ANXIETY PREVALENCE FOLLOWING TRAUMATIC BRAIN INJURY AND DETECTING FEIGNED IMPAIRMENT WITH THE DENVER ATTENTION TEST

by

Emma Johnson

A thesis submitted to the University of Birmingham in part completion of the degree of
DOCTOR OF CLINICAL PSYCHOLOGY

Centre for Applied Psychology
School of Psychology
College of Life and Environmental Sciences
University of Birmingham

July 2024

Thesis Overview

This thesis consists of four chapters. Chapter one is a meta-analytic review examining the prevalence data for generated anxiety disorder (GAD) and clinically significant cases of anxiety in adults with non-penetrating traumatic brain injury (TBI). The review is an update of a previous meta-analysis published in 2014 and explores the past decade of research in this area. Data from 33 studies and a combined sample of 12,063 participants are reported. The impact of time since injury, assessment method, injury severity, and psychiatric history are each examined. Findings reveal that following TBI, GAD and clinically significant anxiety are almost four and six times respectively, more prevalent than in the general population. These findings underscore the importance of assessing mental health needs following TBI.

Chapter two is an empirical study which explores the utility of a newly developed performance validity measure (PVT) called the Denver attention Test (DAT). A simulator study design is used to examine the DATs ability to discriminate between a group of participants who are instructed to do their best and another who are instructed to feign cognitive impairment. Results find that the DAT is a rapid, easy to administer PVT that provides a robust measure of performance validity. It demonstrates excellent ability to detect feigned cognitive impairment in a simulator sample, though further investigation is needed with clinical groups.

Chapters three and four are press releases for the meta-analytic review and empirical study. This research hopes to bring new insights to the field of neuropsychology.

Dedications

This thesis is dedicated to all those who have supported, encouraged and believed in me along my journey into Clinical Psychology. This thesis is especially dedicated to my mum, Sue. Her endless love and support have kept me buoyant on tougher days, from patiently proofreading hours of my work over the years, to bringing me cups of tea and slices of homemade cake to keep me going. My stepdad Mick has also been a rock, and both mum and Mick have frequently stepped in to make other bits of my life a little bit easier whilst I have been busy spinning plates of the doctorate. My sister Amy and best friend Kate have also been integral to this process by being my go-to safe spaces to laugh, vent, cry, and cheer me on. So too have my nieces Violet and Olivia; although they are currently too young to understand this journey, their pure love and joy has always carried me through. A special mention also goes to my little dog Moritz, who has spent a significant part of his life curled up beside me at my desk, blissfully unaware of what any of this means or how much his companionship has bolstered me. I also dedicate this work those whose love and belief in me has spurred me on, even though they are sadly no longer with us: my dad, Bob and nanna, Brenda. I also dedicate this to my cohort of course mates who I have shared this unique journey with, and those of whom I have the pleasure of calling friends for life. Finally, I dedicate this thesis to my younger self who never imagined she was capable of such an achievement. Her courage and determination to tread difficult and unfamiliar territory in the face of adversity has led me to here and for that I'm forever grateful.

Acknowledgments

Firstly, I would like to thank the University of Birmingham for the opportunity to achieve my dream of becoming a Clinical Psychologist, and to all the staff and visiting lecturers who have contributed to my invaluable learning experience over the three years, as well as the staff behind the scenes who have helped everything run as smoothly as possible. I would like to thank my appraisal tutor Gary Law for overseeing my journey and ensuring I got the most from the experience. I would like to thank all my clinical supervisors to date who each have played an integral role in developing my clinical skills and ensuring placements were as rich in learning as possible. I would like to extend a huge thank you to my research supervisors at the Queen Elizabeth Hospital in Birmingham; Chris Jones, David Hacker, and Carl Krynicki, who have taught me so much about the world of neuropsychology. I am especially grateful to Chris for his wisdom and patience in teaching all things statistics and meta-analysis. The research team have made what is quite a daunting venture, an interesting and enjoyable one despite the setbacks we faced. I would also like to thank the extended research team, including Zoe Clowes and Emily Paton, for their valuable time and input on the project. Finally, I would like to express thanks to all those who participated in my research, for their generosity in time and willingness to engage with an hour of testing (and acting for some!) in the name of research.

Table of Contents

List of Figures

List of Tables

Chapter 1: Prevalence of Anxiety Following Adult Traumatic Brain Injury: A Meta-Analysis Update Comparing Measures, Samples and Post-injury Intervals

Abstract

Background Anxiety is a common psychiatric condition following traumatic brain injury (TBI), but the reported prevalence rates are inconsistent and make meaningful interpretation of them difficult. This meta-analysis sought to identify how methodological variables and sample characteristics may impact on prevalence rates. This study is an update of a meta-analysis by Osborn et al. and reviews literature published from May 2014.

Methods Data from 33 studies that reported prevalence data for generalised anxiety disorder (GAD) or clinically significant cases of anxiety from adults with non-penetrating TBI were analysed. A combined sample of 12,063 participants were included. The impact of time since injury, assessment method, injury severity and psychiatric history were each examined.

Results The overall prevalence rate for GAD was 15% and 23% for anxiety. Prevalence rates for both conditions varied across subgroup factors. GAD was influenced by factors including time since injury (4%-34%), assessment method (3%-26%), and psychiatric history (2%-21%), but unaffected by injury severity (8%-16%). Anxiety remained consistent across factors including time since injury (19%-29%), assessment method (22%-24%), and psychiatric history (22%-30%). The impact of injury severity on anxiety prevalence was unclear (16%-71%).

Conclusion The past decade of research indicates that following TBI, GAD and clinically significant anxiety are almost four and six times respectively, more prevalent than in the general population. As such, assessment of mental health needs following TBI is essential. Findings underscore the significant role clinicians can provide in supporting individuals to cope with the psychological impact of their injuries following TBI.

Introduction

Traumatic brain injury (TBI) results from damage caused to the brain from an external force, leading to changes in cognitive and/or behavioural functioning. It is a major worldwide source of health loss and disability, with an estimated global annual incidence between 27 and 69 million (Williamson & Venkatakrishna, 2024). The rising worldwide incidence of TBI is largely associated with the increased use of motor vehicles, either as an occupant of a vehicle or as another road user, such as a pedestrian, cyclist or motorcyclist (Roozenbeek et al., 2013). However, in recent decades an aging population in higher income countries has led to falls surpassing road traffic incidents as the leading cause of TBI (Roozenbeek et al., 2013). The impact of TBI can vary significantly, from shorter term impairment to long-lasting or even permanent changes, depending upon the severity and circumstances of the trauma (Roebuck-Spencer & Cernich, 2014). Injury severity is assessed through clinical examination and neuroimaging techniques to determine whether it can be classified as mild, moderate, or severe (Savitsky et al., 2016). TBI is typically associated with deficits of memory, attention, processing speed, and executive functioning (Stocchetti & Zanier, 2016).

Psychiatric problems and disorders are also commonly observed following TBI, such as depression, anxiety, and psychosis, as well as maladaptive behaviours such as substance misuse (Albrecht et al., 2020; Zgaljardic et al., 2015). The duration of psychiatric symptoms can vary across individuals; some may only experience symptoms during the acute phase post-injury, or for others their symptoms may persist. Long-term psychiatric disorders, together with the cognitive and physical sequelae of TBI, can pose significant challenges for patients and their caregivers by interfering with rehabilitation participation and functional community independence (Zgaljardic et al., 2015). They have also been associated with concurrent unemployment, pain, a deterioration in quality of life, and maladaptive coping skills (Gould et al., 2011). Where an individual may experience substantial changes to their

self-identity, a detrimental impact on psychological well-being may endure (Villa et al., 2020). Furthermore, personality change following TBI can have a dramatic impact on personal relationships, moving couples into unfamiliar dynamics and roles. This may reduce emotional wellbeing and impact relationship quality and satisfaction (Bodley-Scott & Riley, 2015). A perceived need for support around stress and emotional disorders is frequently expressed by TBI patients (Andelic et al., 2014; Corrigan et al., 2004; Ruet et al., 2019).

According to the International Statistical Classification of Diseases and Related Health Problems (ICD-11; World Health Organization, 2019), anxiety is classified as an anticipation or fearfulness of future danger or misfortune which is accompanied by preoccupation, distress, or somatic symptoms of tension. The focus of which may be either internal or external. Anxiety disorders are reported to be the world's most common psychological disorders, with an estimated global prevalence of 4% (World Health Organization, 2023). Given the prevalence rate in the general population, the relationship between TBI and anxiety is complex and it can be difficult to ascertain pre- and post-morbid functioning (Zgaljardic et al., 2015). Whether the cognitive and psychiatric consequences of TBI arise from specific brain lesions, psychological reactions to the trauma, pre-existing psychiatric conditions, or a combination of these factors has been much debated (Scicutella, 2019). Furthermore, brain injuries frequently occur within traumatic or emotionally charged situations, such as motor vehicle collisions or assaults and therefore the context within which the injury was sustained may understandably elicit anxiety reactions (Zgaljardic et al., 2015). The emergence of an anxiety disorder following TBI has found to be strongly predictive of social, personal, and work dysfunction (Mallya et al., 2015). A prospective study looking at participation outcomes following moderate to severe brain injury (Wise et al., 2010), found that at one year post-injury, many individuals had reduced their engagement with leisure activities which could support with recovery following TBI. Participants were typically involved with more

sedentary and less social activities and a substantial portion of participants were reported to be dissatisfied with these changes.

Across the literature, prevalence rates for anxiety are inconsistent and can vary substantially (Ponsford et al., 2018). Whilst variability in prevalence rates across multiple estimates may undermine the value of each individual estimate, there may be numerous factors which contribute to this variation, including diagnostic criteria, assessment method, severity of TBI, and time since injury. To preserve the clinical utility of prevalence data, it is prudent to explore factors which may impact upon it. Regarding assessment method, anxiety may be diagnosed by clinicians through structured clinical interviews, or it may be inferred by meeting thresholds for clinically significant 'caseness' of anxiety via self-report methods that demonstrate good reliability and validity. The thresholds used to determine clinically significant anxiety reflect the extent to which an individual is affected by the anxiety, with chronic difficulties in daily functioning and well-being. However, as noted by Osborn et al. (2016), self-report measures of anxiety sometimes include symptoms that could be related to TBI, such as impaired concentration and memory or sleep disturbances. As a result, these questionnaires may overestimate the prevalence of anxiety cases by including TBI-related symptoms. In a study exploring whether mode of administration can affect results (Moum, 1998), an identical scale of anxiety and depression was administered to a sample of 13,850 Norwegian adults, either by experienced interviewers or via self-administered questionnaires. Two to three times as many 'probable cases' of psychological distress were identified by the self-report method compared to interview. The authors suggest that interviews may be more vulnerable to socially desirable responding compared with self-report methods. Capturing anxiety prevalence though a range of self-report measures and clinical interviews may shed light on the variability of prevalence across assessment methods.

With reference to time since injury, some research suggests that psychiatric disorders are most likely to emerge within the first-year post-injury, although a delayed onset may be associated with severe injury (Ponsford et al., 2018). It has also been suggested that mild TBI may predispose the brain towards heightened fear learning post-injury, via a molecular mechanism (Reger et al., 2012). It is difficult to determine how long anxiety disorders may persist following TBI, though long-term studies suggest they can be enduring (Albicini & McKinlay, 2018; Dahm & Ponsford, 2015; Stenberg et al., 2022). A prospective, long-term study by Ruet et al. (2019) examined outcomes following severe TBI and found that around 25% of the sample had clinically significant anxiety at eight years post-injury. The presence of psychiatric history prior to TBI may also affect prevalence rates. A study by Gould et al. (2011) found that having pre-existing anxiety disorders was statistically predictive of developing post-injury anxiety. However, the number of individuals with anxiety diagnoses was found to have more than doubled following TBI, indicating that anxiety disorders are also likely to emerge post-injury across individuals without pre-existing anxiety.

A further factor to consider in understanding prevalence rates of anxiety is how comparable findings are to other trauma populations or similar patient groups. In a large study including 9,428 individuals with mild TBI, affective disorders including anxiety, depression, and adjustment disorders were observed in 23% of the mild TBI group. This contrasted with 14% in the control group at 12-month post-injury, which had been randomly selected and individually matched based on demographic data and medical comorbidities (Delmonico et al., 2022). Another study found that individuals with TBI are more likely than other trauma participants to report mental health difficulties and an increased need for mental health services (Ouellet et al., 2009). Further control group data is needed to contextualise these findings.

