a pooled mean difference for the complete case analysis and imputation strategies; calculating the proportion of patients who experienced an important treatment effect; and judging the impact of estimates on quality of evidence due to missing participant data. We applied our approach to an example systematic review of respiratory rehabilitation for chronic obstructive pulmonary disease.

Conclusions: Our extended approach provides quantitative guidance for addressing missing participant data in systematic reviews of trials using different instruments to measure the same construct.

CHAPTER 21

INTRODUCTION

Randomized controlled trials (RCTs) often suffer from missing participant data [1]. Missing participant data increases risk of bias in both individual trials and meta-analyses. This is especially concerning in positive trials (i.e., those with a significant treatment effect) if, in the intervention group, the outcomes of participants with missing data are worse than the outcomes of those with available data.

The Cochrane Collaboration Handbook has proposed a strategy for handling missing participant data for dichotomous outcomes in systematic reviews. The strategy suggests conducting a complete (available) case analysis complemented by sensitivity analyses of various assumptions regarding outcomes of participants with missing data [2]. One of the most common approaches is to adopt the “worst case scenario”, in which one assumes that participants with missing data in the intervention group had the worst outcome, and those in the control group had the best possible outcome [2]. Although this assumption tests the robustness of the pooled estimates for complete case analyses, it is typically implausible [1].

Our research group has proposed additional strategies that use a range of plausible assumptions that increasingly challenges the results [1]. These strategies make the assumption that those in the intervention group with missing data do relatively worse than those with available data, and those in the control group with missing data do relatively better than those with available data. We have applied these strategies to RCTs [1] and to systematic reviews [3] reporting dichotomous outcomes.

Until recently, no methods were available (including the Cochrane Handbook) for addressing missing participant data for continuous outcomes in systematic reviews or for assessing its impact on quality of evidence. We have addressed this gap by extending our work in dichotomous outcomes [3], and proposing an approach involving conducting a complete case analysis complemented by sensitivity analyses that apply a series of increasingly stringent assumptions about results in patients with missing continuous outcome data [4]. This approach is limited to systematic reviews in which all trials used the same measurement instrument. In this paper, we extend the approach to systematic reviews pooling trials using different instruments to measure the same construct (e.g., dyspnea, fatigue, emotional function). To further enhance interpretability, we illustrate how to calculate important treatment effects from the derived estimates using the minimally important difference (MID), the smallest difference that patients perceive as important [5,6] and, in our discussion, show how the results can be interpreted in the context of a clinical practice guideline.

CHAPTER 22

METHODS

In developing our initial approach, we reviewed the available literature (including the Cochrane Handbook) on missing participant data [1,2,7,8,9], and conducted an iterative consultative process involving the nine authors of this paper, including clinical epidemiologists, methodologists, and biostatisticians.

Our approach addresses analyses for aggregate trial-level data for conducting a meta-analysis, and not analyses of individual participant data meta-analyses or missing study-level characteristics (for subgroup or meta-regression analyses). Key assumptions of our approach are that reasons for missing data and characteristics of patients with missing data are similar across studies.

We defined missing participant data as data unavailable to systematic review authors even after contacting the trial investigator(s), or unreported data of those who were non-adherent to the trial protocol (participants who received the incorrect intervention, a disallowed co-intervention, or were not compliant with the protocol) [10].

Our approach also applies strategies to enhance interpretability of pooled estimates using the minimally important difference (MID), the smallest difference that patients perceive as either an important benefit or harm [5,6,11]. Clinicians and researchers may incorrectly treat the MID as a cutoff threshold of importance for the mean difference: if the pooled mean difference is less than the MID, the treatment is considered to have an unimportant effect. This interpretation assumes that all patients gain the same benefit from the treatment. Contrary to this assumption, even if the pooled mean difference is less than the MID, a substantial proportion of treated patients may experience an important benefit. In our approach, we calculate the proportion of patients who have an important treatment effect based on the MID.

CHAPTER 23

RESULTS

Our proposed approach consists of the following steps:

- Choosing the reference measurement instrument
- Converting scores from different instruments to the units of the reference instrument
- Imputing measures of effect and their precision
- Combining observed and imputed data
- Calculating the proportion of patients who have an important treatment effect
- Judging the impact of missing participant data on quality of evidence

23.1 Choosing the reference measurement instrument

The reference measurement instrument is typically the one that is most familiar to the target audience (largely based on most commonly used) and/or has the best measurement properties. Our criteria for prioritizing the instrument of choice for presentation of data for this approach include, in possible order of importance, include: an anchor-based MID estimate (i.e., MID derived from methods that examine the relationship between the measure of interest and an independent measure, or anchor, to help interpret the meaning of a particular degree of change) [12], excellent

measurement properties (in particular, responsiveness) [13], and familiarity to the target audience.

23.2 *Converting scores from different instruments to the reference instrument*

In order to best handle continuous data measuring the same construct using different instruments, we convert all results into the units of the reference instrument. When the MID is available, this approach will facilitate interpretability of results to relevant knowledge users. We will begin by a generalization in which two instruments are used to generate continuous data measuring the same construct in which instrument *A* represents the reference instrument and instrument *B* represents the alternative instrument. We used the following approach:

For each trial, we first convert the means and standard deviations of the change scores from instrument B to the units of instrument A [11]. For this calculation, we use the following formulas to convert the mean response (*M*) and standard deviation (*SD*) into the scale of instrument A: $M_{Ai} = (M_{Bi} - L_{Bi}) * (R_{Ai} \div R_{Bi}) + L_{Ai}$ and $SD_{Ai} = SD_{Bi} * (R_{Ai} \div R_{Bi})$, where L_A and L_B represent the worst possible outcome score of instrument A and B respectively, R_A and R_B the ranges (highest possible outcome score minus lowest possible outcome score) for instruments A and B respectively and *i* the trial. We apply these formulas separately for the intervention group and the control

group of each trial (example provided in *Appendix*). The remainder of our approach will use the converted scores from the reference instrument.

23.3 *Imputing measure of effect and its precision*

For imputing means and SDs for participants with missing data, we consider data from five sources that reflect observed outcomes in participants followed-up in the individual trials included in the meta-analysis [4]:

- [A] The best mean score among the intervention arms of included trials.
- [B] The best mean score among the control arms of included trials.
- [C] The mean score from the control arm of the same trial.
- [D] The worst mean score among the intervention arms of included trials.
- [E] The worst mean score among the control arms of included trials.

The worst mean score represents the poorest outcome and the best mean score represents the best outcome. These could be higher or lower scores depending on the instrument used.

In addition to the complete case analysis, we offer four progressively more stringent imputation strategies for participants with missing data that increasingly challenge

the robustness of the pooled estimates of the intervention effect (Table 23.1) [4]. Strategy 1 imputes source C for those with missing data in both the intervention and control arm. Strategy 2 imputes source D for those with missing data in the intervention arm, and source B for those with missing data in the control arm. Strategy 3 imputes source E for those with missing data in the intervention arm, and source B for those with missing data in the control arm. Strategy 4, the most stringent assumption, imputes source E for those with missing data in the intervention arm, and source A for those with missing data in the control arm. We impute the median SD from the control arms of all included trials for each of our strategies, given it is the most plausible assumption, and a previous test between the smallest, median and largest SD showed negligible differences between the three [4].

23.4 Combining observed and imputed continuous data

To combine the observed and imputed data, we suggest a three-step method for each imputation strategy [4]. In the first step, for each arm in each trial, we combine the observed means and SDs of the participants with available data with the imputed means and SDs for participants with missing data. In the second step, we use the combined estimates from each arm to calculate a treatment effect (mean difference) for each study. In the third step, we perform a standard random-effects meta-analysis to pool the newly calculated mean difference of all included studies in the systematic review.

Table 23.1. Matrix of assumptions for participants with missing data for continuous outcomes in intervention and control arms

		Assumptions about the means of participants in intervention arm		
		<i>C. Mean score from the control arm of the same trial</i>	<i>D. Worst mean among the intervention arms</i>	<i>E. Worst mean among control arms of included trials</i>
Assumptions about the means of participants in control arm	<i>A. Best mean among intervention arms</i>			Strategy 4
	<i>B. Best mean among the control arms</i>		Strategy 2	Strategy 3
	<i>C. Mean score from the control arm of the same trial</i>	Strategy 1		

23.5 Enhancing interpretability using MID

If the MID is available, we calculate, for the complete case analysis and each imputation strategy, the proportion of patients who benefit from the treatment. For the control group and intervention group in each trial, we first calculate the probability (ρ_{Ci} and ρ_{Ei} respectively) of obtaining a treatment effect greater than or equal to the MID: $p_{Ci} = 1 - \Phi\left(\frac{MID - m_{Ci}}{sd_{Ci}}\right)$ and $p_{Ei} = 1 - \Phi\left(\frac{MID - m_{Ei}}{sd_{Ei}}\right)$, where Φ denotes the standard normal cumulative distribution function, *MID* the minimally important difference for the reference instrument, m_{Ci} and m_{Ei} the control and intervention mean, and sd_{Ci} and sd_{Ei} the control and intervention standard deviation for the given

trial i [11]. For these calculations, we use the means and SDs derived from step 1 in the section *combining observed and imputed continuous data*.

We derive the risk difference (RD) for each trial using $RD_i = p_{Ei} - p_{Ci}$, and the

corresponding standard errors (SE) using $SE(RD_i) = \sqrt{\frac{(p_{Ei} * (1 - p_{Ei}))}{n_{Ei}} + \frac{(p_{Ci} * (1 - p_{Ci}))}{n_{Ci}}}$,

where n_{Ei} and n_{Ci} represent the intervention and control sample size for the given trial i (including those with missing data) [11]. We perform a standard random-effects meta-analysis using the inverse variance method to pool the calculated trial RDs.

23.6 *Judging impact of missing participant data on quality of evidence*

The GRADE framework for judging our confidence in the pooled estimates (quality of evidence) has four levels - high, moderate, low, or very low confidence [14]. One reason we may lose confidence is due to high risk of bias as a result of missing participant data. To judge whether to rate down confidence in estimates as a result of missing participant data, systematic reviewers would be required to make two assessments.

The first assessment is to consider and apply the most plausible imputation strategies for the systematic review under consideration. For some systematic reviews, authors

may consider the most stringent assumption to be plausible; for others, they may not. If the most stringent strategy were considered implausible, systematic review authors would not conduct the sensitivity analysis corresponding to that strategy.

The second assessment aims to enhance interpretability by considering the proportion of patients who gain an important benefit from the intervention. This proportion is calculated using the MID threshold. As the threshold involves a value and preference judgment, it is typically in the domain of guideline developers. Systematic review authors need to tailor their approach to a wide range of users who are likely to have different values and preferences. Consequently, they should test a reasonable range of thresholds that guideline panels may adopt as an important effect.

A guideline panel, when considering a clinical recommendation, needs to make a specific judgement of values and preferences. A direct translation of this judgment is the choice of a threshold they adopt as an important effect for their intended population. After the guideline panel has adopted a threshold, if - using a particular imputation approach - the entire confidence interval were above the threshold (i.e., indicating an effect greater than the threshold), a guideline panel would not rate down their confidence in estimates due to missing participant data. If, however, the confidence interval associated with that imputation strategy includes the threshold, the guideline panel would lose confidence that the treatment effect is large enough to

warrant its use [15]. The reason for rating down confidence would be on the basis of risk of bias due to missing participant data.

23.7 Application of approach

We applied our approach to a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease (COPD) [16]. The systematic review included 31 RCTs, and had a median missing participant data rate of 17% (range 0% to 38%).