As well as being understood in the broad sense, the term anxiety can be used to denote specific clinical disorders. Generalised anxiety disorder (GAD) is common following TBI but its prevalence remains unclear, with varying estimates of between 1% and 27% (Ponsford et al., 2018). GAD is classified in the ICD-11 as marked and persistent symptoms of anxiety lasting at least several months. Symptoms involve general apprehension or excessive worry directed toward varying day-to-day events such as family, health, finances, or work. Additional symptoms may include muscular tension, subjective nervousness, difficulty maintaining concentration, irritability, or sleep disturbances (World Health Organization, 2019). The Generalised Anxiety Disorder 7 (GAD-7; Spitzer et al., 2006) is a self-report measure of symptoms of GAD over the past 14 days, which unlike clinical interviewing does not rely on prolonged symptom duration. When a score of 10 or more is applied as a cutoff to indicate the probable presence of GAD, the measure produces sensitivity of 89% and specificity of 82% (Williams, 2014). Post-traumatic stress disorder (PTSD) is also frequently cited following TBI with prevalence rates ranging from 2.6 to 36% (Van Praag et al., 2019). However, PTSD is classified as a stress disorder rather than an anxiety disorder within the ICD-11 and a trauma- and stressor-related disorder in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013), and therefore is not included within this review. Given the variability in presentations, it is important to study the prevalence of anxiety and GAD independently.

With appropriate and prolonged care, the consequences of TBI may be attenuated (Stocchetti & Zanier, 2016). It is therefore imperative that clinicians support patients to cope with both the psychological and physical consequences of their injury to improve long-term psychosocial outcomes (Draper et al., 2007; Schönberger et al., 2011). The aim of this paper is to explore the prevalence of GAD and clinically significant anxiety following TBI, by providing an update of a meta-analysis by Osborn et al. (2016). It will build upon previous

findings by including studies that have been published since May 2014. Given the significant variability in factors influencing anxiety prevalence, the impact of time since injury, assessment method, severity of injury, and psychiatric history will each be examined.

Methods

Identifying primary studies

Search of Electronic Databases

A systematic search of the literature was initially carried out on 27th July 2023 using PsycINFO, PubMed, Scopus, and Web of Science Core Collection. Search terms that were used to identify these studies are detailed in table 1 below. Search terms were applied to titles, abstracts, and Medical Subject Headings only, as full-text searches were too broad and yielded vast numbers of results. Subsequently, all types of anxiety were included within search terms to capture relevant papers.

Table 1

Electronic database search terms

Construct	Free Text Search Terms	Method of Search	Limits
Traumatic brain injury	"traumatic brain injur*" "TBI" "head injur*" "brain injur*" "brain damage" "head trauma" "craniocerebral trauma" "cranio-cerebral trauma" "cranio cerebral trauma"	Search terms within each construct (TBI and anxiety) were combined with *OR* whilst search terms between each construct were combined with *AND*. The * was used to search for all forms of the word that start with the same letters.	Peer reviewed articles 2014 - July 2023 Human studies English
Anxiety	"anxiety" "psychiatric diagnos*" "psychological sequelae" "affective disorder" "generalised anxiety disorder" "generalized anxiety disorder" "social anxiety disorder" "acute stress disorder" "post-traumatic stress disorder" "posttraumatic		

Construct	Free Text Search Terms	Method of Search	Limits
	stress disorder" "post traumatic stress disorder" "PTSD" "social phobia		

Inclusion and Exclusion Criteria

The full inclusion and exclusion criteria are described in Table 2.

Table 2

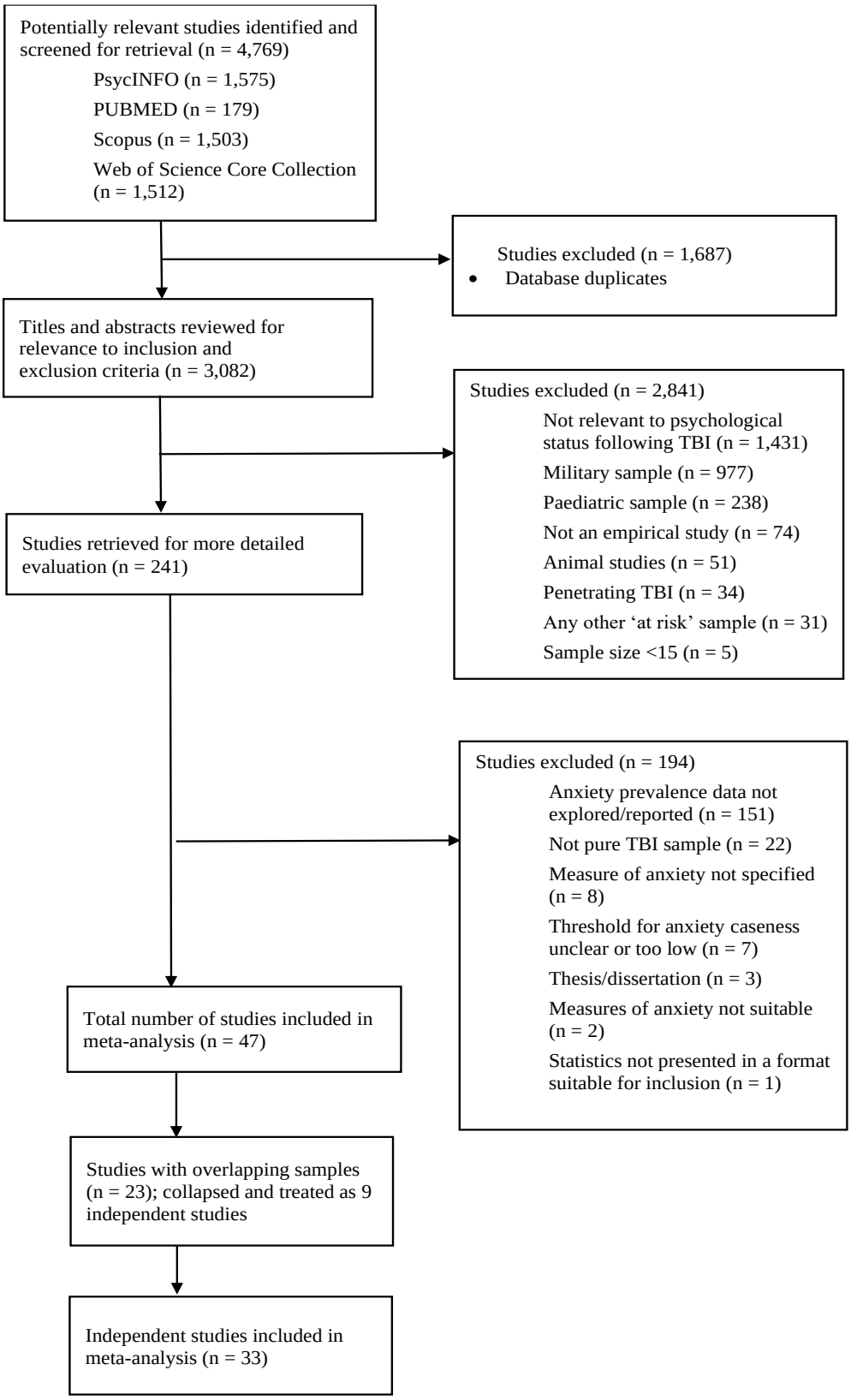
Inclusion and exclusion criteria

Inclusion/exclusion criteria	Justification
Inclusion	
It examined diagnoses of clinically significant anxiety or GAD following non-penetrating TBIs.	In accordance with the previous meta-analysis (Osborn et al., 2016).
The study population were adults (including studies that focus on adult populations but include participants from age 16 and above).	TBI assessment presents different challenges for children and young people and may not be generalisable to an adult population.
The prevalence of current diagnoses of GAD and/or clinically significant 'cases' of anxiety are assessed and reported using a common and specific measure of anxiety with a clinically validated cut-off level (excludes quality of life and general function measures).	To support validity and reliability of findings and to provide robust cut-offs for clinical caseness of anxiety or GAD.
Data were provided for a TBI sample (single sample) or both a TBI and control group (independent samples).	To avoid contamination of findings with non-TBI participants within TBI group scores.
It was published in a journal in English.	To enable a thorough review of research articles.
It reported original data.	To avoid duplication of data from original studies.
The sample size was >15.	Excludes very small studies (which are likely to report zero incident rates) and case studies.
Exclusion	
Highly specific or at-risk TBI populations, such as psychiatric patients, prisoners, military personnel, or victims of large-scale trauma/terrorism.	These studies may skew results towards higher levels of anxiety.
Only pre-treatment data were analysed if a study examined treatment efficacy.	It would be difficult to know if treatment has impacted anxiety scores.
Focus on current/post-injury anxiety, rather than dispositional/trait anxiety	Trait anxiety may obscure the impact of TBI.

The application of the inclusion and exclusion criteria to the results of the systematic search are presented in figure 1. The electronic search yielded a total of 4,769 articles, which resulted in 3,082 articles once duplicates were removed. The titles and abstracts of these articles were then screened using the criteria described in table 2. The three most common reasons for removal were: irrelevance to psychological status following TBI (n = 1,431), a high-risk sample (e.g., military) (n = 977), and paediatric sample (n = 238). A full text review of the remaining 241 articles was then carried out against the exclusion criteria. 47 studies satisfied the criteria for inclusion within this meta-analysis. Of these, 23 studies had overlapping samples either due to reporting different epochs of a longitudinal study or were multi-centre trials reporting different study variables. The data from these 23 articles were combined, resulting in nine independent samples (see appendix 2). Thus, a total of 33 independent studies were included in the meta-analysis.

Figure 1

Results of the systematic search and the application of the inclusion criteria



Data extraction

All data were extracted by the author. It is anticipated that event rates will be reported as the number of participants with or without a clinically significant level of anxiety or GAD. This is determined by diagnosis following clinical interview, or by observing the cutoff scores on self-report measures for their respective clinical thresholds of probable anxiety. Multiple reporting of event rates may occur where primary studies report multiple measures of the same outcome or where they report the same prevalence data in different subgroups. Where possible, multiple outcomes will be combined into a single quantitative prevalence rate figure using the methods described by Borenstein et al. (2009). However, if it is not possible to combine the effects into a single quantitative measure, multiple effects from the same primary study will be included. The inclusion of multiple effects from the same study may cause an inflation of the apparent sample size for the weighted average prevalence rate and, therefore, a slight reduction in confidence intervals for the meta-analytic synthesis.

Defining problematic variance

A study-level effect is considered heterogeneous when it exhibits variation from the meta-analysis synthesis that cannot be attributed to the true variation in the distribution of effect within the population. Heterogeneity may arise due to methodological differences across studies, measurement errors, or uncontrolled individual differences within the literature. Higgins I^2 serves as a commonly used measure of heterogeneity. Higher I^2 values indicate variation in the effect that cannot be attributed to true variation in distribution of effect within the population. Due to the substantial methodological diversity used across primary studies, problematic heterogeneity is defined as a Higgins I^2 value greater than 75%. Where unacceptable or problematic heterogeneity is observed, it becomes incongruous to combine these quite disparate outcomes and, therefore, subsequent analyses should focus the

empirical identification of the sources of heterogeneity among the estimates of anxiety prevalence in the primary studies.

Risk of Bias Assessment

A set of quality criteria was developed to evaluate risk of methodological bias within the included studies. These criteria were adapted from existing risk of bias frameworks, including The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) and its generalisation to non-randomised studies (i.e., the Risk of Bias Assessment Tool for Nonrandomised Studies (Kim et al., 2013)). The current framework assesses the risk of bias across six domains: selection bias, performance bias, detection bias, statistical bias, reporting bias, and generalization. The criteria for Low, Unclear, or High risk within these six domains are described in table 3.

Table 3

Domains of risk of bias and the criteria for ratings of low, unclear, or high risk

Domain	Details	Risk of Bias
Selection Bias	Were efforts made to minimise selection bias in the intervention studies such as consecutive sampling? Are the characteristics of the study population adequately reported?	High Risk- Includes an unacceptable (reporting less than 30% of the data) level of non-response rate. The characteristics of the study population are not reported (e.g. time since injury). Convenience or purposive sampling is used. Unclear Risk- Non-response rate is not reported. The characteristics of the study population are not clearly reported. For example, the country, setting, location, population demographics were not adequately reported (e.g. mean age not reported). The recruitment process of individuals is unclear or has not been reported. Low Risk- The characteristics of the study population are clearly described and without evidence of bias. Non-response rate is reported and of an acceptable level (set at 50%). The source population is well described, and the study reports the characteristics of the sample e.g. the study details subgroups. The recruitment method is clearly

Domain	Details	Risk of Bias
		reported and well defined. The article provides some reassurance that there is no selection bias.
Performance Bias	Were participants rewarded for their participation and therefore extrinsically motivated?	High Risk- Participants were rewarded for their participation in the study. Participants were told what questionnaires they were completing and why and any proposed hypotheses. Unclear Risk- It is not clear if participants were rewarded for their participation (e.g. motivation to respond in a certain way). It is unclear how much information was provided to the participant prior to taking part in the study. Low Risk- Participants were not rewarded for their participation in the study. Information and procedures are provided in a way that does not differentially motivate participants.
Detection Bias	Were valid and reliable outcome measures used?	High Risk- The outcome measures were implemented differently across participants. The outcome measures used had poor reliability and validity reported e.g. Cronbach's Alpha < 0.6. and/or test/retest reliability <0.6. States that the measure has been translated but does not detail how this was conducted or clear problems in translation. Unclear Risk- Information regarding the outcome measures are either not reported or not clearly reported e.g. definition, validity, reliability. Cronbach's Alpha for outcome measures is between 0.6 and 0.7. Test-retest reliability for outcome measures is between 0.6 and 0.7. It is not clear if the measure was implemented consistently across all participants. It is unclear if measures have been translated. Low Risk- The outcome measures are clearly defined, valid and reliable, and are implemented consistently across all participants. Translated versions of measures report process of translation and reliability/validity of translated measure.
Statistical Bias	Have appropriate statistical methods been used? Is there incomplete data due to attrition?	High Risk- Statistics were not reported. Attrition rate – data loss is reported at analysis at an unacceptable level. Greater than 30% attrition. Unclear Risk- Unclear what statistical test was used. Data loss is not reported at analysis and is therefore unclear. Low Risk- Appropriate statistical testing was used. Attrition rate – data loss is reported at analysis at an acceptable level. Less than 30% attrition.
Reporting Bias	Are outcome measures reported as outlined in the method?	High Risk- Not reported full outcome measures that are stated in the method section. Prevalence data is reported as

Domain	Details	Risk of Bias
	Is outcome data presented in a clear and unambiguous manner?	percentage rather than a whole number, leaving some ambiguity in interpreting the results. Unclear Risk- Not all descriptive and/or summary statistics are presented. There is a description in the results but does not record statistics. Reported more than one prevalence rate. Low Risk- Reported all results of measures as outlined in the method. Prevalence data is reported as a number.
Generalisation	Can research findings be generalised to settings beyond the original study context? Are there any differences between the study participants and the broader population to which the review findings apply?	High Risk- Small sample with or without idiosyncratic feature. The sample size is inadequate to detect an effect: <30 participants. Unclear Risk- The study has a sample size of >30 to <80 participants. Low Risk- Sufficient sample for generalisation and representative of target population. The sample size is adequate to detect an effect: ≤ 80 participants.

The application of the risk of bias criteria to the included studies is reported in figure 2. The use of suffixes (a, b, c, d) indicate multiple outcomes from independent samples within a study, including different diagnoses, timepoints, and injury severities. An overall quality index score was calculated by reviewing the study's design and risk of bias rating for each domain. A study hierarchy was established in which prospective case cohort designs were awarded the most points towards their overall quality index score. Retrospective case cohort designs, case control studies, and cross-sectional studies received incrementally less points, with case series designs being assigned the fewest.

Figure 2

Ratings of risk of bias

Study	*Selection Bias*	*Performance Bias*	*Detection Bias*	*Statistical Bias*	*Reporting Bias*	*Generalisability*	*Overall Quality Index*
Albicini & McKinlay (2018) [a]	Low risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	48%

Study	*Selection Bias*	*Performance Bias*	*Detection Bias*	*Statistical Bias*	*Reporting Bias*	*Generalisability*	*Overall Quality Index*
Albicini & McKinlay (2018) [b]	Low risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	48%
Albicini & McKinlay (2018) [c]	Low risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	48%
Albicini & McKinlay (2018) [d]	Low risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	48%
Al-Kader et al. (2022)	High risk	Unclear risk	Unclear risk	Low risk	Low risk	Unclear risk	61%
Alway et al. (2016) [a]	Low risk	High risk	Low risk	Unclear risk	Low risk	Low risk	89%
Alway et al. (2016) [b]	Low risk	High risk	Low risk	Unclear risk	Low risk	Low risk	89%
Alway et al. (2016) [c]	Low risk	High risk	Low risk	Unclear risk	Low risk	Low risk	89%
Alway et al. (2016) [d]	Low risk	High risk	Low risk	Unclear risk	Low risk	Low risk	89%
Anke et al. (2015)	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	93%
Auclair-Pilote et al. (2021)	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	93%
Chaurasiya et al. (2021)	High risk	Low risk	Unclear risk	Low risk	Low risk	High risk	73%
Curvis et al. (2018)	High risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	43%
Dahm & Ponsford (2015)	Low risk	High risk	Low risk	Low risk	High risk	Low risk	86%
de Koning et al. (2016)	Low risk	Low risk	Unclear risk	Low risk	High risk	Low risk	89%
Giustini et al. (2014)	Unclear risk	Low risk	Unclear risk	Low risk	High risk	Low risk	41%
Hart et al. (2014)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Julien et al. (2017)	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	80%
Lamontagne et al. (2022) [a]	Unclear risk	High risk	Low risk	Low risk	Low risk	Low risk	43%
Lamontagne et al. (2022) [b]	Unclear risk	High risk	Low risk	Low risk	Low risk	Low risk	43%
Lamontagne et al. (2022) [c]	Unclear risk	High risk	Low risk	Low risk	Low risk	Low risk	43%
Lamontagne et al. (2022) [d]	Unclear risk	High risk	Low risk	Low risk	Low risk	Low risk	43%
Leong Bin Abdullah et al. (2018)	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk	48%
Maestas et al. (2014)	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	45%
Marinkovic et al. (2020)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Mascialino et al. (2022)	Unclear risk	Low risk	Low risk	Low risk	High risk	Unclear risk	41%
Mikolić et al. (2021) [a]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Mikolić et al. (2021) [b]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
O'Donnell et al. (2016)	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	91%
Osborn et al. (2017)	Unclear risk	Low risk	Low risk	Low risk	Low risk	Low risk	93%
Ponsford et al. (2019)	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	91%
Popov et al. (2022)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	84%
Shields et al. (2016)	Low risk	Low risk	Low risk	Low risk	High risk	Unclear risk	89%
Silverberg et al. (2018)	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	50%
Simon et al. (2020)	High risk	Low risk	Low risk	High risk	Low risk	Low risk	41%
Singh et al. (2019)	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	91%
Singh et al. (2019)	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	91%

Study	*Selection Bias*	*Performance Bias*	*Detection Bias*	*Statistical Bias*	*Reporting Bias*	*Generalisability*	*Overall Quality Index*
Stenberg et al. (2015) [a]	Unclear risk	Low risk	Unclear risk	Low risk	Low risk	Unclear risk	89%
Stenberg et al. (2015) [b]	Unclear risk	Low risk	Unclear risk	Low risk	Low risk	Unclear risk	89%
Stenberg et al. (2022)	Low risk	Low risk	Unclear risk	Low risk	Low risk	High risk	89%
Theadom et al. (2016) [a]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Theadom et al. (2016) [b]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Theadom et al. (2016) [c]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	95%
Tölli et al. (2018)	Low risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	93%
Vikane et al. (2019)	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	93%
Yilmaz et al. (2017)	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	93%
Zahniser et al. (2019) [a]	High risk	Low risk	Unclear risk	Low risk	High risk	Low risk	84%
Zahniser et al. (2019) [b]	High risk	Low risk	Unclear risk	Low risk	High risk	Low risk	84%
Zhu et al. (2016)	High risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	43%

Note: Red indicates high risk of bias, amber marks an unclear risk of bias, and green signifies a low risk of bias

Selection Bias

Most studies (n = 22) were rated low risk of bias with clear details about demographic and injury detail of participants, and how they were recruited to the study. Six studies were rated as high risk, as some characteristics of the study population or injury detail were not reported (Al-Kader et al., 2022; Curvis et al., 2018; Chaurasiya et al., 2021; Simon et al., 2020). Convenience sampling was used in three studies (Al-Kader et al., 2022; Chaurasiya et al., 2021; Zahniser et al., 2019). Five studies were rated as unclear risk of bias for reasons which included limited data regarding age (Stenberg et al., 2015) and time since injury (Giustini et al., 2014; Osborn et al., 2017). Two studies used participants from a larger study sample, and it was unclear how they had been selected for the smaller analysis (Mascialino et al., 2022; Lamontagne et al., 2022).

Performance Bias

Most studies (n = 27) were rated low risk for performance bias because data was collected routinely in clinical settings following traumatic brain injury. Three studies were rated as high risk of bias as participants were recruited from TBI admissions to a hospital providing rehabilitation under a no-fault accident compensation system (Alway et al., 2016; Dahm & Ponsford, 2015) or received monetary compensation for their participation (Lamontagne et al., 2022). Thus, it was unclear if participants were extrinsically motivated to perform on anxiety measures. Three studies were rated as unclear risk, as it was not reported whether participants were rewarded for their participation (Al-Kader et al., 2022; Curvis et al., 2018; Leong Bin Abdullah et al., 2018).

Detection Bias

The majority of studies (n = 22) were rated as low risk of bias with clearly defined, valid and reliable measures which were implemented consistently across all participants. An unclear risk of bias was observed in 11 studies. Nine of these were due to a lack of clarity as whether the anxiety measure has been translated (Al-Kader et al., 2022; Anke et al., 2015; Auclair-Pilote et al., 2021; Chaurasiya et al., 2021; de Koning et al., 2016; Giustini et al., 2014; Stenberg et al., 2022; Vikane et al., 2019; Yilmaz et al., 2017). In one study it was not clear whether the anxiety measure was implemented consistently across all participants (Zahniser et al., 2019). One study did not provide sufficient information about the anxiety measure (Zhu et al., 2016).

Statistical Bias

Thirty studies were rated as low risk of bias with appropriate statistical testing implemented and acceptable levels of attrition. Two studies were rated high risk of bias due to

a lack of reporting around missing data and non-response rate (Julien et al., 2017; Simon et al., 2020). One study was rated as unclear risk of bias as the reason for attrition was not specified (Alway et al., 2016).

Reporting Bias

Most studies (n = 23) were rated as low risk of bias, with outcome measures reported as outlined in the method. Data was presented in a clear and unambiguous manner. Ten studies were rated as high risk of bias, as they reported prevalence of anxiety in percentages rather than a whole number, thus leaving some ambiguity in interpreting the results (Dahm & Ponsford, 2015; de Koning et al., 2016; Giustini et al., 2014; Maestas et al., 2014; Mascialino et al., 2022; O'Donnell et al., 2016; Ponsford et al., 2019; Shields et al., 2016; Singh et al., 2019; Zahniser et al., 2019).

Generalisability

The majority of studies had a sufficient sample for generalisation and were representative of an adult TBI population due to sample sizes of over 80 participants (n = 26). Two studies were rated high risk of bias due to sample sizes smaller than 30 participants (Chaurasiya et al., 2021; Stenberg et al., 2022). Six studies were rated as unclear risk due to sample sizes between 30 and 80 participants (Albicini & McKinlay, 2018; Al-Kader et al., 2022; Mascialino et al., 2022; Stenberg et al., 2015; Shields et al., 2016; Tolli et al., 2018).

Summary

Overall, there was a predominantly low risk of bias across the studies included in the meta-analysis. Five studies did not report any unclear or high risk of bias in any of the quality criteria (Marinkovic et al., 2020; Mikolic et al., 2021; Popov et al., 2022; Silverberg et al.,

2018; Theadom et al., 2016). Due to the relatively large number of studies included in this synthesis and the overall quality of data, the results of this meta-analysis are expected to provide a robust representation of anxiety prevalence following TBI in an adult population.