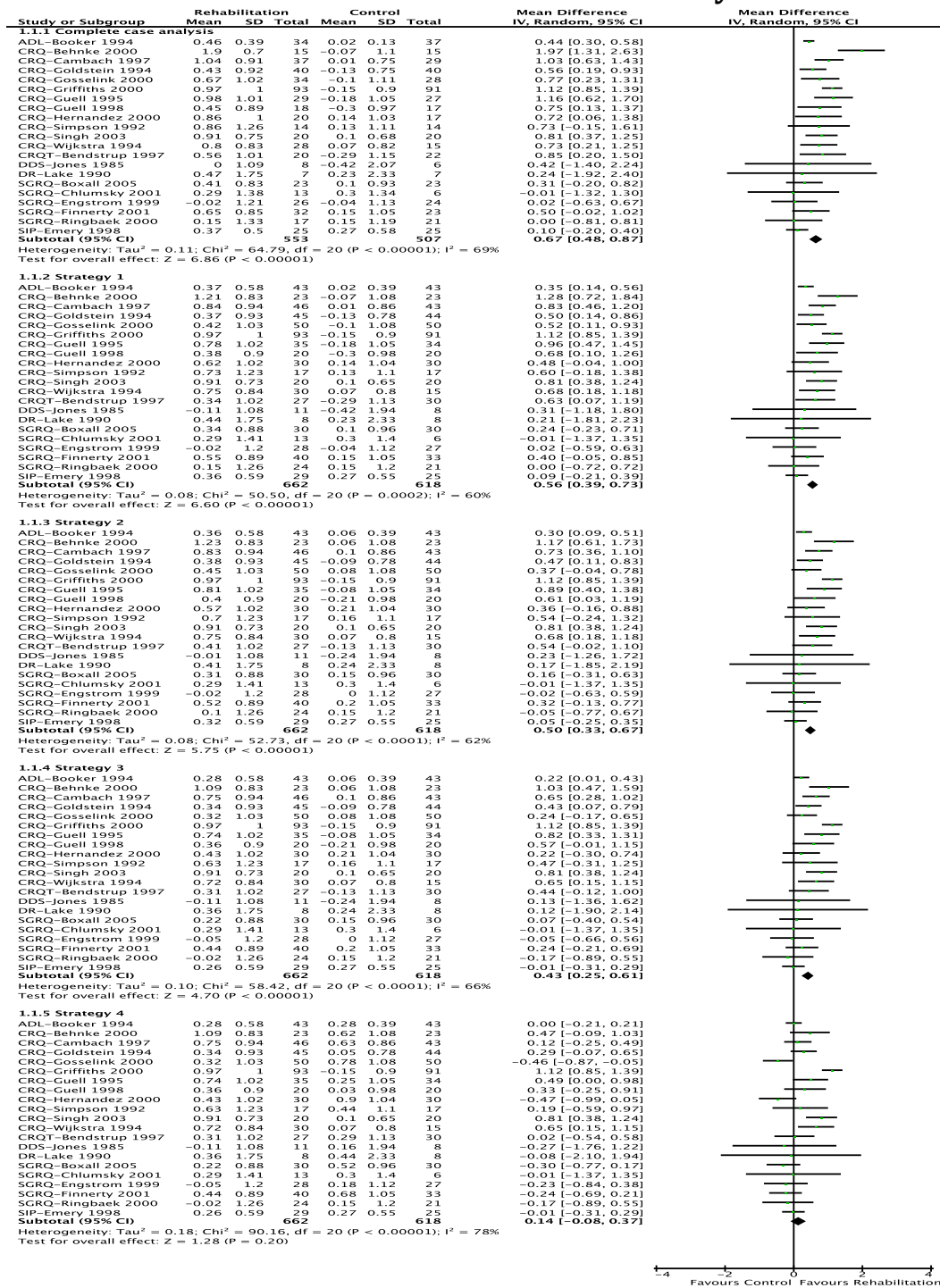
Of the 31 RCTs, 21 reported on health-related quality of life (HRQoL), of which 11 employed the Chronic Respiratory Disease Questionnaire (CRQ; 0 to 7 scale; higher scores represent better outcome) [17,18,19,20,21,22,23,24,25,26,27,28], 6 the St. Georges Respiratory Questionnaire (SGRQ; 0 to 100; higher scores represent better outcome) [22,29,30,31,32,33], 1 the Overall CRQ total score (0 to 140; higher scores represent better outcome) [34], 1 the Activities of Daily Living (ADL) Questionnaire (0 to 100; higher scores represent better outcome) [18], 1 the Sickness Impact Profile (SIP; 0 to 100; higher scores represent worse outcome) [35], 1 the Dyspnea Daily Diary score (0 to 5; higher scores represent worse outcome) [36], and 1 the Dyspnea rating (1 to 4; higher scores represent worse outcome) [37]. One study reported both the CRQ and SGRQ [22].

We began by choosing the reference instrument. Both the CRQ and SGRQ have an established MID, are widely used, and have well-established excellent measurement properties. The CRQ is, however, more responsive [38], and on this basis, we chose it as the reference instrument. The MID on the CRQ is 0.5 CRQ units on a 1 to 7 scale [39].

Of the 21 studies, 17 reported reasons for missing data. These included exacerbation of current condition, medical illness, changes in treatment, time constraints, lack of cooperation, social breakdown, personal or familial reasons, work-related reasons, admission to a nursing home, lost interest, transportation issues, and death. We did not find important differences in reasons for missing data across study arms or studies. One study reported that those lost to follow-up were more likely to be younger and male than those followed [35], and another study reported that those lost to follow-up were older than those followed [27].

The complete case analysis showed a pooled mean difference of 0.67 (95% confidence interval [CI] of 0.48 to 0.87) CRQ units. Applying Strategy 1 to 3 resulted in some loss of effect but maintained statistical significance. Strategy 4 (the most stringent) resulted in loss of statistical significance (Figure 23.1). Sensitivity analyses including and excluding the two studies that reported age and sex differences between those followed and those lost to follow-up showed no important differences.

Figure 23.1: Complete case analysis and sensitivity analyses using the four strategies for handling participants with missing data for continuous outcomes in a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease



Calculation of the proportion of patients who achieved an improvement equal to or greater than the MID of 0.5 resulted in an estimate of 29% (95% CI of 21% to 37%) in the complete case analysis. Strategies 1 to 3 results in point estimates of 24% to 18%, with lower confidence limits from 17% to 11% (Figure 23.2).

Figure 23.3 presents the summary effects of each strategy (showing the proportion of patients who achieved an improvement equal to or greater than the MID of 0.5) and provides a range of thresholds that guideline panels may adopt as an important effect, i.e., a significant proportion that benefit from the treatment. At the 20% threshold, confidence that the effect is important is secure only for the complete case analysis. At the 15% threshold, confidence that the effect is important is secure for all but the two most stringent strategies. At the 10% threshold, confidence that the effect is important is secure for all but the most stringent strategy.

Figure 23.2: Interpreting important treatment effects using the four strategies for handling participants with missing data for continuous outcomes in a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease

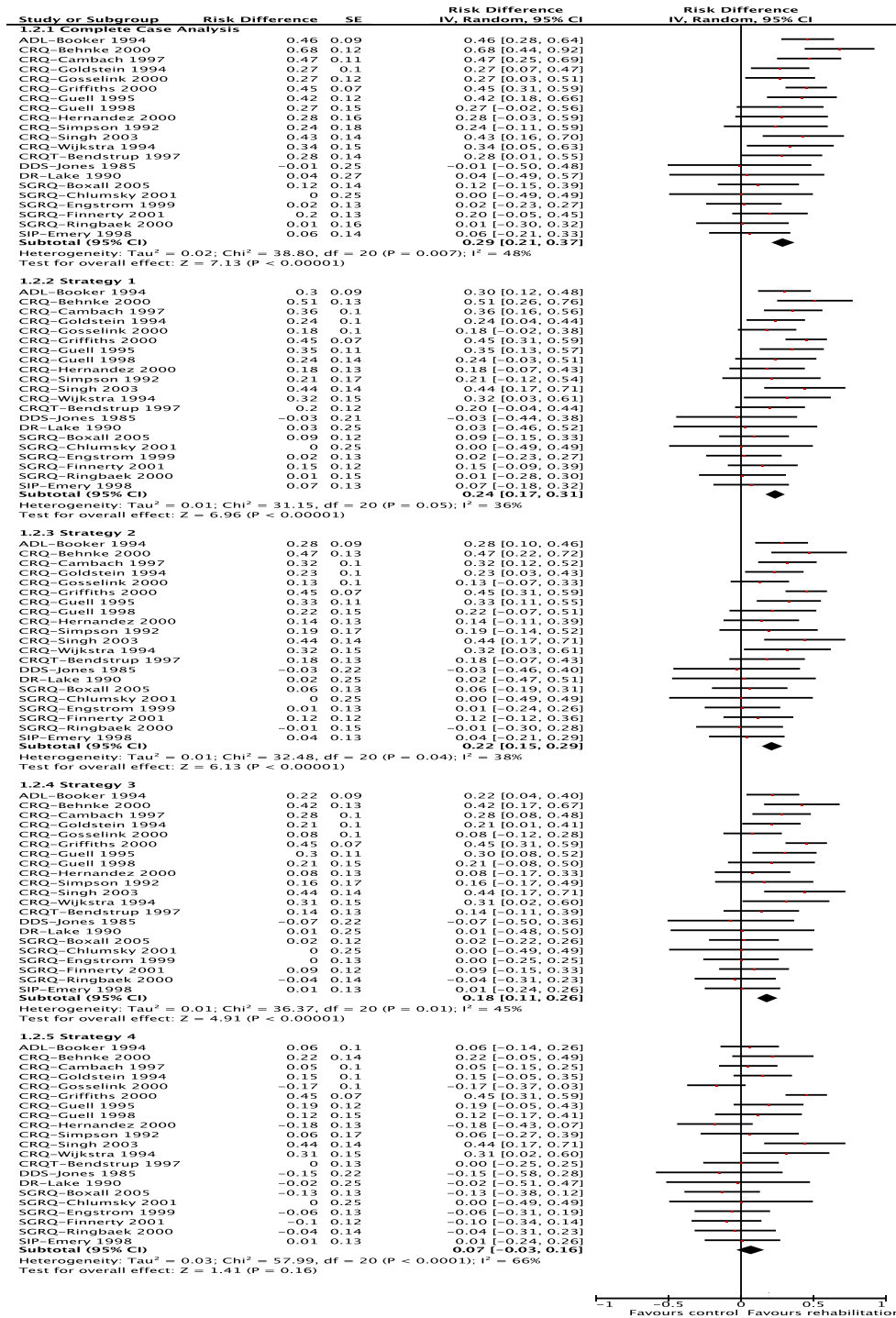
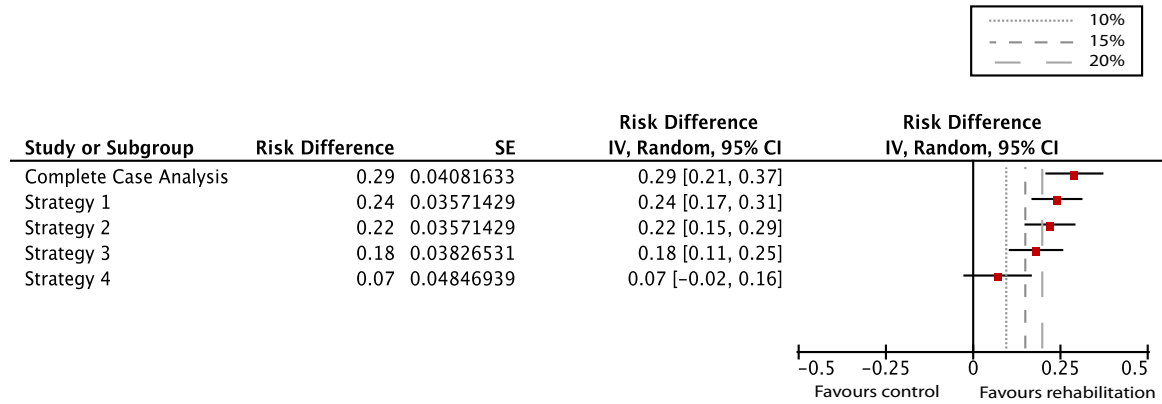


Figure 23.3: Interpreting important treatment effects using different thresholds in a systematic review of respiratory rehabilitation in Chronic Obstructive Pulmonary Disease



CHAPTER 24

DISCUSSION

24.1 Summary of findings

We developed a generalization of our previous approach for systematic reviewers to judge the risk of bias associated with individual trial missing data to situations in which trials have used different instruments to measure the same construct. This generalization involves choosing the reference instrument (the instrument with the best combination of an established anchor-based MID, excellent measurement properties, and high familiarity to the target audience), converting scores to units of the reference instrument, and then conducting a complete case analysis and four increasingly stringent imputation strategies [4]. When applying our approach to an example systematic review, we found that the pooled effect estimates withstood all but the most stringent assumption regarding missing data though, as we discuss below, the same may not be true of the lower boundary of the confidence intervals.

24.2 Implications on quality of evidence

When applying the GRADE/Cochrane Handbook criteria on how missing participant data impacts confidence in effect estimates (quality of evidence) [14,40], one needs to first assess whether the fourth imputation strategy (worst control mean score for

those with missing data in the intervention arm, and best intervention mean score for those with missing data in the control arm) is considered plausible, or too stringent and outside the realm of plausibility. Using our example of rehabilitation in patients with COPD, if a guideline panel considered that the fourth strategy is plausible, they are likely to consider confidence in estimates of effect seriously as a result of missing data. As a result, they would rate down confidence in estimates of effect for that outcome [14,40].

If, however, a guideline panel were to consider the most stringent assumption to be implausible, they would still need to consider whether the confidence interval around the effect for the other three strategies crosses a threshold between an effect that warrants recommending the intervention to an effect that does not. The minimal important difference on the CRQ is 0.5 units [39]. The lower confidence limit for the first three strategies is 0.39, 0.33 and 0.25 (Figure 23.1). A naive interpretation of these results would suggest that these effects represent trivial impact, thus warranting rating down confidence in an important effect of treatment.

This reasoning, however, assumes that all patients gain the same expected benefit from the intervention. Our method of interpreting the MID by calculating the proportion of patients who benefit from the treatment suggests, in the complete case analysis, that 29% of patients will have an important benefit from rehabilitation

(Figure 23.2). Further, the lower confidence limit suggests that at least 21% will achieve an important improvement (number needed to treat $100/21 \approx 5$ patients) (Figure 23.2), a magnitude of benefit that is likely to be considered important [41]. The conclusion would be similar for strategy 2 given a similar lower confidence limit around proportions achieving an important benefit (Figure 23.2).

The lower confidence limit in strategy 3 (11% benefiting and an NNT ≈ 9) may raise greater concern. To the extent that a guideline panel felt that even 10% of patients improving by the MID or greater would constitute an important treatment effect and any of these proportions benefiting would lead to similar inferences regarding the desirability of the intervention, they would not rate down confidence in estimates for risk of bias due to missing data for any of the three less stringent imputation strategies (Figure 23.3). If they considered that 11% benefiting from treatment would not be worth its undesirable consequences, they would rate down the quality of evidence for risk of bias (Figure 23.3).

24.3 Strengths, limitations and other issues

The main strength of our approach is its applicability to systematic reviews including trials that used different instruments to measure continuous data of the same construct. Additional strengths include the use of four increasingly stringent

strategies to test a range of assumptions regarding missing data, the use of the MID to calculate the proportion of patients who experience an important benefit, and compatibility with GRADE and Cochrane handbook guidance for addressing confidence in estimates on the basis of risk of bias as a result of missing participant data.

Our approach has limitations. The following are some possible situations in which one can encounter counterintuitive results: 1) using a combination of both change and end-of-study scores in the pooled analysis. One should use either one method or the other, preferably the change scores (if available or can be calculated) to potentially avoid extreme scores; 2) the best or worst means for imputations may not be plausible if these results come from very small trials. In these cases, one option is using the 2nd best or 2nd worst means for imputations.

We discussed other limitations of our approach previously [4]. These relate to some arbitrariness in the choices we made such as choosing 4 particular strategies from a potential 25 combinations, and assuming similar reasons for missing data across reviews. In our example systematic reviews, 17 of 21 studies reported reasons for missing data, which were similar across studies, and 2 studies reported important prognostic differences between those followed and those not followed. Our sensitivity

analyses including and excluding these 2 studies showed no important differences in results. We anticipate that our approach will be generalizable across most reviews.

Systematic review authors will find challenges applying our approach when authors of primary studies fail to adequately report missing participant data. We provide guidance for the most likely scenarios. For trials in which authors do not report missing participant data rates, we suggest using the median missing participant data rate from the systematic review. If one feels this assumption is too stringent, we propose performing a sensitivity analysis using a missing participant data rate of zero in both arms. For trials in which the authors report total missing participant data only, we suggest assuming that missing data was equally distributed in both arms. For trials in which the authors handled missing participant data and reported imputed analyses only, we suggest using the imputed results for the meta-analysis. If the authors reported both the imputed analysis and the complete case analysis, we suggest applying our approach to the trial's complete case analysis. Each of these suggestions assumes that systematic review authors have attempted, and failed, to obtain the missing information directly from authors of the individual trials.

24.4 Conclusions

In summary, we provide an approach to addressing missing participant data in systematic reviews in which trials use different instruments to measure and report continuous data on the same construct. This approach will facilitate increased rigor in the conduct of systematic reviews of continuous outcomes with missing participant data.

ACKNOWLEDGEMENTS, FUNDING, AND AUTHOR CONTRIBUTIONS

Conflicts of interest

None.

Funding

There were no sources of funding for this study. SE is supported by a Canadian Institutes of Health Research (CIHR) Doctoral Award. PA-C is funded by a Miguel Servet investigator contract from the Instituto de Salud Carlos III (CP09/00137).

Authors' contributions

SE and GHG conceived the study. SE, EAA, RAM, XS, SDW, DHA, PAO, BCJ and GHG contributed to the study design and developed the approach. SE and BCJ extracted the outcome estimates, missing data, and missingness mechanisms from the individual studies included in the example systematic review. SE applied the approach to the example systematic review. SE completed the first draft of the manuscript. SE is the first author and GG is the senior author of the study. All authors read, critically revised and approved the final manuscript to be published.

Authors' information

SE is a doctoral candidate in the Department of Clinical Epidemiology and Biostatistics at McMaster University. BCJ is an epidemiologist in the Department of Anaesthesia & Pain Medicine at The Hospital for Sick Children, Scientist in the Child Health Evaluative Sciences at the SickKids Research Institute, and a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. EAA is an Associate Professor in the Department of Internal Medicine at the American University of Beirut, and a part-time Associate Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. RAM is an assistant professor in the Department of Medicine, Division of Nephrology at University of Missouri-Kansas City, and a doctoral candidate in the Department of Clinical Epidemiology and Biostatistics at McMaster University. XS is a senior researcher at the Center for Clinical Epidemiology and Evidence-based Medicine in Xiqiao Hospital, Third Military Medical University in China and a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. SDW is a Professor Emeritus, in the Department of Clinical Epidemiology & Biostatistics at McMaster University. DHA is a part-time Assistant Professor in the Department of Clinical Epidemiology & Biostatistics at McMaster University. PA-C is a researcher in the Iberoamerican Cochrane Centre at the CIBERESP-IIB Sant Pau in Barcelona, Spain. GHG is a Distinguished Professor in the Departments of Medicine and Clinical Epidemiology and Biostatistics at McMaster University.

REFERENCES

1. Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, Mulla S, Lamontagne F, Bassler D, Vera C, Alshurafa M, Katsios CM, Zhou Q, Cukierman-Yaffe T, Gangji A, Mills EJ, Walter SD, Cook DJ, Schunemann HJ, Altman DG, Guyatt GH (2012) LOST to follow-up Information in Trials (LOST-IT): Potential impact on estimated treatment effects. *BMJ* 344: e2809.
2. The Cochrane Collaboration (2012) Further issues in meta-analysis: Intention to treat issues. The Cochrane Collaboration open learning material. London, UK: The Cochrane Collaboration.
3. Akl EA, Johnston BC, Alonso-Coello A, Neumann I, Ebrahim S, Briel M, Cook DJ, Guyatt GH (2013) Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers *PLoS ONE* 8: e57132.
4. Ebrahim S, Akl EA, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D, Alonso-Coello P, Johnston BC, Guyatt GH (2012) Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers. *J Clin Epidemiol* (Accepted for publication).
5. Schunemann HJ, Guyatt GH (2005) Commentary--goodbye M(C)ID! Hello MID, where do you come from? *Health Services Research* 40: 593-597.
6. Johnston BC, Thorlund K, Schünemann HJ, Xie F, Murad MH, Montori VM, Guyatt GH (2010) Improving the interpretation of quality of life evidence in meta-

- analyses: the application of minimal important difference units. *Health Qual Life Outcomes* 8: 116.
7. Groenwold RHH, Donders RT, Roes KCB, Harrell FE Jr, Moons KGM (2011) Dealing With Missing Outcome Data in Randomized Trials and Observational Studies. *American Journal of Epidemiology* 175: 210-217.
 8. Higgins JPT, White IR, Wood AM (2008) Imputation methods for missing outcome data in meta-analysis of clinical trials. *Clinical Trials* 5: 225-239.
 9. The Cochrane Collaboration (2011) Chapter 8: Risk of Bias. In: Higgins JPT, Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0)*. Oxford, UK: Cochrane Collaboration.
 10. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, et al (2010) CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 340: c869.
 11. Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH (2012) Pooling health-related quality of life outcomes in meta-analysis—a tutorial and review of methods for enhancing interpretability. *Research Synthesis Methods* 2: 188-203.
 12. Norman GR, Sridhar FG, Guyatt GH, Walter SD (2001) Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 39: 1039-1047.
 13. Guyatt GH, Feeny DH, Patrick DL (1993) Measuring health-related quality of life. *Ann Intern Med* 118: 622-629.

14. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336: 924-926.
15. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJH, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JW Jr, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schunemann HJ (2011) GRADE guidelines 6. Rating the quality of evidence-imprecision. *J Clin Epidemiol* 64: 1283-1293.
16. Lacasse Y, Goldstein R, Lasserson TJ, Martin S (2006) Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 18: CD003793.
17. Behnke M, Taube C, Kirsten D, Lehnigk B, Jörres RA, Magnussen H (2000) Home-based exercise is capable of preserving hospital-based improvements in severe chronic obstructive pulmonary disease. *Respiratory Medicine* 94: 1184-1191.
18. Booker HA (1984) Exercise training and breathing control in patients with chronic airflow limitation. *Physiotherapy* 70: 258-260.
19. Cambach W, Chadwick-Straver RVM, Wagenaar RC, van Keimpema ARJ, Kemper HCG (1997) The effects of a community-based pulmonary rehabilitation programme on exercise tolerance and quality of life: a randomized controlled trial. *European Respiratory Journal* 10: 104-113.

20. Elías Hernandez MT, Montemayor Rubio T, Ortega Ruiz F, Sanchez Riera H, Sanchez Gil R, Castillo Gomez J (2000) Results of a home-based training program for patients with COPD. *Chest* 118: 106-114.
21. Goldstein RS, Gort EH, Stubbing D, AvendanoMA, Guyatt GH (1994) Randomised controlled trial of respiratory rehabilitation. *Lancet* 344: 1394-1397.
22. Griffiths TL, Burr ML, Campbell IA, Lewis-Jenkins V, Mullins J, Shiels K, et al. (2000) Results at 1 year of outpatient multidisciplinary pulmonary rehabilitation: a randomised controlled trial. *The Lancet* 355: 362-368.
23. Güell R, Morante F, Sangenis M, Casan P (1995) Effects of respiratory rehabilitation on quality of life of patients with chronic obstructive pulmonary disease. *Annals de Medicina* 81: 9.
24. Güell R, Gonzalez Y, Martorell B, Gonzalez A, Sotomayor C, Sangenis M, et al (1998) Impact of pulmonary rehabilitation on personality traits and styles in COPD patients *European Respiratory Journal* 12: 228S.
25. Simpson K, Killian K, McCartney N, et al (1992) Randomised controlled trial of weightlifting exercise in patients with chronic airflow limitation. *Thorax* 47: 70-75.
26. Singh V, Khandelwal DC, Khandelwal R (2003) Pulmonary rehabilitation in patients with chronic obstructive pulmonary diseases. *Indian Journal of Chest Diseases and Allied Sciences* 45: 13-17.