Results

Participant details

There were 33 studies reporting a total of 49 effects in a combined sample of 12,063 participants. The summary demographic and injury characteristics for these studies are reported in Table 4. On average, participants were middle aged adult males (65.5%). These findings are similar to those reported in the Osborn et al. (2017) meta-analysis where a mean age of 38.2 years (*SD* 7.6 years) was observed and 69% males. The average interval between TBI and assessment of anxiety was 2.2 years. Whilst only seven studies provided a mean GCS score, 32 studies provided categorical data relating to injury severity. Though most studies examined outcomes of anxiety within mild TBI (16 studies), a large number of studies provided mixed data from mild, moderate, and severe TBI (12 studies), which meant that impact of injury severity could not be calculated within this group.

Recruitment took place equally across inpatient and outpatient settings, with small numbers taking place across both inpatient and outpatient sessions, and a small number taking place across community settings. Most studies included participants with a history of clinical anxiety or depression prior to their TBI (18 studies), whilst seven studies excluded participants with a psychiatric history and a further eight studies did not specify pre-morbid anxiety or depression. Eight studies reported included participants with a pre-injury history of TBI which equated to 14.6% of the overall sample. Seven studies excluded participants with prior TBI (44.3%), and 18 studies did not state whether pre-injury TBI was present in their sample or not. Medication use was only reported in five studies; four included participants

who were prescribed psychotropic medication (15.8% of the overall sample), and one study excluded participants who were prescribed medication (1.8%). The remaining 28 studies did not specify if participants were prescribed medication or not. Finally, only five studies assessed a control group alongside their TBI sample. Four studies recruited a medical control group (orthopaedic injury or general trauma) and one study recruited significant others.

Table 4

Summary demographic and injury characteristics for the studies

Variable	N studies	N participants	%	Mean	*SD*
Sample size	33	12,063		354.8	542.5
Age (years)	32	12,032		40.2	6.27
Gender (males)	33	7,901	65.5		
Time since injury (months)	28	11,090		26.4	40.6
Injury severity (GCS)	7	755		9.81	3.4
Injury severity					
Mild	16	6,335	52.5		
Mild, moderate	2	448	3.7		
Mild, moderate, severe	12	4,761	39.5		
Moderate, severe	3	237	2		
Severe	1	197	1.6		
Not specified	1	85	0.7		
Recruitment source					
Inpatients	12	5,833	48.4		
Outpatients	17	5,843	48.4		
Both inpatients and outpatients	2	165	1.4		
Community	2	222	1.8		
Pre-injury history of depression or anxiety					
Participants with history included	18	9,258	76.7		
Participants with history excluded	7	1,080	9		
Not specified	8	1,725	14.3		
Pre-injury history of TBI					
History of prior TBI	8	1,760	14.6		
No history of prior TBI	7	5,338	44.3		
Not specified	18	4,965	41.1		
Medication					
Psychotropic medication	4	1,901	15.8		
Participants excluded if using medication	1	220	1.8		
Not specified	28	9,942	82.4		
	N Studies	N TBI	%	N Control	%
Type of control group					
Medical	4	643	70.9	264	29.1
Significant others	1	31	50	31	50
TOTAL		674		295	

Note. N Studies and N Participants indicate the total number of studies and participants for which data were available. GCS = Glasgow Coma Scale.

An overview of individual study characteristics is presented in Table 5. Participants were selected from 15 different countries, whilst one study obtained data from multiple centres across Europe. Post-injury intervals were categorised as early post-injury (≤6 months), short term (>6 months to ≤2 years), medium term (>2 years to ≤5 years), and (long term (>5 years). In total, 16 studies assessed anxiety in the early stage (*M* = 3.07 months, *SD* = 1.9); 11 in the short-term (*M* = 13.9, *SD* = 4.3); two studies assessed anxiety in the medium-term (*M* = 45.1, *SD* = 9.8); and five in the long-term (*M* = 110.6, *SD* = 39.5). Five studies did not report the time at which anxiety was assessed following TBI. Of the 33 studies, ten provided prevalence data for GAD specifically, whilst 26 studies provided data for clinically significant anxiety. Measures of anxiety included Composite International Diagnostic Interview (CIDI; Kessler et al., 1998), Generalized Anxiety Disorder 7 (GAD-7; Spitzer et al., 2006), Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First et al., 1997), Hospital Anxiety and Depression Scale (HADS, (Zigmond & Snaith, 1983), Brief Symptom Inventory (BSI; Derogatis, 1975), Brief Symptom Inventory 18 (BSI-18; (Derogatis, 2001); Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), Golberg Anxiety Scale (GAS; Goldberg et al., 1988), and Self-Rating Anxiety Scale (SAS; Zung, 1971). Eight studies provided data for multiple different variables, such as severity of TBI, time since injury, or type of anxiety.

Table 5

Demographic of individual studies

Study/Sample	Country	Age (mean)	Gender (% males)	Time since injury (months)	Post-injury interval	TBI severity	Type of anxiety	Anxiety measure
Albicini & McKinlay (2018) [a]	New Zealand	24.5	66.2	146.04	Long-term	Mild	GAD	CIDI
Albicini & McKinlay (2018) [b]	New Zealand	24.5	54.1	181.56	Long-term	Moderate-severe	GAD	CIDI
Albicini & McKinlay (2018) [c]	New Zealand	24.5	66.2	146.04	Long-term	mild	Anxiety	CIDI
Albicini & McKinlay (2018) [d]	New Zealand	24.5	54.1	181.56	Long-term	Moderate-severe	Anxiety	CIDI
Al-Kader et al. (2022)	Pakistan	-	64.5	-	-	Mild	GAD	GAD-7
Alway et al. (2016) [a]	Australia	35.02	78.3	24	Short-term	Moderate-severe	Anxiety	SCID-I
Alway et al. (2016) [b]	Australia	35.02	78.3	36	Medium-term	Moderate-severe	Anxiety	SCID-I
Alway et al. (2016) [c]	Australia	35.02	78.3	48	Medium-term	Moderate-severe	Anxiety	SCID-I
Alway et al. (2016) [d]	Australia	35.02	78.3	60	Medium-term	Moderate-severe	Anxiety	SCID-I
Anke et al. (2015)	Norway	40.1	78	12	Short-term	Severe	Anxiety	HADS
Auclair-Pilote et al. (2021)	Canada	39.97	35.8	1.45	Early post-injury	Mild	Anxiety	HADS
Chaurasiya et al. (2021)	India	27.59	68.2	-	-	Mild-moderate	Anxiety	HADS
Curvis et al. (2018)	United Kingdom	42.4	63.5	92.64	Long-term	-	Anxiety	HADS
Dahm & Ponsford (2015)	Australia	35.9	60	87.59	Long-term	All	Anxiety	HADS
de Koning et al. (2016)	Netherlands	40	62.9	0.53	Early post-injury	Mild	Anxiety	HADS
Giustini et al. (2014)	Italy	31.6	77.6	-	-	All	Anxiety	HADS
Hart et al. (2014)	USA	40.2	74	6	Early post-injury	All	Anxiety	BSI
Julien et al. (2017)	Canada	40.37	41.9	0.46	Early post-injury	Mild	Anxiety	BAI
Lamontagne et al. (2022) [a]	Canada	41.3	75.8	4	Early post-injury	Mild	Anxiety	HADS
Lamontagne et al. (2022) [b]	Canada	41.3	75.8	12	short-term	Mild	Anxiety	HADS
Lamontagne et al. (2022) [c]	Canada	41.3	75.8	4	Early post-injury	Mild	GAD	MINI
Lamontagne et al. (2022) [d]	Canada	41.3	75.8	12	Short-term	Mild	GAD	MINI
Leong Bin Abdullah et al. (2018)	Malaysia	37	84	16	Short-term	All	GAD	SCID-I
Maestas et al. (2014)	USA	33.35	76	3	Early post-injury	Mild	Anxiety	BSI
Marinkovic et al. (2020)	Finland	40.5	57.3	3.8	Early postinjury	Mild	GAD	SCID-I
Mascialino et al. (2022)	Ecuador	36	93.94	6	Early post-injury	All	GAD	GAD-7
Mikolić et al. (2021) [a]	Europe (63 centres)	54	64	6	Early post-injury	Mild	GAD	GAD-7
Mikolić et al. (2021) [b]	Europe (63 centres)	50	74	6	Early post-injury	All	GAD	GAD-7
O'Donnell et al. (2016)	Australia	39.5	71.5	72	Long-term	Mild	Anxiety	MINI
Osborn et al. (2017)	Australia	40.8	70.2	-	Medium-term	All	Anxiety	GAS
Ponsford et al. (2019)	Australia	54	54.5	6.9	Short-term	Mild	Anxiety	HADS
Popov et al. (2022)	Canada	39.74	35.2	14.8	Short-term	Mild	GAD	GAD-7
Shields et al. (2016)	Australia	42.3	72	36.58	Medium-term	All	Anxiety	BSI-18
Silverberg et al. (2018)	Canada	41.5	44.3	2.69	Early post-injury	Mild	Anxiety	MINI
Simon et al. (2020)	USA	43	36	-		All	GAD	GAD-7
Singh et al. (2019)	United Kingdom	46.9	68.7	2.3	Early post-injury	All	Anxiety	HADS
Singh et al. (2019)	United Kingdom	46.9	68.7	12	Short-term	All	Anxiety	HADS
Stenberg et al. (2015) [a]	Sweden	42	86	3	Early post-injury	Severe	Anxiety	HADS
Stenberg et al. (2015) [b]	Sweden	42	86	12	Short-term	Severe	Anxiety	HADS
Stenberg et al. (2022) [c]	Sweden	46	67	84	Long-term	Severe	Anxiety	HADS
Theadom et al. (2016) [a]	New Zealand	37.5	58.9	1	Early post-injury	Mild	Anxiety	HADS
Theadom et al. (2016) [b]	New Zealand	37.5	58.9	6	Early post-injury	Mild	Anxiety	HADS
Theadom et al. (2016) [c]	New Zealand	37.5	58.9	12	Short-term	Mild	Anxiety	HADS

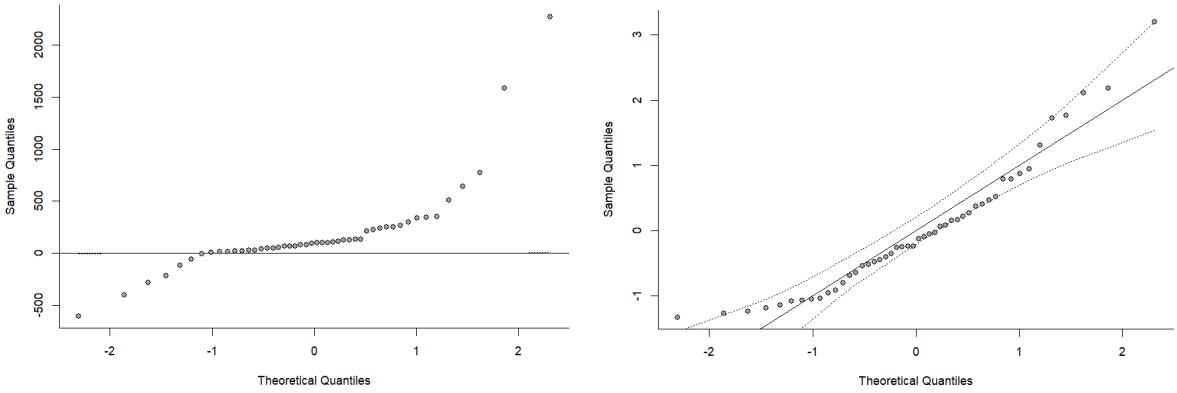
Study/Sample	Country	Age (mean)	Gender (% males)	Time since injury (months)	Post-injury interval	TBI severity	Type of anxiety	Anxiety measure
Tölli et al. (2018)	Sweden	47.1	73.2	12	Short-term	Moderate-severe	Anxiety	HADS
Vikane et al. (2019)	Norway	33	63	1.61	Early post-injury	Mild	Anxiety	HADS
Yilmaz et al. (2017)	Netherlands	45	63	3	Early post-injury	Mild-moderate	Anxiety	HADS
Zahniser et al. (2019) [a]	USA	39.97	65.4	0.46	Early post-injury	Mild	Anxiety	BSI-18
Zahniser et al. (2019) [b]	USA	39.97	65.4	3	Early post-injury	Mild	Anxiety	BSI-18
Zhu et al. (2016)	China	45	76.8	19.07	Short-term	All	Anxiety	SAS

Selection of the meta-analytic model

The distribution of study-level effects is shown in figure 3, using the fixed effects model and the random effect model (using the restricted maximum likelihood estimator of between studies variation).