27. Troosters T, Gosselink R, Decramer M (2000) Short- and long-term effects of outpatient rehabilitation in patients with chronic obstructive pulmonary disease: a randomized trial. *American Journal of Medicine* 109: 207-212.
28. Wijkstra PJ, van Altena R, Kraan J, et al (1994) Quality of life in patients with chronic obstructive pulmonary disease improves after rehabilitation at home. *European Respiratory Journal* 7: 269-273.
29. Boxall AM, Barclay L, Sayers A, Caplan GA (2005) Managing chronic obstructive pulmonary disease in the community: A randomized controlled trial of home-based pulmonary rehabilitation for elderly housebound patients. *Journal of Cardiopulmonary Rehabilitation* 25: 378-385.
30. Chlumsky J, Sterbova L, Smolikova L, Matous M, Salajka F (2001) The effect of pulmonary rehabilitation on exercise tolerance and quality of life in patients with COPD. Preliminary data. *European Respiratory Journal* 18: 223S.
31. Engström CP, Persson LO, Larsson S, Sullivan M (1999) Long-term effects of a pulmonary rehabilitation programme in outpatients with chronic obstructive pulmonary disease: A randomized controlled study. *Scandinavian Journal of Rehabilitation Medicine* 31: 207-213.
32. Finnerty JP, Keeping I, Bullough I, Jones J (2001) The effectiveness of outpatient pulmonary rehabilitation in chronic lung disease. A randomized controlled trial. *Chest* 119: 1705-1710.

33. Ringbaek TJ, Broendum E, Hemmingsen L, Lybeck K, Nielsen D, Andersen C, et al (2000) Rehabilitation of patients with chronic obstructive pulmonary disease. Exercise twice a week is not sufficient! *Respiratory Medicine* 94: 150-154.
34. Bendstrup KE, Ingemann Jensen J, Holm S, Bengtsson B (1997) Outpatient rehabilitation improves activities of daily living, quality of life and exercise tolerance in chronic obstructive pulmonary disease. *European Respiratory Journal* 10: 2801-2806.
35. Emery CF, Schein RL, Hauck ER, MacIntyre NR (1998) Psychological and cognitive outcomes of a randomised trial of exercise among patients with chronic obstructive pulmonary disease. *Health Psychology* 17: 232-240.
36. Jones DT, Thomson RJ, Sears MR (1985) Physical exercise and resistive breathing in severe chronic airways obstruction - are they effective? *European Journal of Respiratory Diseases* 67: 159-166.
37. Lake FR, Henderson K, Briffa T, et al (1990) Upper-limb and lower-limb exercise training in patients with chronic airflow obstruction. *Chest* 97: 1077-1082.
38. Puhan MA, Guyatt GH, Goldstein R, Mador J, McKim D, Stahl E, Griffith L, Schünemann HJ. (2007) Relative responsiveness of the Chronic Respiratory Questionnaire, St. Georges Respiratory Questionnaire and four other health-related quality of life instruments for patients with chronic lung disease. *Respir Med* 10: 308-316.

39. Schünemann HJ, Puhan M, Goldstein R, Jaeschke R, Guyatt GH (2005)
Measurement properties and interpretability of the chronic respiratory disease questionnaire (CRQ). *COPD* 2: 81-89.
40. The Cochrane Collaboration (2011) Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.10)*. Oxford, UK: Cochrane Collaboration.
41. Ries AL, Bauldoff GS, Carlin BW, Casaburi R, Emery CF, Mahler DA, Make B, Rochester CL, Zuwallack R, Herrerias C (2007) Pulmonary Rehabilitation: Joint ACCP/AACVPR Evidence-Based Clinical Practice Guidelines. *Chest* 131: 4S-42S.

THESIS CONCLUSIONS

CHAPTER 25

CONCLUSIONS

25.1 Overview

This thesis focused on issues related to determining optimal management of patients in receipt of disability benefits, specifically focusing on the effectiveness of therapies for depression in patients receiving disability benefits, and addressing a key methodological issue relevant to all systematic reviews including patients on disability benefits: missing participant data for continuous outcomes. In the remainder of this chapter, I discuss key findings, limitations and future directions under two separate headings: management of patients in receipt of disability benefits, and addressing missing participant data.

25.2 Management of patients in receipt of disability benefits

With the support of a large, private, Canadian disability insurer, I completed two studies to explore the effectiveness of therapies for the management of depression in patients receiving disability benefits.

In the first study, we systematically reviewed the evidence from randomized controlled trials on the effectiveness of CBT in patients with depression in receipt of disability benefits in comparison to those not receiving disability benefits. We found 92 potentially eligible trials, of which none reported whether they had enrolled patients in receipt of disability benefits. This is a significant limitation in the literature, as trialists are either not considering this question even though up to 14% of disability burden worldwide is due to depression, or not reporting it in their published report. After successful contact with 80% of trialists, we found that three trials had actually captured disability benefit status but did not include this information in their published reports. We successfully obtained raw data from 2 of the 3 trials to complete a more powerful individual patient data meta-analysis.

Our analysis of 227 patients enrolled in these 2 studies found that the effect of CBT for depression (measured by the Beck Depression Inventory-II) was not significantly different between those receiving disability benefits versus those not receiving disability benefits. The confidence interval around our effect suggested a possible greater effect in those receiving disability benefits but our findings were limited by the small number of patients (34 patients receiving disability benefits versus 193 not receiving disability benefits) that were available for analyses. It thus appears that the disability benefit status of depressed patients have been largely ignored in existing trials of CBT for depression.

In the second study, we explored a disability insurer's administrative data to evaluate the association between the provision of psychotherapy (including CBT) and time to both STD and LTD claim closure (a surrogate for return to work). In these analyses, we also examined the association between other possible determinants of STD and LTD claim closure.

We found that LTD claimants receiving psychotherapy closed their claim 1.42x (99% CI of 1.30 to 1.55) faster than LTD claimants not receiving psychotherapy.

Paradoxically, however, STD claimants not receiving psychotherapy closed their claims 1.23 times (99% CI of 1.03 to 1.47) faster than STD claimants receiving psychotherapy. From these results, it is unclear if STD claimants are in fact different from LTD claimants, if our analysis was affected by selection bias, or if the effects of psychotherapy may not be realized during the limited time that this therapy is provided among STD claims.

We also found other important factors that were associated with both STD and LTD claim closure: older age, having a primary diagnosis of recurrent depression versus major depression, having a secondary psychological or non-psychological diagnosis and having an administrative services only or refund policy versus non-refund policy

were associated with longer claim closure, and residing in the Prairies versus Ontario was associated with faster claim closure. We found some inconsistent associations between STD and LTD claim closure. For example, longer time to claim approval was associated with faster STD claim closure but longer time to LTD claim closure, and residing in Quebec versus Ontario was associated with longer time to STD claim closure but faster LTD claim closure. These findings warrant further exploration. Additionally, two factors predictive of claim closure, time to claim approval and type of funding policy, are novel findings that have not been reported in previous literature.

Our review of the administrative data noted a number of issues that limit the strength of our findings (e.g., many existing variables were not captured in a systematic manner such as capturing psychotherapy instead of CBT, a number of important factors such as illness severity and patient beliefs were not captured).

Our two studies provide little support for the hypothesis that CBT is less effective in patients receiving disability benefits than in other patients. Our studies, however, were limited by the very small number of patients enrolled in the relevant RCTs, by the limitations of the administrative database, and by the limitations of any inferences from observational studies. Thus, our findings highlight the need for well-designed prospective studies to establish factors associated with recovery among depressed

patients in receipt of disability benefits and randomized trials to definitively establish the effect of CBT in depressed patients in receipt of disability benefits. These however would be best completed with support from disability insurers. This research ought to be carried out before we can confidently decide whether clinicians and payers should decrease, continue or expand the use of psychotherapy in managing these patients.

Our research group has already begun setting the stage for completing a large international multicentre cohort study to establish factors associated with prolonged recovery in patients receiving disability benefits, in collaboration with multiple insurers in Canada, the Netherlands, Switzerland and Australia. This study will result in the first validated instrument capable of identifying claims at risk of prolonged recovery across all clinical conditions, which will facilitate effective triaging and assignment of resources to claimants' treatments based on the severity of their condition. This study will also provide essential information for the design of randomized controlled trials to test interventions designed to improve the prognosis of high-risk claims.

25.3 Addressing missing participant data for continuous outcomes in systematic reviews

Issues related to missing participant data in continuous outcomes arose out of the systematic review evaluating CBT for depression in patients receiving disability benefits. Trials in our CBT review had a significant percentage of missing participant data (median 21%, range 0 to 41%). We found that there was no guidance in the literature or the Cochrane handbook on how systematic review authors should address missing participant data for continuous outcomes and what its implications are on risk of bias, aside from rating down on the Cochrane Risk of Bias tool. With no guidance, most systematic review authors ignore participants with missing data and carry out a complete case analysis on those with available data.

We addressed this gap by developing an approach consisting of a range of increasingly stringent plausible assumptions of what may have happened to those with missing participant data. We provided detailed guidance for conducting these analyses, and applied it to two example reviews. Our examples illustrated three properties of our approach: 1) small effects with borderline significance are more likely to lose significance than larger, precise effects, 2) the greater the percentage of participants with missing data, the greater the risk of losing significance, and 3) the more extreme the scores of imputation in the proposed strategies, the more likely the newly calculated effects will deviate from the complete case analysis results. Our approach

was not restricted to patients on disability, and is applicable to systematic reviews of randomized trials across all disciplines.

Our fourth paper extended our approach to address missing participant data in systematic reviews in which different instruments report continuous data measuring the same construct. This involved an initial conversion of scores from all instruments to the units of a reference instrument that measured the same construct, and then application of our imputation strategies. In this study, we also provided guidance to systematic review authors on how to calculate the proportion of patients who experience a clinically important treatment effect derived from the strategies, and how to assess quality of evidence on the basis of risk of bias due to missing participant data.

There were limitations to our approach including the arbitrariness of choosing 4 strategies out of a possible 25 combinations, choice of best/worst means for our sources of imputations of effect, and how to address missing data that may be due to differences in prognostic characteristics.

These papers helped us strategize future directions. Our next objective is to empirically test our approach with a large number of systematic reviews. This would

help uncover problems or issues that may not have been initially identified in the examples used in our papers. Second, we are planning to extend our approach to address harm outcomes (e.g., adverse events). Third, we plan to describe the methods of reporting (e.g., author's plan to collect frequency of and reasons for missing participant data), handling (e.g., author's description of methods for addressing missing participant data), and judging quality of evidence (e.g., author's assessment of risk of bias and confidence in estimates, and its implications on findings) associated with missing participant data for continuous outcomes in systematic reviews of RCTs. These steps will help advance the rigor of systematic reviews.

APPENDICES

APPENDIX A

Search strategies of the IPDMA of CBT for depression in patients receiving disability benefits

MEDLINE

#3 Search #1 AND #2 Limits: Humans, Randomized Controlled Trial, All Adult: 19+ years, Publication Date from 2011/01/01 10:22:43

#2 Search Depression OR depressive

#1 Search behavior therapy OR biofeedback OR cognitive analytic therapy OR cognitive behavior therapy OR counseling OR family therapy OR marital therapy OR psychoanalytic therapy OR psychotherapy OR relaxation therapy

EMBASE

1 psychotherapy.mp.

2 depression.mp.

3 random*.mp.

4 1 and 2 and 3

5 limit 4 to yr="2011 - Current"

PSYCINFO

1 clinical trials/

2 (depression or depressive).mp.

3 1 and 2

4 limit 3 to yr="2011"

CENTRAL

#1 (Behavior-therapy OR Biofeedback OR Cognitive analytic therapy OR Cognitive behavior therapy OR Cognitive-behavior-therapy OR Cognitive behaviour therapy OR Counselling OR Counseling OR Family therapy OR Marital therapy OR Psychoanalytic therapy OR Psychoanalysis OR Psychotherapy OR Relaxation therapy) AND (depression OR depressive)

#2 (#1), in 2011

CINAHL

- S19 S17 or S18
- S18 S3 AND S15 Limiters - Clinical Queries: Review - High Sensitivity
Search modes - Boolean/Phrase
- S17 S3 and S15 Limiters - Clinical Queries: Review - High Sensitivity
Search modes - Boolean/Phrase
- S16 S3 and S15
- S15 S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14
- S14 CBASP
- S13 "functional analytic psychotherapy"
- S12 "meta cognitive therapy"
- S11 "metacognitive therapy"
- S10 "acceptance and commitment therapy"
- S9 "mindfulness"
- S8 "rational emotive therapy"
- S7 (MH "Reality Therapy")
- S6 "cognitive behavio#r* therapy"
- S5 (MH "Behavior Therapy")
- S4 (MH "Cognitive Therapy")
- S3 S1 or S2
- S2 "depressive disorder"
- S1 (MH "Depression+")

AMED

- 1 depression/
- 2 depressive disorder/
- 3 dysthymia.mp.
- 4 cognitive therapy/
- 5 behavior therapy/
- 6 cognitive behavio?ral therapy.mp.
- 7 reality therapy/
- 8 rational emotive therapy.mp.
- 9 (acceptance and commitment therapy).mp. [mp=abstract, heading words,
title]
- 10 or/1-3
- 11 or/4-9
- 12 10 and 11

APPENDIX B

Published Article- IPDMA of CBT for depression in patients receiving disability benefits

Ebrahim S, Montoya L, Hsu S, Kamal el Din M, Truong W, Carrasco-Labra A, Busse JW, Walter SD, Heels-Ansdell D, Couban R, Patelis-Siotis I, Bellman M, de Graaf LE, Dozois D, Bieling P, Guyatt GH. Effectiveness of Cognitive Behavioural Therapy for depression in patients receiving disability benefits: A systematic review and individual patient data meta-analysis. PLoS ONE. 2012;7(11):e50202

**PLOS journals apply the Creative Commons Attribution License (CCAL) to all works they publish. Under the CCAL, authors retain ownership of the copyright for their article, but authors can allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in PLOS journals, so long as the original authors and source are cited. No permission is required from the authors or the publishers.*

Effectiveness of Cognitive Behavioral Therapy for Depression in Patients Receiving Disability Benefits: A Systematic Review and Individual Patient Data Meta-Analysis

Shanil Ebrahim^{1*}, Luis Montoya², Wanda Truong³, Sandy Hsu⁴, Mostafa Kamal el Din⁵, Alonso Carrasco-Labra^{1,6}, Jason W. Busse⁷, Stephen D. Walter^{1,8}, Diane Heels-Ansdell¹, Rachel Couban⁷, Irene Patelis-Siotis^{9,10}, Marg Bellman¹¹, L. Esther de Graaf¹², David J. A. Dozois¹³, Peter J. Bieling⁹, Gordon H. Guyatt^{1,14}

1 Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada, **2** Department of Dentistry, Santo Tomas University, Bogota D.C., Colombia, **3** McMaster Integrative Neuroscience Discovery & Study program, McMaster University, Hamilton, Canada, **4** Department of Oncology, McMaster University, Hamilton, Canada, **5** Ain Shams University, Faculty of Medicine, Cairo, Egypt, **6** Evidence-Based Dentistry Unit, Faculty of Dentistry, University of Chile, Santiago, Chile, **7** Department of Anesthesia, McMaster University, Hamilton, Canada, **8** Department of Mathematics and Statistics, McMaster University, Hamilton, Canada, **9** Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, Canada, **10** Mood Disorders program, St. Joseph's Healthcare, Hamilton, Canada, **11** National Disability Services, Policy & Procedure department, Sun Life Financial, Toronto, Canada, **12** Department of Medical Psychology and Psychotherapy, Erasmus MC University Medical Center, Rotterdam, The Netherlands, **13** Department of Psychology, University of Western Ontario, London, Canada, **14** Department of Medicine, McMaster University, Hamilton, Canada

Abstract

Objectives: To systematically summarize the randomized trial evidence regarding the relative effectiveness of cognitive behavioural therapy (CBT) in patients with depression in receipt of disability benefits in comparison to those not receiving disability benefits.

Data Sources: All relevant RCTs from a database of randomized controlled and comparative studies examining the effects of psychotherapy for adult depression (<http://www.evidencebasedpsychotherapies.org>), electronic databases (MEDLINE, EMBASE, PSYCINFO, AMED, CINAHL and CENTRAL) to June 2011, and bibliographies of all relevant articles.

Study Eligibility Criteria, Participants and Intervention: Adult patients with major depression, randomly assigned to CBT versus minimal/no treatment or care-as-usual.

Study Appraisal and Synthesis Methods: Three teams of reviewers, independently and in duplicate, completed title and abstract screening, full text review and data extraction. We performed an individual patient data meta-analysis to summarize data.

Results: Of 92 eligible trials, 70 provided author contact information; of these 56 (80%) were successfully contacted to establish if they captured receipt of benefits as a baseline characteristic; 8 recorded benefit status, and 3 enrolled some patients in receipt of benefits, of which 2 provided individual patient data. Including both patients receiving and not receiving disability benefits, 2 trials (227 patients) suggested a possible reduction in depression with CBT, as measured by the Beck Depression Inventory, mean difference [MD] (95% confidence interval [CI]) = -2.61 (-5.28, 0.07), $p=0.06$; minimally important difference of 5. The effect appeared larger, though not significantly, in those in receipt of benefits (34 patients) versus not receiving benefits (193 patients); MD (95% CI) = -4.46 (-12.21, 3.30), $p=0.26$.

Conclusions: Our data does not support the hypothesis that CBT has smaller effects in depressed patients receiving disability benefits versus other patients. Given that the confidence interval is wide, a decreased effect is still possible, though if the difference exists, it is likely to be small.

Citation: Ebrahim S, Montoya L, Truong W, Hsu S, Kamal el Din M, et al. (2012) Effectiveness of Cognitive Behavioral Therapy for Depression in Patients Receiving Disability Benefits: A Systematic Review and Individual Patient Data Meta-Analysis. *PLoS ONE* 7(11): e50202. doi:10.1371/journal.pone.0050202

Editor: Robert S. Phillips, University of York, United Kingdom

Received: August 5, 2012; **Accepted:** October 17, 2012; **Published:** November 29, 2012

Copyright: © 2012 Ebrahim et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: There were no funds received for this study. Shanil Ebrahim was supported by MITACS Accelerate scholarship (partnered with Sun Life Financial Assurance Company of Canada) and the Canadian Institutes of Health Research Doctoral Award during the completion of this study. Jason W. Busse is supported by a New Investigator Award from the Canadian Institutes of Health Research and Canadian Chiropractic Research Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have read the journal's policy and have the following conflicts. Sun Life was supporting the MITACS scholarship for Shanil Ebrahim, and did not have a role in the study design, data collection and analysis, or preparation of the manuscript. Jason W. Busse acts as a consultant to Prisma Health Canada, a private incorporated company funded by employers and insurers that consults on and manages long-term disability claims. Marg Bellman is the assistant vice-president of the National Disability Services, Policy & Procedure department at Sun Life Financial. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials. All other authors have no affiliations with or involvement in any organization or entity with a direct interest of any kind in the subject matter or materials discussed in the manuscript. No financial support has been received for the creation of this manuscript.

* E-mail: shanil.ebrahim@utoronto.ca

Introduction

Major Depressive Disorder (henceforth referred to as depression) results in immense human suffering and an enormous socioeconomic burden. Depression accounts for 11% of disability worldwide and an estimated productivity loss of \$17 to \$44 billion in the USA [1,2]. Depression is expected to become the second leading cause of disease burden worldwide by the year 2020 [3].

The National Institute for Health and Clinical Excellence (NICE) in the UK recommends that health care professionals provide pharmacological treatments and/or high-intensity psychological interventions for individuals suffering from depression. Pharmacological treatments may accelerate recovery from depression, particularly when symptoms are severe [4] and, over the last few decades, their use has increased dramatically in Western nations [5,6]. NICE guidelines suggest psychological therapies should be offered to individuals suffering from persistent subthreshold symptoms of depression, mild to moderate depression, and those with a high risk of relapse or those declining pharmacological treatment for severe depression [5,6].

Cognitive Behavioral Therapy (CBT) is a common non-pharmacological treatment for depression [5,7]. CBT is based on three fundamental propositions: cognitive activity affects behavior, cognitive activity can be monitored and altered, and desired behavior change may be affected through cognitive change [7]. Twelve systematic reviews evaluating CBT in individuals suffering from depression have demonstrated that CBT reduces depressive symptoms [8,9,10,11,12,13,14,15,16,17,18,19], with the most current and rigorous meta-analysis reporting a pooled standardized mean difference (SMD) of 0.69 (95% confidence interval [CI] of 0.59 to 0.79) [13].

In North America, depression is one of the most frequent reasons for receiving disability benefits [20,21], and disability claims for mental health disorders incur greater costs compared to other disorders [22]. In those receiving disability benefits, individuals suffering from mental health disorders require more treatment and have greater difficulty returning to work than those suffering from other conditions [23]. Although CBT is one of the most frequently reimbursed therapies by insurers, its utilization by insurance companies still remains relatively low at approximately 3% for short-term disability claimants and 15% for long-term disability claimants [24].

CBT may be less effective, or ineffective, in patients receiving disability benefits, because their circumstances or psychological status may interfere with its successful implementation [25]. This may also be associated with the compensation process [26], secondary gain from financial benefits (benefits of assuming a sick

role) [27], or the adversarial nature of litigation [28]. A recent meta-analysis of 129 studies in surgical populations that found a substantially greater risk of an unsatisfactory outcome (functional, quality of life, pain and patient satisfaction) after surgery in compensated patients (odds ratio [95% CI] = 3.79 [3.28 to 4.37]) provides indirect evidence for this hypothesis [29]. The effectiveness of CBT for depression in patients receiving disability benefits has received little attention.

Objectives

The purpose of our study was to perform a systematic review and an individual patient data meta-analysis of all randomized controlled trials (RCTs) that compared the effectiveness of CBT to minimal/no treatment, or care-as-usual, in patients with depression receiving versus those not receiving disability benefits.

Questions

In adult patients with depression, is there a difference in the effect of CBT on depression between those receiving disability benefits compared those not receiving disability benefits?

Methods

We used the PRISMA guidelines [30] to report our findings.

Protocol and registration

We developed a protocol prior to conducting the study but did not register it.

Eligibility criteria

Eligible studies met the following criteria: 1) random allocation of adult patients to CBT or a control arm consisting of minimal/no treatment, treatment as usual (TAU) or pharmacotherapy if it was equally balanced in the treatment groups (e.g. CBT plus pharmacotherapy versus pharmacotherapy alone), and 2) inclusion of patients with depression, classified as Major Depressive Disorder by any edition of the Diagnostic and Statistical Manual (DSM), International Classification of Diseases (ICD), Research Diagnostic Criteria (RDC) or other diagnostic system [31].

Information sources

We identified all relevant RCTs from a database of randomized controlled trials and comparative studies examining the effects of psychotherapy for adult depression (<http://www.evidencebasedpsychotherapies.org>) [32]. This database consisted of 281 trials and was identified from searching the following

electronic databases in all languages: PUBMED, EMBASE, PsycINFO and Cochrane Central Register of Controlled Trials, from inception until January 1, 2011 [11]. In addition to the 281 trials, we updated the search with the assistance of an experienced academic librarian (RC) until June 13, 2011 for each electronic database, and also searched AMED and CINAHL. We hand searched the reference lists of all relevant RCTs for additional eligible trials.

Search

Our search strategy including keywords and MESH headings are provided in Appendix A.

Study selection

Three teams of reviewers (SE, SH, LM, WT, MK, ACL) worked in pairs and screened titles and abstracts of identified citations, independently and in duplicate, using a standardized, pilot-tested screening form. The same reviewers independently applied eligibility criteria to the full text of potentially eligible studies. One psychiatrist (IPS) and one psychologist (RM), blinded to study results, independently reviewed and confirmed eligibility of therapies that were not explicitly described by trial authors as CBT. We measured agreement for the full text review stage, and interpreted the agreement statistics using the guidelines proposed by Landis and Koch [33]. Kappa values of 0 to 0.20 represented slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61 to 0.80 substantial agreement, and greater than 0.80 almost perfect agreement.