Figure 3

QQ plot of the distribution of prevalence within the primary studies using the fixed effects model and the random effects model



(a) Fixed effects Model *(b) Random effects model*

As can be seen from figure 3, there is clear evidence of non-linearity and non-normality in the distribution of the prevalence of anxiety using the fixed effects model, which was absent when the distribution of effects was calculated using the random effects model.

Therefore, the random effects model using the restricted maximum likelihood estimate of between studies variation is an appropriate method for calculating the weighted average prevalence rate.

The omnibus tests

The study level prevalence of anxiety reported in the studies is presented in Table 6.

Table 6

Study level effect sizes, standard errors and 95% confidence intervals

Study	Effect	Standard Error	CI Lower	CI Upper	Weight (random)
Albicini & McKinlay (2018) [a]	0.07692	0.033051	0.012143	0.1417	42.22
Albicini & McKinlay (2018) [b]	0.08197	0.035122	0.013129	0.15081	41.97
Albicini & McKinlay (2018) [c]	0.18462	0.048124	0.090295	0.27894	40.14
Albicini & McKinlay (2018) [d]	0.2623	0.056321	0.151908	0.37268	38.81
Al-Kader et al. (2022)	0.41935	0.088627	0.245649	0.59306	32.84
Alway et al. (2016) [a]	0.24444	0.036988	0.17195	0.31694	41.73
Alway et al. (2016) [b]	0.27826	0.04179	0.196355	0.36017	41.08
Alway et al. (2016) [c]	0.23364	0.040907	0.153468	0.31382	41.21
Alway et al. (2016) [d]	0.16832	0.037229	0.095349	0.24128	41.7
Anke et al. (2015)	0.13934	0.031353	0.077893	0.2008	42.41
Auclair-Pilote et al. (2021)	0.5419	0.03724	0.46891	0.61489	41.7
Chaurasiya et al. (2021)	0.23077	0.067466	0.098538	0.363	36.84
Curvis et al. (2018)	0.70588	0.049422	0.609018	0.80275	39.94
Dahm & Ponsford (2015)	0.34091	0.05053	0.241872	0.43995	39.77
de Koning et al. (2016)	0.0625	0.014677	0.033733	0.09127	43.84
Giustini et al. (2014)	0.48299	0.041215	0.402212	0.56377	41.16
Hart et al. (2014)	0.03446	0.006841	0.021051	0.04787	44.17
Julien et al. (2017)	0.18471	0.030971	0.124011	0.24542	42.46
Lamontagne et al. (2022) [a]	0.28333	0.041136	0.202709	0.36396	41.18
Lamontagne et al. (2022) [b]	0.18333	0.035323	0.114102	0.25256	41.94
Lamontagne et al. (2022) [c]	0	0.001	-0.00196	0.00196	44.26
Lamontagne et al. (2022) [d]	0.04167	0.018242	0.005914	0.07742	43.62
Leong Bin Abdullah et al. (2018)	0.0198	0.013863	-0.007369	0.04697	43.89
Maestas et al. (2014)	0.29947	0.033494	0.233818	0.36511	42.17
Marinkovic et al. (2020)	0.02913	0.016569	-0.003349	0.0616	43.73
Mascialino et al. (2022)	0.18182	0.067141	0.050225	0.31341	36.9
Mikolić et al. (2021) [a]	0.058	0.004369	0.049438	0.06657	44.22
Mikolić et al. (2021) [b]	0.04876	0.005899	0.037201	0.06032	44.19
O'Donnell et al. (2016)	0.06419	0.010073	0.044446	0.08393	44.06
Osborn et al. (2017)	0.16058	0.031367	0.099105	0.22206	42.41
Ponsford et al. (2019)	0.14869	0.01921	0.111036	0.18634	43.55
Popov et al. (2022)	0.552	0.044479	0.464823	0.63918	40.7
Shields et al. (2016)	0.1	0.042426	0.016846	0.18315	40.99
Silverberg et al. (2018)	0.35443	0.053817	0.24895	0.45991	39.23
Simon et al. (2020)	0.34087	0.019052	0.303532	0.37821	43.56
Singh et al. (2019)	0.48865	0.013748	0.461708	0.5156	43.89
Singh et al. (2019)	0.36537	0.01386	0.338203	0.39253	43.89
Stenberg et al. (2015) [a]	0.21333	0.047304	0.12062	0.30605	40.27
Stenberg et al. (2015) [b]	0.21622	0.047855	0.122422	0.31001	40.19

Study	Effect	Standard Error	CI Lower	CI Upper	Weight (random)
Stenberg et al. (2022)	0.14286	0.07636	-0.006806	0.29252	35.18
Theadom et al. (2016) [a]	0.20235	0.021756	0.159705	0.24499	43.35
Theadom et al. (2016) [b]	0.24633	0.023333	0.200602	0.29207	43.22
Theadom et al. (2016) [c]	0.25513	0.023607	0.208863	0.3014	43.19
Tölli et al. (2018)	0.29268	0.071058	0.153412	0.43195	36.17
Vikane et al. (2019)	0.12346	0.025846	0.0728	0.17411	42.99
Yilmaz et al. (2017)	0.11736	0.015914	0.086168	0.14855	43.77
Zahniser et al. (2019) [a]	0.20781	0.015153	0.178112	0.23751	43.81
Zahniser et al. (2019) [b]	0.15342	0.013459	0.127038	0.1798	43.91
Zhu et al. (2016)	0.05909	0.015897	0.027933	0.09025	43.77

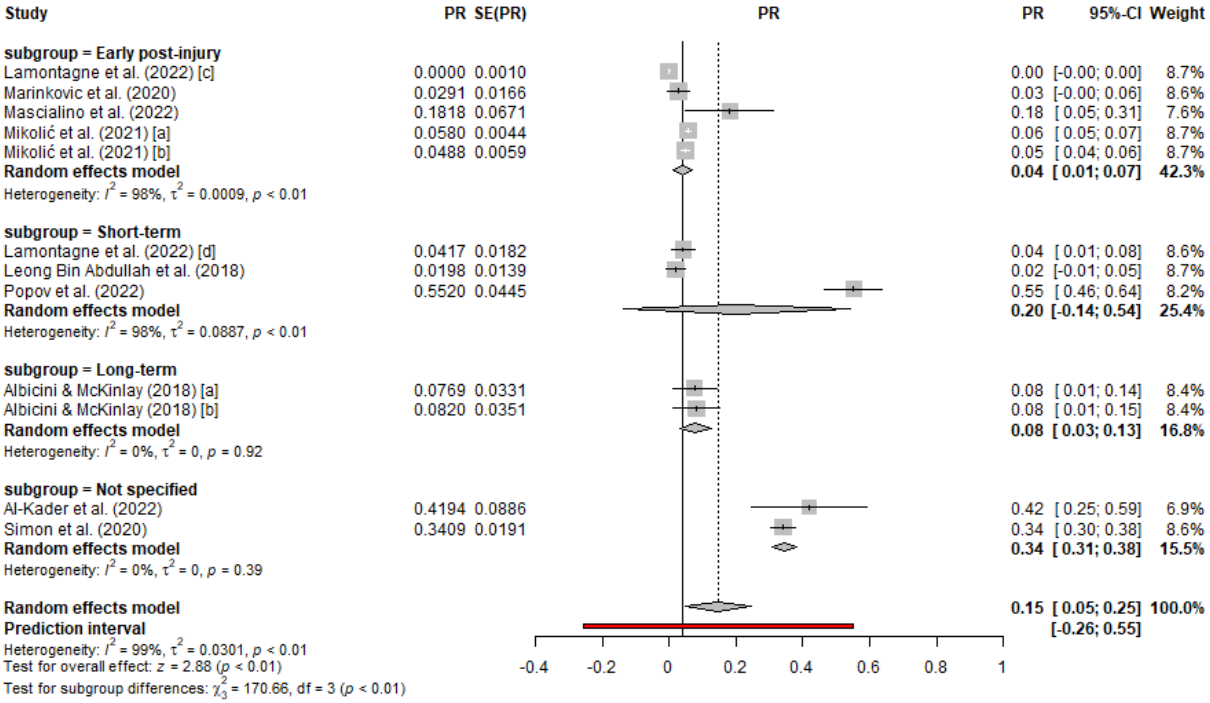
A random effects model was calculated using the generic inverse variance method. The random effects model, inclusive of both anxiety types and across all timepoints and TBI severity groups, evidenced a weighted average prevalence of anxiety = 0.21 ($z = 9.64$ $p = <0.0001$) and a 95% confidence interval of between 0.17 to 0.26.

Prevalence of GAD following TBI

The weighted average prevalence rate for GAD, sub-grouped by time since injury, is shown in figure 4.

Figure 4

Prevalence of GAD following TBI sub-grouped by post-injury interval



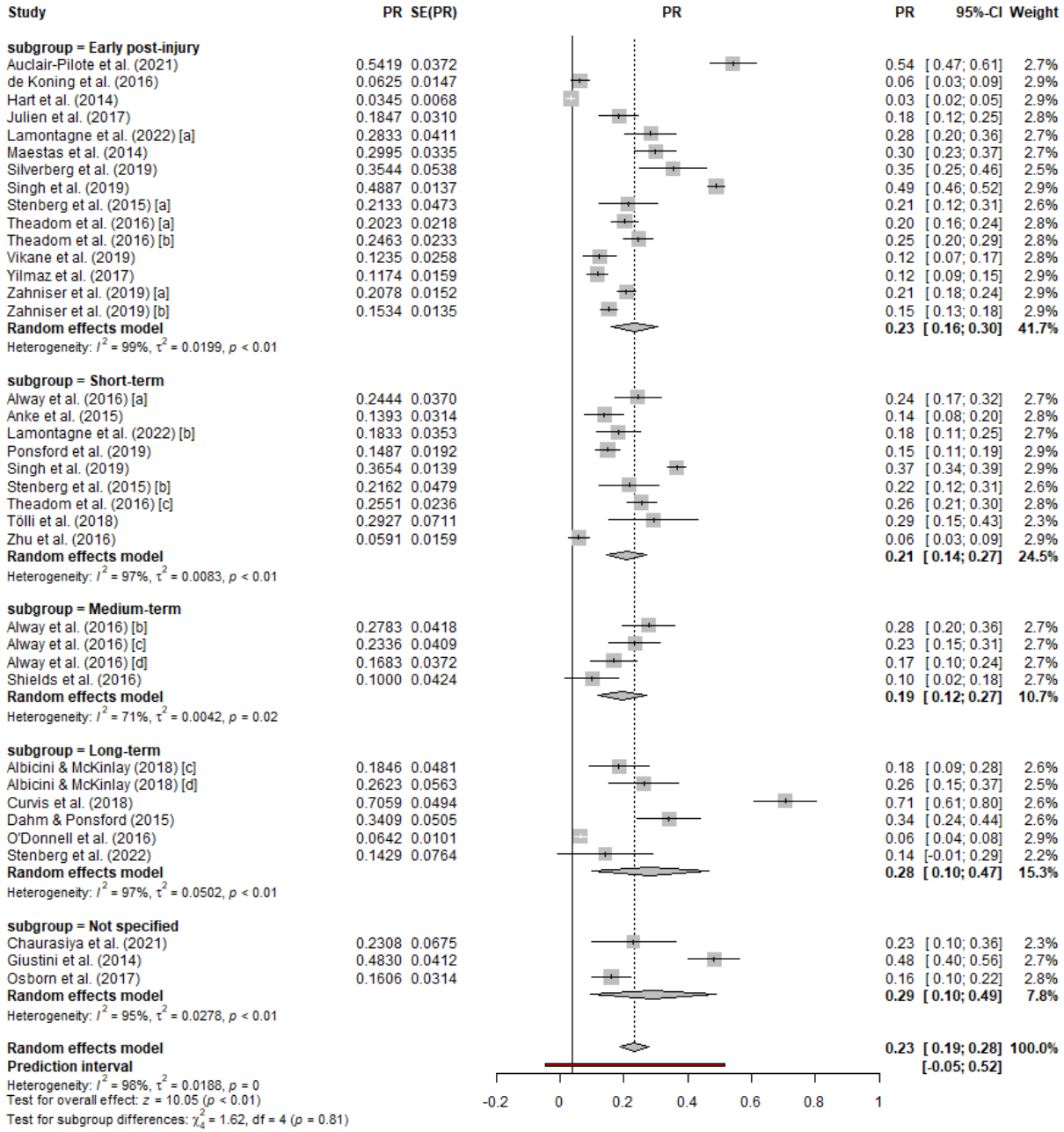
In the early post-injury stage, the prevalence rate for GAD was estimated at 0.04 (95% CI 0.01 to 0.07). The short-term interval was estimated at 0.20 (95% CI -0.14 to 0.54). The long-term interval was estimated at 0.08 (95% CI 0.03 to 0.13). Where the post-injury interval was not specified, the estimated prevalence rate was at 0.34 (95% CI 0.31 to 0.38). There is a statistically significant difference reported across post-injury intervals (x^2 = 170.66, p = <0.01), suggesting that time since injury has an impact on the prevalence of GAD following TBI. However, the largest significant finding in this analysis has emerged where the post-injury interval has not been specified, which presents a challenge with interpreting the significance of post-injury interval.