Reviewers grouped eligible articles into one of the four categories: (i) studies that did not explicitly state if they included or excluded patients receiving disability benefits, (ii) studies that explicitly excluded patients receiving disability benefits, (iii) studies that explicitly included patients receiving disability benefits but did not separately report outcomes based on receipt of disability benefits, and (iv) studies that explicitly included patients receiving disability benefits and reported outcomes separately based on receipt of disability benefits. Disability benefits were defined as wage replacement benefits administered by a third party (e.g. insurer).

Contacting authors of eligible studies

We identified 88 studies in category i, 4 in category iii, and none in either category ii or iv. Contact information was not reported and not available through an Internet search for authors of 22 (24%) trials. We attempted to contact authors of the remaining 70 trials by email and requested information on whether they had an eligibility stipulation for disability status. If authors included patients on disability benefits, we requested their trial data to facilitate an individual patient data meta-analysis (IPDMA). To maintain patient confidentiality, authors removed any personal identifiers from their dataset prior to transferring it to our center. We clarified uncertainties or discrepancies in the data sets with the study authors and combined individual patient data for variables that were similar across the trials. Based on authors' replies, we classified studies into four groups: (A) those that did have some data specific to patients on disability benefits, (B) those that confirmed that they had no patients on disability benefits, (C) those that did not have an eligibility criterion for disability status and did not collect information on disability status, and (D) unknown or did not respond.

Data collection process

Using piloted standardized forms and a detailed instruction manual to extract data, the same teams of reviewers extracted

data, independently and in duplicate, from studies in groups A and B. We did not abstract data from groups C and D.

Data abstracted included patient characteristics, treatment effect on depression, frequency and timing of follow-up, details of depression (including diagnostic classification system used, severity of depression, and duration of depression), and CBT intervention details (including the type of CBT administered, expertise of providers administering CBT, and frequency of CBT). Reviewers abstracted data from the following study arms: CBT, TAU and minimal or no treatment. Data comparing CBT only to active comparators were not abstracted, unless the active comparator was equally balanced between both the treatment and control group.

Risk of Bias in individual studies

Using a modified Cochrane risk of bias instrument, reviewers assessed risk of bias for each eligible trial on the following domains: sequence generation; allocation concealment; blinding of participants, investigators, data collectors, outcome assessors, and data analysts; incomplete outcome data; selective outcome reporting; and other sources of bias (e.g. bias of study design, trial stopped early, extreme baseline imbalance, and fraudulent trial) [34,35]. Reviewers used response options of “definitely yes”, “probably yes”, “probably no”, and “definitely no” with definitely and probably yes ultimately assigned high risk of bias and probably and definitely no assigned low risk of bias [35]. The reviewers resolved disagreements by discussion, and an arbitrator (JWB) adjudicated any remaining conflicts.

Synthesis of results

For our IPDMA, we compared the effects (mean difference) of CBT on depression, measured by the most commonly reported instrument [Beck Depression Inventory (BDI-II)], in patients receiving disability benefits versus patients not receiving disability benefits. We used a one-stage method [36], and included the following variables in our model: study arm, receipt of disability benefits, interaction term of study arm and receipt of disability benefits, trial as a categorical variable, age and baseline BDI-II score. To guard against multiplicity of data [37], we used the most common follow-up time point of 3 months for our analysis.

Our secondary analyses evaluated whether there were differences in patients not in receipt of disability benefits between trials that included patients in receipt of disability benefits (group A) and trials with aggregate data that did not include patients receiving disability benefits (group B). We compared the following: 1) the effects of CBT between group A and B; 2) the effects of CBT between group A and B that compared CBT plus pharmacotherapy versus pharmacotherapy alone; 3) the effects of CBT between group A and B that compared CBT to TAU.

For our secondary analyses, we used the 2-stage method [38]. In the first stage, we aggregated the IPD data of the patients not receiving disability benefits in group A and in the second stage, pooled the aggregate data of studies in group A and B using a random-effects model.

We used the means and standard deviations (SDs) of the end of study scores for our secondary pooled analyses. To pool data across trials and to facilitate interpretation for clinicians and other stakeholders, we calculated the mean difference (MD) and its associated 95% confidence interval (CI) of the natural units of the most familiar instrument across trials, the BDI-II. For this calculation, we used the following formulas to convert mean estimates (M) and standard deviations (SD) into the scale of the most familiar instrument: $M_A = (M_B - L_B) (R_A/R_B) + L_A$ and $SD_A = SD_B (R_A/R_B) + L_A$, where A represented the most familiar

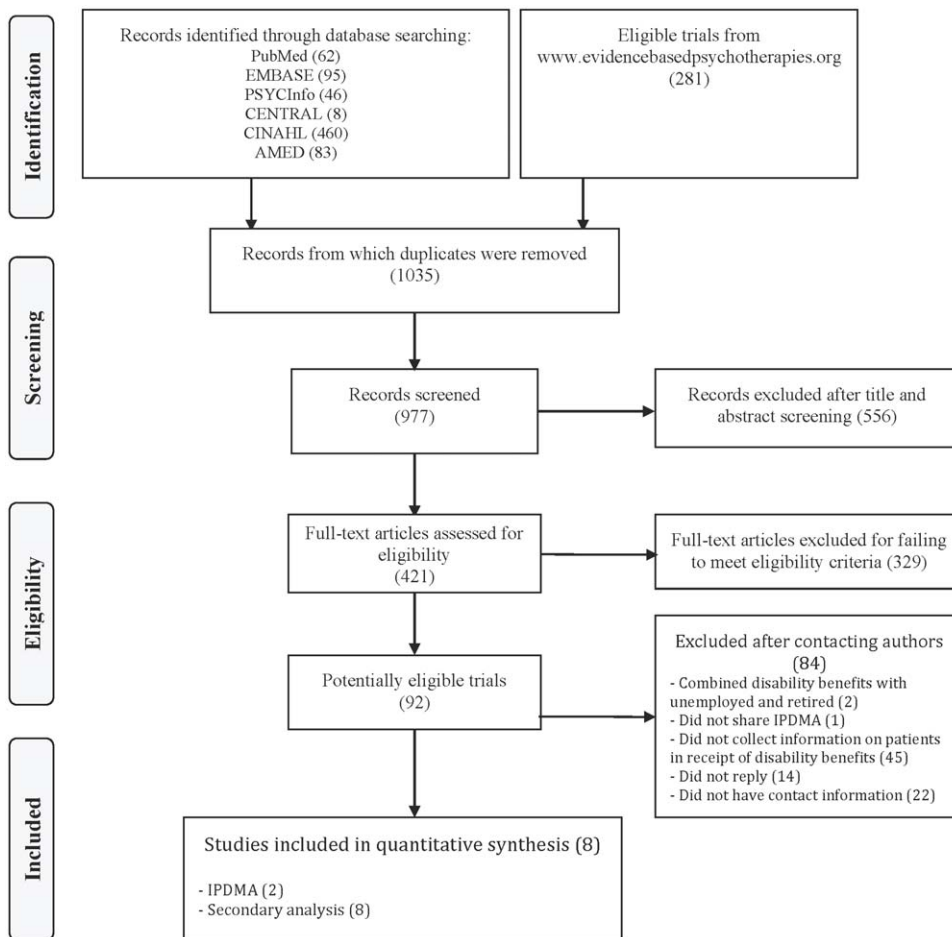


Figure 1. Flow chart of study eligibility.
doi:10.1371/journal.pone.0050202.g001

instrument and B represented the alternative instrument, L_A and L_B represent the lower range of instrument A and B respectively, and R_A and R_B represented the ranges for instruments A and B respectively [39].

We examined heterogeneity using both a chi-squared test and the I^2 statistic [40]. Heterogeneity defined by an I^2 of 0% to 40% was interpreted as 'might not be important', 30% to 60% as 'moderate heterogeneity', 50% to 90% as 'substantial heterogeneity', and 75% to 100% as 'considerable heterogeneity' [40]. We generated the following a priori hypotheses to explain variability between studies in our secondary analyses: studies using in-person CBT will have greater effects than studies using computer administered-CBT, and studies with high risk of bias will demonstrate larger effects compared to studies with low risk of bias.

We performed analyses using SPSS version 20 and the Cochrane Collaboration Review Manager software (RevMan version 5.1.2).

Results

Study selection

We screened 977 citations and retrieved 421 articles in full text; 329 studies did not meet inclusion criteria and 92 trials were deemed eligible. The kappa (95% CI) chance-corrected agreement

on assessing full text eligibility was 0.74 (0.66 to 0.81), representing substantial agreement.

After establishing author contact for 56 of the 70 trials for which we acquired contact information, we found that 45 trials did not have an eligibility criterion based on disability benefit status or collect information on disability status, 6 trials did not enrol any patients in receipt of disability benefits, and 5 trials enrolled some patients in receipt of disability benefits. Authors of 4 of the 5 trials that included patients in receipt of disability benefits agreed to provide individual patient data. Two of these trials combined patients who were disabled with unemployed and retired individuals and information specific to receipt of disability benefits were uncertain; these trials were therefore excluded from our IPDMA. Our primary analysis consisted of the 2 remaining trials that included some patients in receipt of disability benefits [41,42], and our secondary analyses consisted of 8 trials, i.e., 6 trials that did not enrol any patients in receipt of disability benefits [43,44,45,46,47,48], and 2 trials that included some patients in receipt of disability benefits (Figure 1) [41,42].

Study characteristics

Seven studies were parallel group RCTs [41,42,43,44,45,46,47], and one was a cluster RCT [48]. Table 1 describes the characteristics of the 8 eligible trials, and Table 2 provides details regarding their interventions.

Table 1. Characteristics of studies.

Study	Country	Sample size		Age (mean ± SD)		Patient Population	Treatment group	Control group	Depression outcomes reported	Included patients on disability benefits
		CBT	Control	CBT	Control					
De Graaf 2009 [41]	The Netherlands	100	103	45.2±10.9	45.1±12.2	Depression	CBT+TAU	TAU	BDI-II	Yes
Dozois 2009 [42]	Canada	25	23	NR	NR	Depression	CBT+pharmaco-therapy	Pharmaco-therapy alone	BDI-II; HRSD	Yes
Naeem 2011 [47]	Pakistan	17	17	32.35±8.9	33.64±1.0	Depression	CBT+pharmaco-therapy	Pharmaco-therapy alone	HADS; BSI	No
Faramarzi 2007 [43]	Iran	42	40	28.3±3.8	28.4±5.3	Depression in infertile women	CBT	Minimal or no treatment	BDI-II	No
Hollon 1992 [44]	USA	25	57	NR	NR	Nonpsychotic, nonbipolar depressed outpatients	CBT+pharmaco-therapy	Pharmaco-therapy alone	BDI; HRSD	No
Miranda 2003 [45]	USA	90	89	29.8±7.9	29.5±9.1	Depression in predominantly low-income young minority women	CBT	TAU	HRSD	No
Misri 2004 [46]	Canada	19	16	29.5±3.3	30.8±5.9	Postpartum depression	CBT+pharmaco-therapy	Pharmaco-therapy alone	HRSD	No
Rahman 2008 [48]	Pakistan	463	440	26.5±5.2	27.0±5.1	Perinatal depression	CBT	TAU	HRSD	No

CBT – Cognitive Behavioural Therapy; TAU – Treatment As Usual; SD – Standard deviation; NR – Not reported; BDI-II – Beck Depression Inventory-II; HRSD – Hamilton Rating Scale for Depression; HADS – Hospital Anxiety and Depression Scale; BSI – Bradford Somatic Inventory.
doi:10.1371/journal.pone.0050202.t001

Table 2. CBT details from studies.

Study	Mode of administration of CBT	Duration of CBT per visit	Frequency of CBT	Total duration of CBT	Clinical background of the individuals administering CBT	Was there a standardized program or certification process that CBT providers have undergone or had to undergo?
De Graaf 2009 [41]	Computer/internet based CBT	30 minutes	1 per week	9 weeks	Not reported	Not reported
Dozois 2009 [42]	In-person individualized CBT	1 hour	1 per week	15 weeks	Master's level therapist	Not reported
Naeem 2011 [47]	In-person individualized CBT	Not reported	1 to 2 sessions per week	9 weeks	Psychiatrist; psychology graduates	Not reported
Faramarzi 2007 [43]	In-person group CBT	2 hours	1 per week	10 weeks	Psychologist	Not reported
Hollon 1992 [44]	In-person individualized CBT	50 minutes	2 in the first 4 weeks, 1 or 2 in the next 4 weeks, and 1 in the last weeks	12 weeks	Psychologist; social worker	Not reported
Miranda 2003 [45]	In-person individualized CBT	Not reported	1 per week	8 weeks	Psychologist; psychotherapist	Not reported
Misri 2004 [46]	In-person individualized CBT	1 hour	1 per week	12 weeks	Psychologist	Not reported
Rahman 2008 [48]	In-person individualized CBT	Not reported	4 in 1st month, 3 in 2nd month, and 1 per month for next 9 months	11 weeks	Lady health workers	Not reported

CBT –Cognitive Behavioural Therapy.
doi:10.1371/journal.pone.0050202.t002

Risk of bias within studies

Protection against bias was generally poor (Figure 2). All 8 trials reported loss to follow-up (LTFU), ranging from 4% to 40%. Four trials excluded those LTFU and performed a complete case analysis [41,42,43,48], 2 used the last observation carried forward [44,46], 1 used multiple imputation (56), and 1 did not report an approach [47].

IPDMA

Two trials including data on patients receiving disability benefits enrolled a total of 227 patients; 34 in receipt of disability benefits and 193 not receiving disability benefits. The mean (SD) baseline BDI-II score for patients with disability benefits was 32.9 (±8.55) and for patients not receiving disability benefits 26.9 (±7.9).

Pooled results from these 2 trials, including both those receiving and not receiving disability benefits, suggested a possible benefit of CBT on depression (MD = -2.61; 95% CI = -5.28 to 0.07; p = 0.06, minimally important difference [MID] = 5), as did results from both the subgroup of patients in receipt of disability benefits (MD = -6.88; 95% CI = -14.06 to 0.31), and patients not receiving disability benefits (MD = -2.22; 95% CI = -5.07 to 0.63). Results suggested a possible larger effect on reducing depression in those receiving versus not receiving disability benefits, though the confidence interval includes a small reduction in benefit in those receiving benefits (MD = -4.46; 95% CI = -12.21 to 3.30; p = 0.26; MID = 5).

Secondary analyses

There were no significant differences in the effect of CBT on depression among patients not in receipt of disability benefits across studies that enrolled patients receiving disability benefits and studies that did not (p = 0.26) (Figure S1). There were no significant differences in the effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT

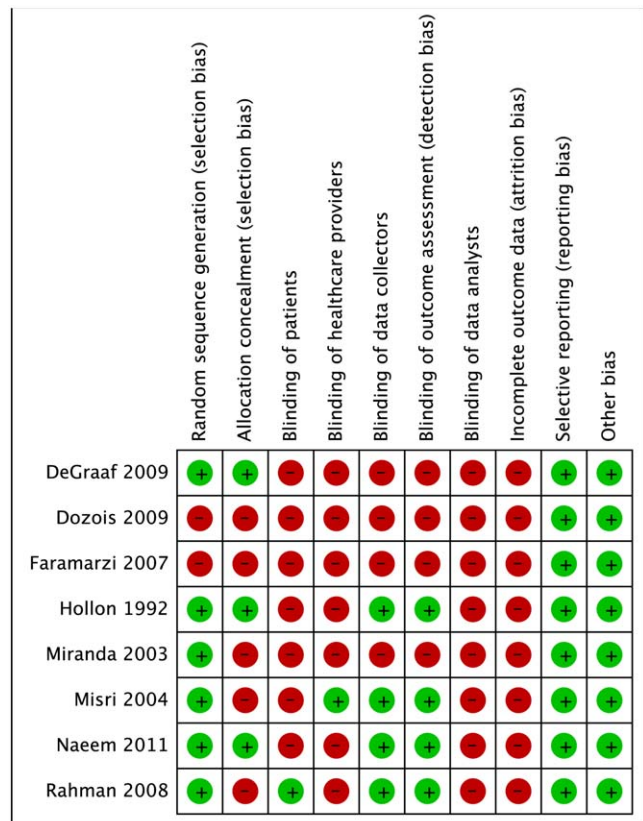


Figure 2. Risk of Bias within studies. ‘+’ denotes low risk of bias, and ‘-’ represents high risk of bias.
doi:10.1371/journal.pone.0050202.g002

plus pharmacotherapy versus pharmacotherapy alone ($p = 0.94$) (Figure S2). There were no significant differences in the effect of CBT on depression within patients not receiving disability benefits in studies comparing CBT versus TAU/standard care ($p = 0.59$) (Figure S3). Our *a priori* subgroup hypotheses failed to explain the heterogeneity observed in our secondary analyses.

Discussion

Summary of evidence

This is the first systematic review comparing the effect of receiving disability benefits on depression following treatment with CBT. We failed to find differences in the effect of CBT on depression between patients receiving disability benefits and patients not receiving disability benefits. The results suggest a possible greater effect in those receiving disability benefits (-4.46 BDI units in which the minimally important difference is 5), and the boundaries of the confidence interval suggest that if there is a decrement in benefit, that decrement is small (no greater than 3.30 BDI-II units). Nevertheless, these data come from only 34 patients receiving disability benefits, so that any inferences regarding relative effect in the two populations are very weak.

The strengths of our review include a comprehensive and transparent search strategy, independent and duplicate eligibility assessment, use of the most commonly reported instrument with the most established reliability and validity (BDI-II) for our pooled analysis, and use of individual patient data from eligible trials, allowing adjustment for potential confounding predictors. We also ensured rigorous data abstraction by using detailed written instructions, conducting formal calibration exercises, conducting in duplicate, and implementing a consensus approach to resolve disagreement. We contacted authors to verify whether they enrolled patients in receipt of disability benefits and achieved an 80% response rate among trials for which we were able to acquire author contact information.

Although no prior reviews have explored the effect of CBT in patients receiving disability benefits, reviews have explored the effect of compensation in other patient populations. A 2005 systematic review found that the presence of compensation was associated with worse outcome (combination of functional, quality of life, pain and patient satisfaction outcome that was rated as satisfactory or unsatisfactory by review investigators) after surgery [29]. This was consistent with findings from systematic reviews regarding chronic pain and closed-head injuries [49,50], which reported a significant effect between compensation and poor outcome. This indirect evidence, however, does not address the relative effect of interventions in the populations (one may have poorer outcomes, but still have larger treatment effects if results without treatment are very poor). In the two trials we examined, patients in receipt of disability benefits had a greater severity of depression than those who were not receiving disability benefits (baseline BDI-II of 32.9 versus 26.9). Although a prior review reported that the effectiveness of CBT was reduced in patients with severe depression compared to those with mild to moderate depression [51], we found no suggestion of a smaller effect of CBT in patients receiving disability benefits.

Limitations

Our study has limitations. First, our IPDMA is based on only 34 patients in receipt of disability benefits and 193 patients not receiving disability benefits. The extent to which findings from this small sample will generalize to a wide population of individuals in

receipt of benefits is uncertain. Second, our secondary analyses showed substantial heterogeneity within subgroups of patients not receiving disability benefits, which could not be explained by our *a priori* hypotheses. Possible explanatory factors that we were unable to explore due to limitations in the reporting of trials include baseline severity of depression, duration of depression, frequency of CBT, and experience of CBT providers. Third, none of the trials evaluated the effect of CBT on return to work (RTW), a critical outcome for patients receiving disability benefits and for insurers providing benefits. It remains possible that CBT may improve BDI-II scores, but may not have any effect on claim resolution or RTW. Future trials should include these outcomes in order to ascertain a BDI-II threshold that is associated with RTW and claim resolution.

Conclusions

If the use of CBT to manage depression among patients receiving disability benefits was less effective than in patients not receiving disability benefits, clinicians and payers might reasonably choose alternative treatment strategies (e.g. pharmacotherapy, other psychotherapies or a combination of both). The limited evidence available, however, provides no support for this hypothesis and suggests that, for the time being, CBT should continue as a recommended approach for addressing depression in patients receiving disability benefits. Secure inference will, however, only be possible after the conduct of much larger comparative trials, conducted with low risk of bias and in collaboration with insurers.

Supporting Information

Figure S1 Effect of cognitive behavioural therapy in patients not receiving disability benefits in studies including patients receiving disability benefits versus those that did not.

(TIF)

Figure S2 Effect of cognitive behavioural therapy on depression within patients not receiving disability benefits in studies comparing CBT plus pharmacotherapy versus pharmacotherapy alone.

(TIF)

Figure S3 Effect of cognitive behavioural therapy on depression within patients not receiving disability benefits in studies comparing CBT versus TAU/standard care.

(TIF)

Checklist S1 PRISMA Checklist.

(DOC)

Acknowledgments

We thank Dr. Randi McCabe for her assistance in reviewing the eligibility of the treatment in trials.

Author Contributions

Conceived and designed the experiments: SEJWB SDW GHG. Performed the experiments: SE LM SH MK WT AC-L RC. Analyzed the data: SE JWB SDW DH-A GHG. Contributed reagents/materials/analysis tools: LED DD PB. Wrote the paper: SE. Knowledge users and reviewed the manuscript: IP-S MB.