Prevalence of anxiety following TBI

The weighted average prevalence rate for clinically significant anxiety, grouped by post-injury interval, is described in figure 5.

Figure 5

Prevalence of anxiety following TBI sub-grouped by post-injury interval



In the early post-injury stage, the prevalence rate for anxiety was estimated at 0.23 (95% CI 0.16 to 0.30). The short-term interval was estimated at 0.21 (95% CI 0.14 to 0.27). The medium-term interval was estimated at 0.19 (95% CI 0.12 to 0.27). The long-term interval was estimated at 0.28 (95% CI 0.10 to 0.47). Where the post-injury interval was not specified, the estimated prevalence rate was at 0.29 (95% CI 0.10 to 0.49). There is no statistically significant difference reported across post-injury intervals ($x^2 = 1.62$, $p = 0.81$). This suggests that time since injury does not meaningfully impact the prevalence of anxiety following TBI.

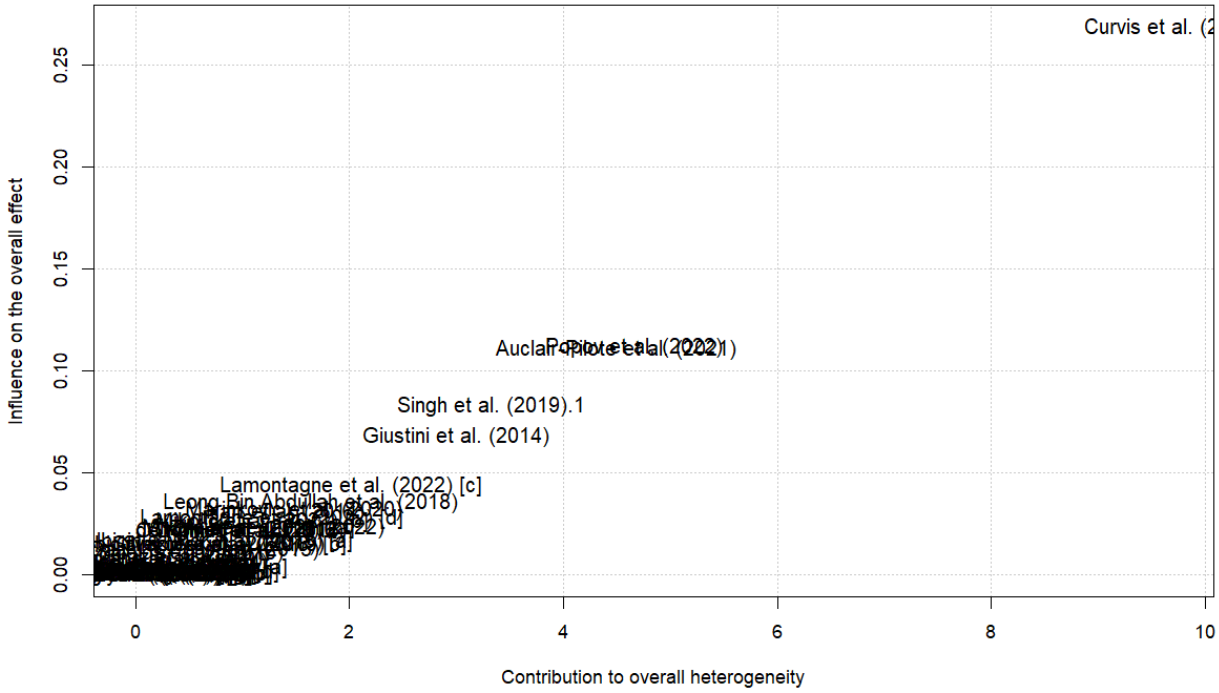
For both GAD and anxiety, there was marked heterogeneity in the study level effects (GAD $I^2 = 99\%$; Anxiety $I^2 = 98\%$). This suggests that the estimates of anxiety prevalence in the primary studies may be influenced by the presence of uncontrolled or confounding factors. Therefore, the focus of the subsequent analyses will be upon the identification of the sources of heterogeneity between the estimates of the prevalence of anxiety in the primary studies.

The impact of influential primary studies

The effect of studies with disproportionate influence was evaluated using a 'leave-one-out' analysis. In this approach, the random effects model was computed repeatedly, with each primary study removed in turn. The resulting changes in the weighted average effect size (representing influence) and the heterogeneity (indicating discrepancy) were recorded. The Baujat plot (Baujat et al., 2002) presents the results of this 'leave-one-out' analysis in figure 6.

Figure 6

Baujat diagnostic plot of sources of heterogeneity



Note: The vertical axis reports the influence of the study on the overall effect and the horizontal axis reports the discrepancy of the study with the rest of the literature.

The Baujat chart indicates that Curvis et al. (2018) is an influential and discrepant study, therefore it was removed, and the model was recalculated. The random effects model for long-term anxiety outcome was recalculated with the study showing disproportionate and discrepant influence removed. The corrected random effects model reported a synthesis of anxiety prevalence = 0.19 (95% CI 0.09 to 0.30) and evidences an approximately 8.9% decrease relative to the uncorrected estimate.

Curvis et al. (2018) was reviewed to identify any methodological factors that might account for its discrepancy from other studies reported in this review. As no such factors were identified, this study was included in subsequent analyses.

The effect of risk of bias in the primary studies

To assess the impact of study level risk of bias upon heterogeneity, subgroup analyses were carried out for the risk of bias ratings of 'low risk' and 'any risk' (i.e., unclear risk and high risk of bias combined) for each of the six types of methodological bias. These are shown in table 7. These analyses included both types of anxiety across all timepoints and injury severity.

Table 7

Effect of risk of bias in primary studies

Type of bias	Low Risk			Any Risk				
	Effect	95% CI	k	Effect	95% CI	k	X^2	P
Selection bias	0.20	0.15 to 0.25	33	0.24	0.15 to 0.33	16	0.51	0.48
Performance bias	0.21	0.16 to 0.25	37	0.24	0.13 to 0.35	12	0.27	0.60
Detection bias	0.21	0.16 to 0.26	35	0.22	0.17 to 0.26	14	0.01	0.91
Statistical bias	0.21	0.16 to 0.26	43	0.24	0.19 to 0.30	6	0.86	0.35
Reporting bias	0.20	0.15 to 0.25	37	0.24	0.15 to 0.33	12	0.51	0.48
Generalisability bias	0.22	0.16 to 0.27	37	0.19	0.13 to 0.24	12	0.80	0.37

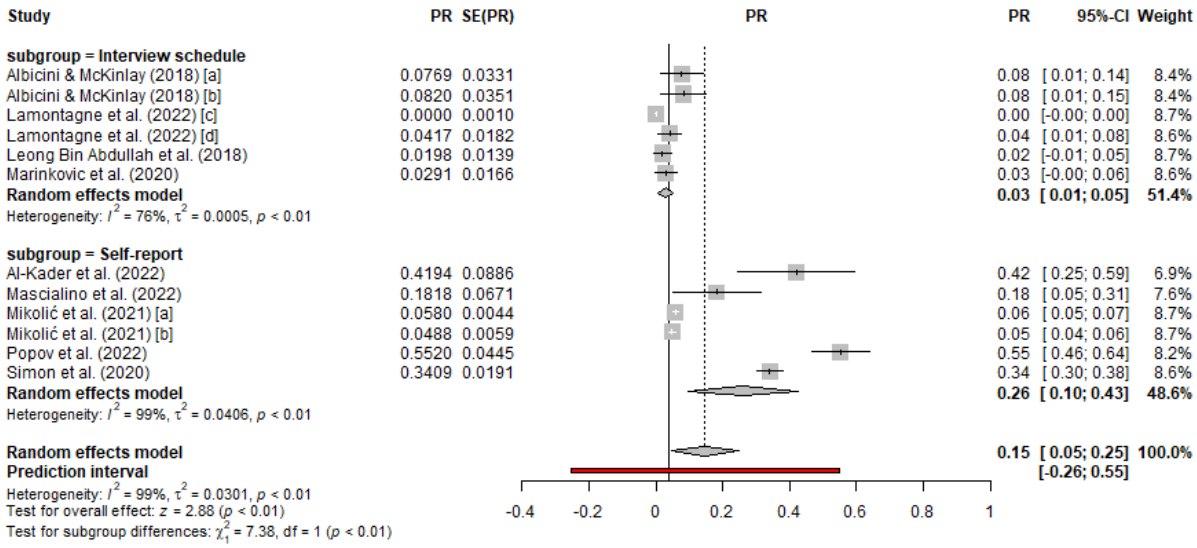
As shown above, there were no significant differences observed between anxiety prevalence estimates across risk of bias ratings. This suggests that inclusion of studies that are at risk of bias of any type, do not contribute to heterogeneity.

The impact of different ways of assessing GAD

A subgroup analysis was undertaken to assess the impact of different ways of assessing GAD. Study level effects were rated to indicate whether they had been derived from clinical interview or self-report. The weighted average prevalence rates for each of these assessment methods is reported in figure 7.