References

- Hirschfeld R, Montgomery SA, Keller MB, Kasper S, Schatzberg AF, et al. (2000) Social functioning in depression: a review. *J Clin Psychiatry* 61: 268–275.
- Stewart WF, Ricci JA, Chee E, Hahn SR, Morganstein D (2003) Cost of lost productive work time among US workers with depression. *JAMA* 289: 3135–3144.
- World Health Organization (WHO) (2002) *Mental Health: New understanding, New hope*. Geneva: WHO.
- Fournier J, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, et al. (2010) Antidepressant drug effects and depression severity: A patient-level meta-analysis. *JAMA* 303: 47–53.
- NICE (2009) *Depression: Treatment management of depression in adults*. London: National Institute for Clinical Excellence.
- Olfson M, Marcus SC (2009) National patterns in antidepressant medication treatment. *Arch Gen Psychiatry* 8: 848–856.
- Dobson KS (2001) *Handbook of cognitive-behavioral therapies*. 2nd Edition. New York; London: Guilford Press.
- Almeida A, Lotufo Neto F (2003) Cognitive-Behavioral therapy in prevention of depression relapses and recurrences: a review. *Rev Bras Psiquiatr* 25: 239–244.
- Bortolotti B, Menchetti M, Bellini F, Montaguti MB, Berardi D (2008) Psychological interventions for major depression in primary care: a meta-analytic review of randomized controlled trials. *General Hospital Psychiatry* 30: 293–302.
- Churchill R, Hunot V, Corney R, Knapp M, McGuire H, et al. (2001) A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technology Assessment* 5: 1–173.
- Cuijpers P, van Straten A, Warmerdam L, Andersson G (2008) Psychological treatment of depression: A meta-analytic database of randomized studies. *BMC Psychiatry* 8.
- Cuijpers P, van Straten A, van Schaik A, Andersson G (2009) Psychological treatment of depression in primary care: A meta-analysis. *Journal of General Practice* 59: e51–e60.
- Cuijpers P, Smit F, Bohlmeijer E, Hollon SD, Andersson G (2010) Efficacy of cognitive behavioural therapy and other psychological treatments for adult depression: a meta-analytic study of publication bias. *Br J Psychiatry* 196: 173–178.
- Gloaguen V, Cottraux J, Cucherat M, Blackburn IM (1998) A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders* 49: 59–72.
- Harrington R, Whittaker J, Shoebridge P, Campbell F (1998) Systematic review of efficacy of cognitive behaviour therapies in childhood and adolescent depressive disorder. *BMJ* 316: 1559–1563.
- Kalenthaler E, Parry G, Beverley C, Ferriter M (2008) Computerised cognitive-behavioural therapy for depression: systematic review. *Br J Psychiatry* 193: 181–184.
- Oei TP, Dingle G (2008) The effectiveness of group cognitive behaviour therapy for unipolar depressive disorders. *J Affect Disorder* 107: 5–21.
- Peng XD, Huang CQ, Chen LJ, Lu ZC (2009) Cognitive behavioural therapy and reminiscence techniques for the treatment of depression in the elderly: a systematic review. *J Int Med Res* 37: 975–982.
- Wampold BE, Minami T, Baskin TW, Callen TS (2002) A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. *Journal of Affective Disorders* 68: 159–165.
- Kouzis AC, Eaton WW (1997) Psychopathology and the development of disability. *Social and Psychiatry and Psychiatric Epidemiology* 32: 379–386.
- Kouzis AC (2000) Psychopathology and the initiation of disability payments. *Psychiatr Serv* 51: 908–913.
- Salkever DS, Goldman H, Purushothaman M, Shinogle J (2000) Disability management, employee health and fringe benefits, and long-term-disability claims for mental disorders: an empirical exploration. *The Milbank Quarterly* 78: 79–113.
- Dewa C, Hoch JS, Goering P, Lin E, Patterson M (2003) Use of antidepressants among Canadian workers receiving depression-related short-term disability benefits. *Psychiatr Serv* 54: 724–729.
- Ebrahim S, Guyatt GH, Walter SD, Heels-Ansdell D, Hanna S, et al. (2012) Association of psychotherapy with disability benefit claim closure among patients disabled due to depression (under review). Hamilton: McMaster University.
- (2003) *Roadblocks in Cognitive-Behavioral Therapy: Transforming Challenges into Opportunities for Change*; Leahy R, editor. New York: The Guilford Press.
- Weighill VE (1983) "Compensation neurosis": a review of the literature. *J Psychosom Res* 27: 97–104.
- Young JN, Shaffrey CI, Laws ER Jr, Lovell LR (1997) Lumbar disc surgery in a fixed compensation population: a model for influence of secondary gain on surgical outcome. *Surg Neurol* 48: 552–558.
- Olney JR, Quenzer DE, Makowsky M (1999) Contested claims in carpal tunnel surgery: outcome study of worker's compensation factors. *Iowa Orthop J* 19: 111–121.
- Harris I, Mulford J, Solomon M, van Gelder JM, Young J (2005) Association between compensation status and outcome after surgery: a meta-analysis. *JAMA* 293: 1644–1652.
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 62: 1006–1012.
- American Psychiatry Association (2010) *Diagnostic and Statistical Manual*. Arlington, VA.
- Cuijpers P (2011) *Psychotherapy randomized controlled and comparative trials*. Amsterdam: vrije Universiteit.
- Landis J, Koch GG, Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174.
- The Cochrane Collaboration (2011) Chapter 8: Risk of Bias. In: Higgins JPT, Green S, editor. *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.10). Oxford, UK: Cochrane Collaboration.
- Akl EA, Sun X, Busse JW, Johnston BC, Briel M, et al. (2011) Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. *J Clin Epidemiol* Dec 24 (Eub ahead of print).
- (1999) Variable intercepts: fixed or random parameters? In: Snijders T, Bosker RJ, editor. *Multilevel Analysis An introduction to basic and advanced multilevel modeling*. London: SAGE Publications Ltd.
- Tendal B, Nüesch E, Higgins JP, Jüni P, Gøtzsche PC (2011) Multiplicity of data in trial reports and the reliability of meta-analyses: empirical study. *BMJ* 343.
- Riley R, Simmonds MC, Look MP (2007) Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J Clin Epi* 60: 431–439.
- Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH (2012) Pooling health-related quality of life outcomes in meta-analysis—a tutorial and review of methods for enhancing interpretability. *Research Synthesis Methods* In press.
- The Cochrane Collaboration (2009) 9.5.2 Identifying and measuring heterogeneity. In: Higgins JPT, Thompson SG, editor. *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.02). Oxford, UK: Cochrane Collaboration.
- de Graaf LE, Gerhards SAH, Arntz A, Riper H, Metsmakers JFM, et al. (2009) Clinical effectiveness of online computerised cognitive behavioural therapy without support for depression in primary care: randomised trial. *The British Journal of Psychiatry* 195: 73–80.
- Dozois DJA, Bieling PJ, Patelis-Siotis I, Hoar L, Chudzik S, et al. (2009) Changes in Self-Schema Structure in Cognitive Therapy for Major Depressive Disorder: A Randomized Clinical Trial. *Journal of Consulting and Clinical Psychology* 77: 1078–1088.
- Faramarzi M, Alipor A, Esmaelzadeh S, Kheirkhah F, Poladi K, et al. (2008) Treatment of depression and anxiety in infertile women: Cognitive behavioral therapy versus fluoxetine. *Journal of Affective Disorders* 108: 159–164.
- Hollon SD, DeRubeis RJ, Evans MD, Wiemer MJ, Garvey MJ, et al. (1992) Cognitive therapy and pharmacotherapy for depression. *Arch Gen Psychiatry* 49: 774–781.
- Miranda J, Chung JY, Green BL, Krupnick J, Siddique J, et al. (2003) Treating depression in low-income young minority women: A randomized controlled trial. *JAMA* 290: 57–65.
- Misri S, Reebye P, Corral M, Milis L (2004) The Use of Paroxetine and Cognitive-Behavioral Therapy in Postpartum Depression and Anxiety: A Randomized Controlled Trial. *Journal of Clinical Psychiatry* 65: 1236–1241.
- Nacem F, Waheed W, Gobbi M, Ayub M, Kingdon D (2011) Preliminary Evaluation of Culturally Sensitive CBT for Depression in Pakistan: Findings from Developing Culturally-Sensitive CBT Project (DCCP). *Behavioural and Cognitive Psychotherapy* 39: 165–173.
- Rahman A, Malik A, Sikandar S, Roberts C, Creed F (2008) Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *The Lancet* 372: 902–909.
- Binder LM, Rohling ML (1996) Money matters: a metaanalytic review of the effects of financial incentives on recovery after closed-head injury. *Am J Psychiatry* 153: 7–10.
- Rohling ML, Binder LM, Langhinrichsen-Rohling J (1995) Money matters: a meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. *Health Psychol* 14: 537–547.
- Haby MM, Donnelly M, Corry J, Vos T (2006) Cognitive behavioural therapy for depression, panic disorder and generalized anxiety disorder: a meta-regression of factors that may predict outcome. *Aust N Z J Psychiatry* 40: 9–19.

APPENDIX C

Description of all variables in administrative database

Variable	Definition or description
Age at disability	Age of claimant at disability
Analyst Name	Name of current case manager
Budget Code	Budget code from where expense was paid from, i.e. if it was paid under a claims case management plan or a rehab (health management) plan
Budget Type	Health Management or Claims
Calculated Definition Change Date	This is the true Change of Definition Date (COD)
Certificate	Certificate number of claimant
City	City where claimant resides
Claim Office	Office where the claim is currently managed
Claim Status	Actively managed or permanent
Control Number	Unique identifier of the claim
Cost Centre	Cost centre number
CPP/QPP Offset Amount	Amount of Canada Pension Plan/ Quebec Pension Plan (CPP/QPP) offset
CPP/QPP Offset Type	Type of Canada Pension Plan/ Quebec Pension Plan (CPP/QPP) offset, if applicable
Date claim was approved	Date the abilities case managers approved the claim
Date claim was received	Date insurer received first part of the claim
Date of Birth	Date of birth of claimant
Date of claim resolution	Effective date of resolution/closure or decline. This field indicates a resolution if the Approval Date is populated. If Approval Date is not populated, indicates a decline decision.
Date of Disability	Date of disability
Disability Code	Diagnostic code used
Entry Date	Date services was paid by the insurer
Expense Amount	Amount of invoice paid to service provider
Expense Description	Detailed expense category description
Expense Type Code	Detailed expense category code
Expiry Date	Date benefit would expire according to the contract. In most of the contracts, it is age 65 but may differ according to the contract.
File Owner	Code for current management (or most recent) of the claim
First Payment Date	Date plan member is eligible for payment, if benefit is approved
Forecast Date	Date insurer believes that the claim will be managed until
Funding Category	Refund or non-refund funding
Funding Type	Non-refund funded policy includes claims where the payments of the

services and treatments are covered by the insurer. Refund funded policy includes claims where the payments of the services and treatments were shared by the insurer and the plan sponsor (e.g. the employer), and ASO funded policy consists of claims where the plan sponsor pays for the services and treatments.

Gender	Gender of claimant
Gross Benefit	Calculated gross benefit amount based on salary and contractual benefit amount. For LTD, this is a monthly amount. For STD, this is a weekly amount.
HMC Acceptance Date	The date the claim was accepted into a HMC program.
HMC Name	If the claim was in a health management program, it indicates the name of the HMC (consultant).
HMC Referral Date	The date the claim was referred by the case manager to a health management program
HMC Status	Status of the HMC
HMC Team	If the claim was in a health management program, it indicates the team of the HMC (consultant).
ICD-10 Primary Diagnosis Code	ICD-10 code
ICD-10 Primary diagnosis name	Description of primary diagnosis
ICD-10 Secondary Diagnosis Code	ICD-10 secondary diagnosis code
ICD-10 Secondary Diagnosis Name	Description of secondary diagnosis
Input ID	User who entered this expense
Internal/External	Indicates if expense was external service or internal. All CBT expenses are external
Last Payment Date	Last date the claim was paid, if resolved or date they have approved payment to before further review if claim is still active.
Major Disability Category	Disability category: Accident, cancer, circulatory, psychological, musculo-skeletal, pregnancy, nervous
Market Segment	This is the segment of market these claims belong to - determined by Plan Sponsor
Maximum claim benefit period	STD max benefit date; blank in LTD
Net Benefit	Gross benefit less offsets
Number of Units	Will all be zero for external expenses
Occupation	Occupation group of claimant
Occupation Industry	Occupation industry of the claimant
Payment Status	Indicates payment status of invoice
PC Firm Likely Code	Indicates if CPP/QPP offset is being offset 1 = actual offset, 2 = estimated. Remaining codes indicate not offset from benefit
Plan ID	ID number of plan

PM Code	Current portfolio management code on the claim or if claim now closed, was the last code on the claim
PM Code Date	Date PM code was set by the case manager
Policy	Policy name
Policy Status	Active or terminated
Policy Type	High level funding of the plan sponsor, Non ASO indicated an insured arrangement, ASO is administrative services only, and Government
Policyholder Name	Name of policyholder
Postal Code	Postal code of claimant
Provider ID	Provider ID
Provider Name	Provider of the particular services if available
Province	Province where claimant resides
Received psychotherapy	Claimant received psychotherapy
Referral for an Independent Medical Evaluation (IME)	Claimant referral for a detail assessment by an independent medical examiner
Referral for Transferrable Skills Analysis (TSA)	Assistance provided by an external provider to assist the claimant in preparing and looking for a work position outside of their employer's work environment.
Referral for Academic Upgrading/ Skills Retraining	Claimant referral for education upgrading or skills retraining through an educational program
Referral to a Work hardening program	Claimant was referral to a multidisciplinary work hardening/conditioning facility that addresses the functional, physical behavioural, cognitive requirements of a job and provides real or simulated work to assist the claimant in the safe transition back to full time employment.
Received Massage Therapy/ Acupuncture/ Chiropractic treatment	Claimant received acupuncture, massage therapy or chiropractic treatment
Rehab Earnings Offset Amount	Amount of Rehab earnings offset
Rehab Earnings Offset Type	Rehab earnings offset type, if applicable
Retirement Pension (RP) Firm Likely Code	Indicates if RP offset is being offset 1 = actual offset, 2 = estimated. Remaining codes indicate not offset from benefit
Retirement Pension Offset Amount	Amount of retirement pension offset
Retirement Pension Type	Retirement pension offset type, if applicable
RH Firm Likely Code	Rehab firm likely code; indicates if RH offset is being offset 1 = actual offset, 2 = estimated. Remaining codes indicate not offset from benefit
Salary	Salary of claimant
Service/Therapy Date	Date service/therapy was performed
Service/Therapy Type	Type of Service/therapy: psychotherapy, independent medical evaluation, functional capacity evaluation, TSA, Academic Upgrading/ Skills Retraining, Massage Therapy/ Acupuncture/ Chiropractic treatment, Work Hardening program
Tax code	Indicates taxability of claim - yes, no or at source.
Team Name	Team that is currently managing the claim

Total Reserve amount	Reserves held on Insured LTD claims only.
Type of claim resolution	Resolution/closure/decline reason
Type of coverage	LTD or STD