Figure 7

The impact of different ways of assessing GAD



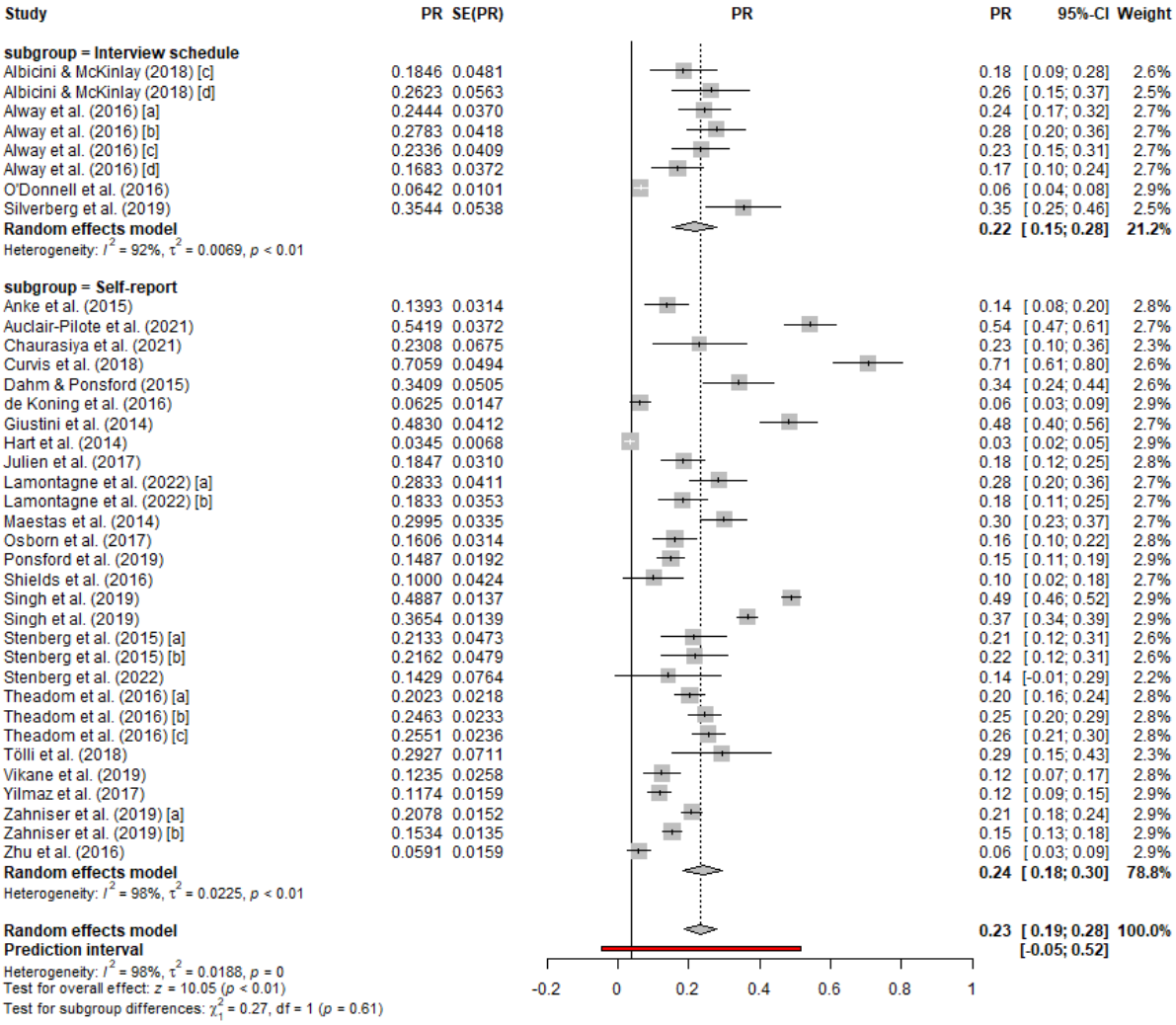
As can be seen above, there was a statistically significant difference between effect sizes assessed using clinical interview and those assessed using self-report methods (x^2 = 7.38, p = <0.01), with clinical interview recording smaller effects sizes. Accordingly, the inclusion of studies that rely upon self-report of anxiety have the effect of inflating the overall reported prevalence rate. It may be that self-report measures of GAD may be more sensitive to sub-threshold symptomology than clinical interview.

The impact of different ways of assessing anxiety

A further subgroup analysis was undertaken to assess the impact of method of assessing clinically significant anxiety. Study level effects were rated to indicate whether they had been derived from clinical interview or self-report. The weighted average prevalence rates for each of these assessment methods is reported in figure 8.

Figure 8

The impact of different ways of assessing anxiety



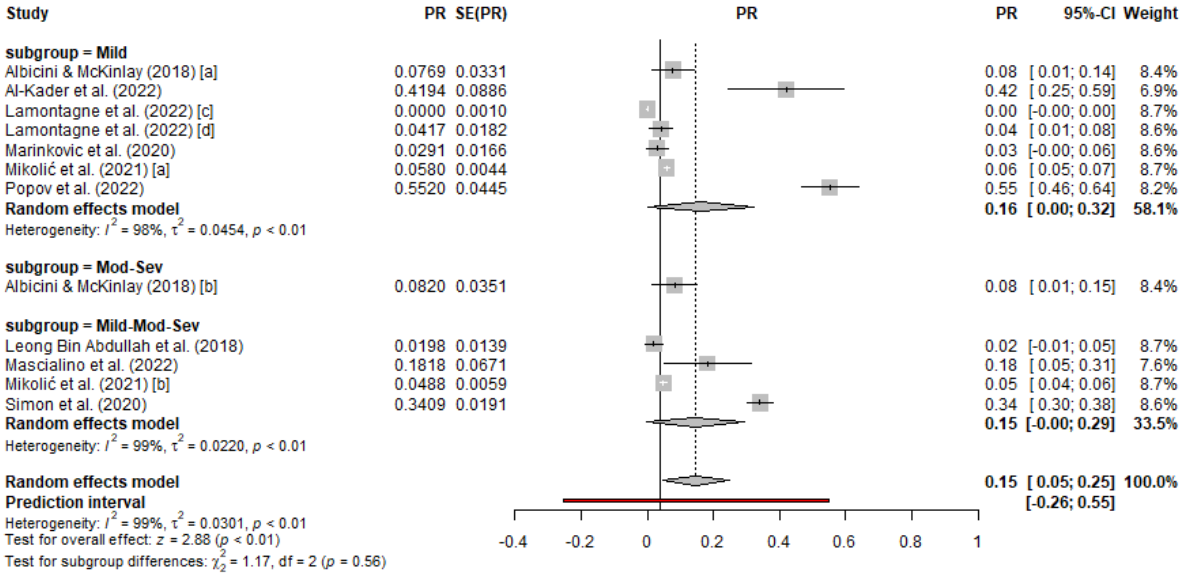
As illustrated above, there was no statistically significant difference observed between effect sizes assessed using clinical interview and those assessed using self-report methods (x^2 = 0.27, p = 0.61). Accordingly, the inclusion of studies that rely upon either clinical interview or self-report of anxiety have no effect on overall prevalence rate.

The impact of injury severity on GAD

To assess the impact of injury severity on GAD, a subgroup analysis was conducted, and study level effects were rated. The weighted average prevalence rates for each category of injury severity are reported in figure 9.

Figure 9

The impact of injury severity on GAD



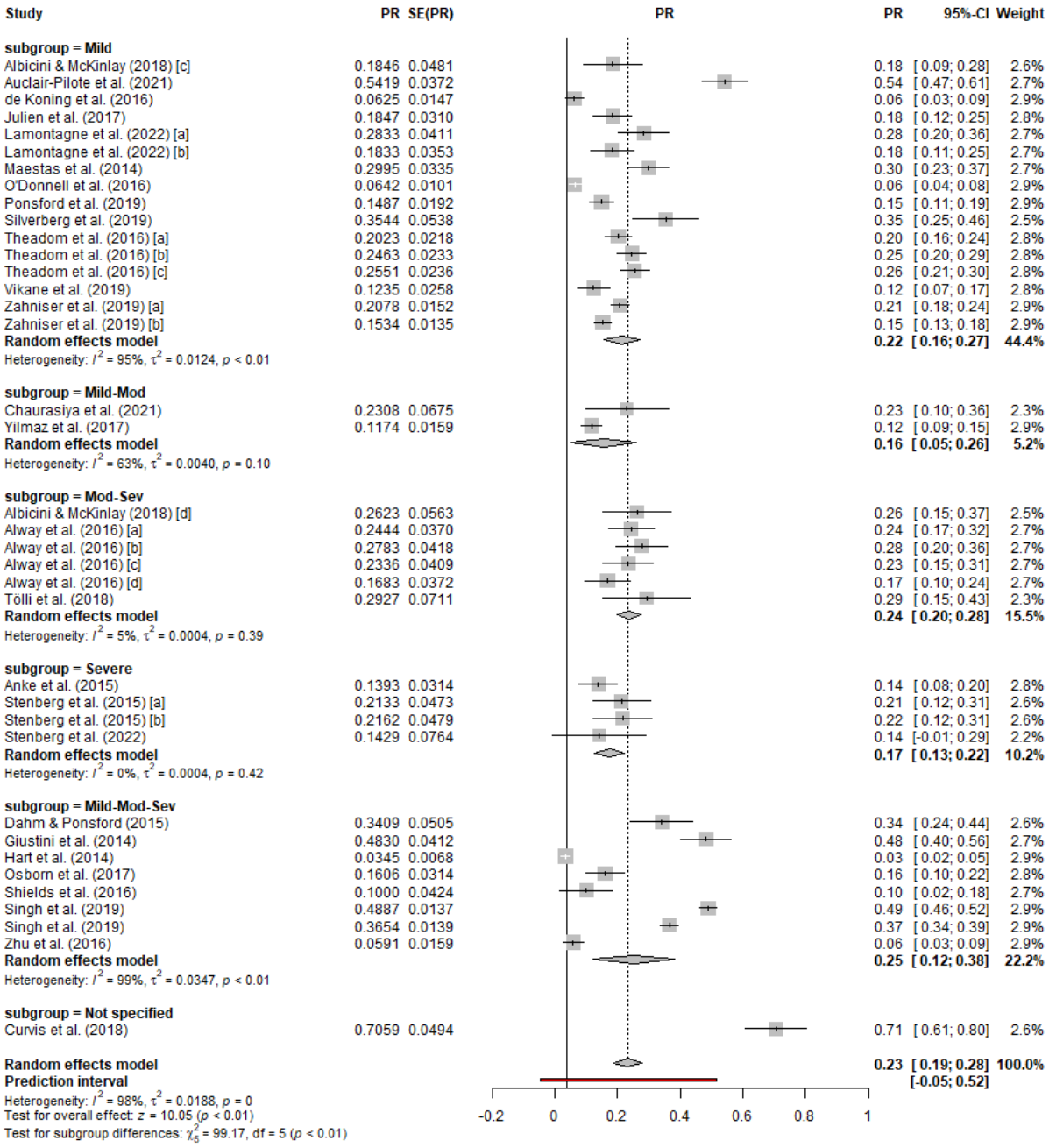
As shown above, there was no statistically significant difference between effect sizes across injury severity ($x^2 = 1.17$, $p = 0.56$). Therefore, the inclusion of participants with different injury severities has no effect on the overall prevalence rate of GAD. However, data is limited pertaining to moderate and severe TBI, which restricts direct comparisons between mild TBI and moderate to severe presentations.

The impact of injury severity on anxiety

To assess the impact of injury severity on clinically significant anxiety, a subgroup analysis was conducted, and study level effects were rated. The weighted average prevalence rates for each category of injury severity are reported in figure 10.

Figure 10

The impact of injury severity on anxiety



As indicated above, a significant difference was observed across injury severity types when the Curvis et al. (2018) paper was included with unspecified injury severity (x^2 = 99.17, p = <0.01). However, when this study was excluded from the analysis, there were no

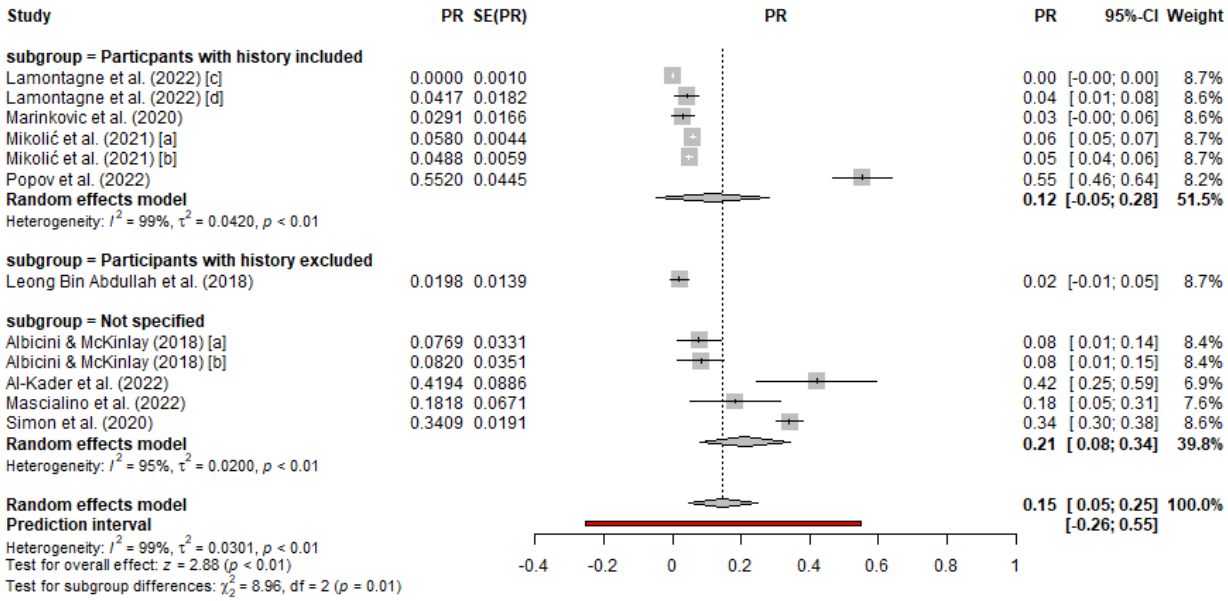
statistically significant differences reported across injury severity types ($x^2 = 5.47$, $p = 0.24$). This indicates that where data pertaining to injury severity has been available, it has had no effect on the overall prevalence rate of anxiety. The lack of clarity regarding the non-specified data makes it difficult to ascertain to what extent injury severity affects anxiety prevalence following TBI.

The impact of previously diagnosed psychiatric conditions on GAD

A subgroup analysis was undertaken to assess the impact of previously diagnosed psychiatric conditions on GAD across all timepoints and injury severity. Study level effects were rated to indicate whether participants had a previous history of mental health conditions or not. The weighted average prevalence rates for history are reported in figure 11.

Figure 11

The impact of psychiatric history on GAD



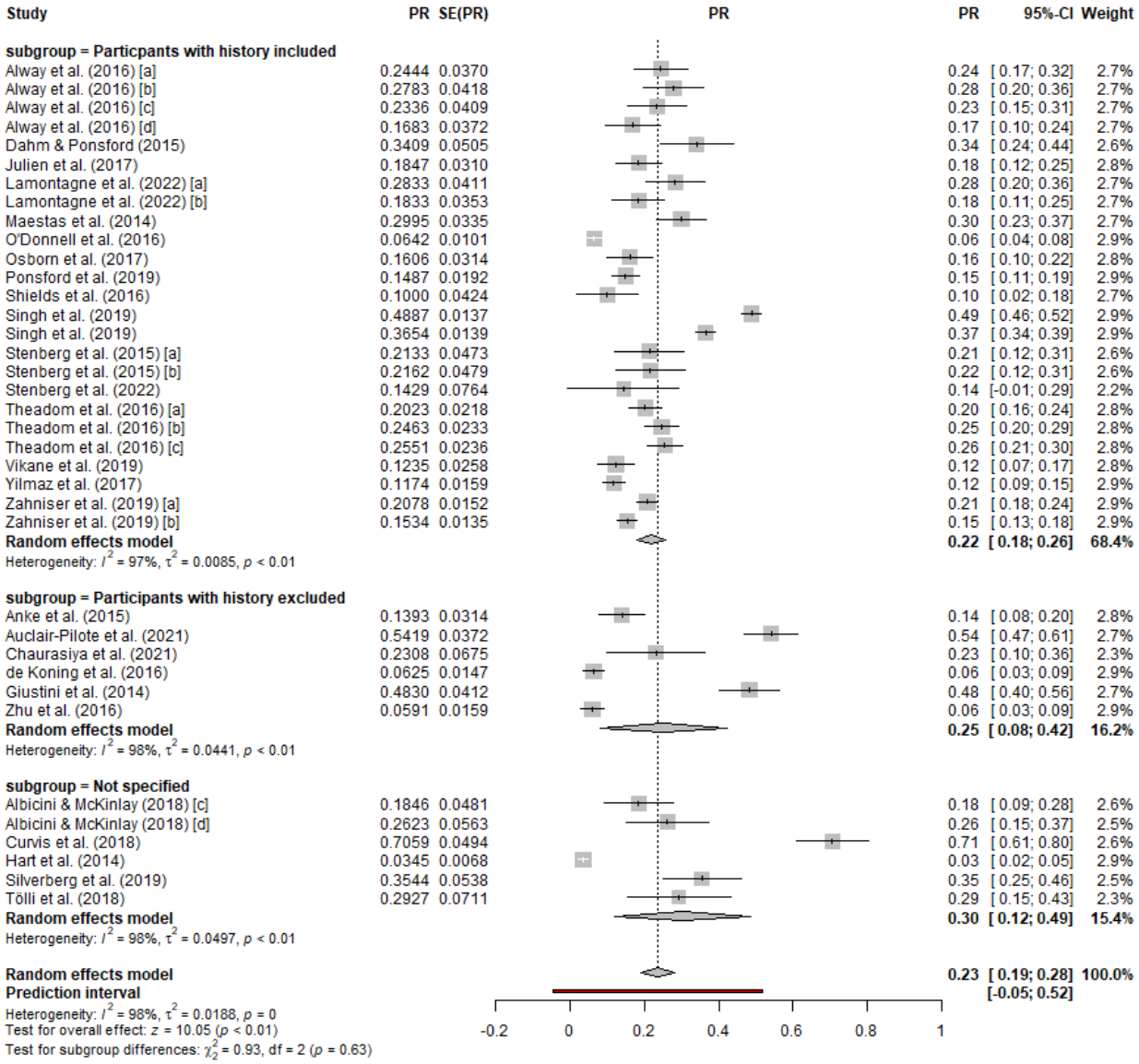
As illustrated above, there is a significant difference between effect sizes across psychiatric history ($x^2 = 8.96$, $p = 0.01$). However, the largest effect size is observed in the group where psychiatric history is not specified, so a firm conclusion about the influence of history cannot be drawn. Only one study reported prevalence data where participants with a psychiatric history were explicitly excluded.

The impact of previously diagnosed psychiatric conditions on anxiety.

A subgroup analysis was undertaken to assess the impact of previously diagnosed psychiatric conditions on anxiety across all timepoints and injury severity. The weighted average prevalence rates for history are reported in figure 12.

Figure 12

The impact of psychiatric history on anxiety



As can be seen above, no statistically significant differences were reported between effect sizes across psychiatric history ($x^2 = 0.93$, $p = 0.63$). Therefore, the inclusion of participants with a previous psychiatric history has no effect on the overall prevalence rate of anxiety, even where psychiatric history is not specified.

Summary of subgroup analyses

A summary of the significance of findings across subgroup analyses for GAD and anxiety is illustrated in table 8.

Table 8

Summary of findings

Subgroup factor	GAD	Anxiety
Post-injury interval	Significant	Not significant
Assessment method	Significant	Not significant
Injury severity	Not significant	Unclear
Psychiatric history	Significant	Not significant

As can been seen above, significant differences were observed across timepoints, assessment methods, and psychiatric histories for GAD, whilst only injury severity was a non-significant factor. Conversely, only one significant finding emerged amongst anxiety prevalence when an influential and discrepant study without injury severity data was included in the analysis (Curvis et al., 2018). This suggests that following TBI, anxiety prevalence is consistent across timepoints, and psychiatric history, but the extent to which injury severity affects anxiety prevalence is less clear.

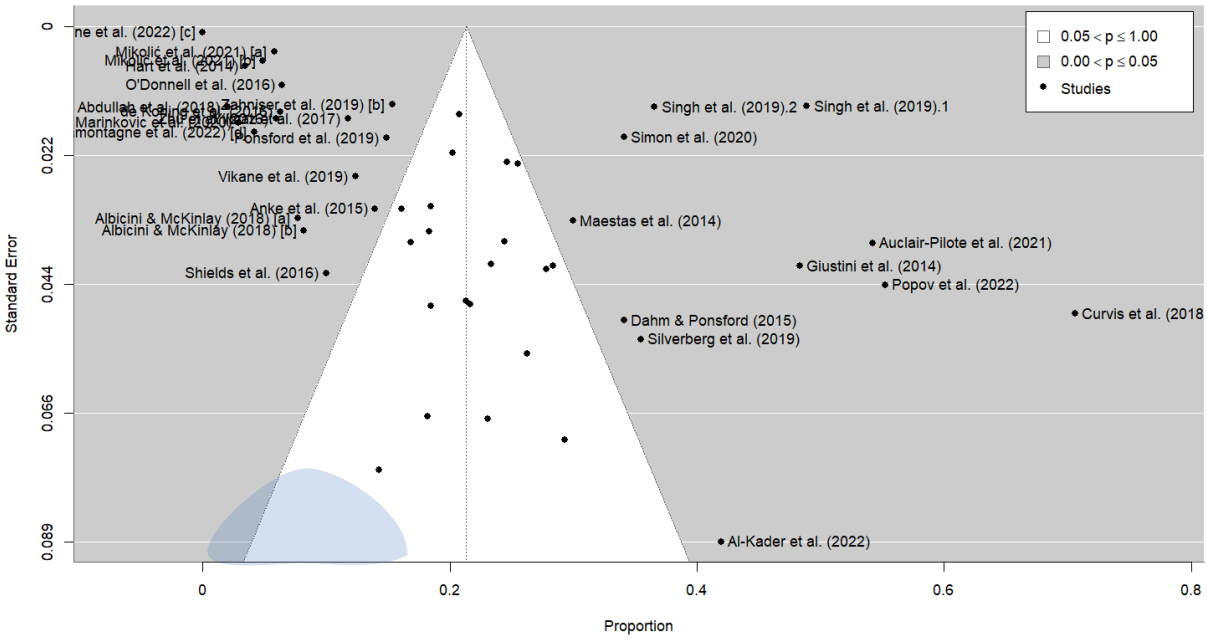
The impact of publication and small study biases

Publication bias arises when journals, publishing platforms, or studies themselves prioritise statistically significant findings and reject or do not report findings that are non-significant. For instance, a systematic review examining outcome reporting bias in randomised controlled trials found that more than half of 283 reviews did not include full data

for the primary outcome of interest (Kirkham et al., 2010). Small study bias occurs when smaller sample sizes demonstrate greater variability in their measurement of the prevalence of anxiety. Both types of bias can be identified in a funnel plot, which indicates the magnitude of the study's prevalence of anxiety and importance within the synthesis. The funnel plot estimates the deviance of each study from the meta-analytic average. In the absence of publication bias, small studies with greater variability in effect size distribute more widely at the bottom of the plot compared to larger studies towards the top, which lie closer to the overall meta-analytic effect and create a symmetrical funnel shape. If studies are absent from the area of the plot that is associated with small sample sizes and non-significant results, it is likely that there is some publication bias leading to an overestimation of the true effect. For this synthesis, the funnel plot of anxiety prevalence following TBI is presented in figure 13.

Figure 13

Funnel plot of the prevalence of anxiety following TBI



Note: The 95% confidence interval of the expected distribution of prevalence of anxiety is shown as an inverted 'funnel'. The estimated area of the funnel plot that is associated with small studies reporting no results is demarcated in blue.

Whilst there is high heterogeneity across the studies, there is also some evidence of publication bias in the distribution of prevalence of anxiety following TBI. This is evident in that small studies tend to report large effects and there is an absence of small studies in the area of the graph associated with null effects (i.e., there is an absence of small studies within the general population rate of anxiety, indicated by a blue area in figure 13). Therefore, the possibility of publication bias must be considered.

To further explore this, a trim and fill procedure (Duval & Tweedie, 2000) was simulated which assumes that publication bias produces an asymmetrical funnel plot. The trim and fill procedure iteratively removes the most extreme small studies from the side of the funnel plot associated with positive effects. At each iteration, the effect size is recomputed until the funnel plot becomes symmetric around the corrected effect size. This method of trimming yields an adjusted effect size and reduces the variance of the effects, leading to biased and narrow confidence intervals. The original studies are then reintroduced into the analysis, and a mirror image is imputed for each point of the funnel plot associated with negative effects. In this analysis, the trim and fill procedure failed to converge on an estimate and was therefore unable to correct for publication bias.

Orwin's (1983) failsafe number procedure was subsequently used which calculates the number of studies with non-significant results that would need to be included in the meta-analysis for the overall effect to be reduced to a minimally interpretable value. Application of different criterion levels for a minimally interpretable effect were determined using the 4% prevalence rate of anxiety in the general population (World Health Organization, 2023). As shown in table 8, this procedure suggests that 163 studies would be required to reduce the observed prevalence of anxiety following TBI from 21% to 8%. For a minimally interpretable effect of 10% to be seen, 93 studies would be required. For a minimally interpretable effect of 12% and treble that of the general population, 57 studies would be required. Assuming a

minimally interpretable effect of around 10%, the observed prevalence rate of 21% appears robust to studies missing due to publication bias.

Table 9

Application of different criterion levels for a minimally interpretable effect

Criterion level for a minimally interpretable effect	Average null effect size	Number of studies required*
0.08	0.04	163
0.10	0.04	93
0.12	0.04	57

Note: *Number of studies required to change the observed average effect to the minimally interpretable effect.

Discussion

This meta-analysis update reviewed studies which have reported the prevalence of GAD and clinically significant anxiety following TBI since May 2014. Given the disparity across reported prevalence rates and factors which may affect them, impact of time since injury, assessment method, injury severity, and psychiatric history were each examined. The results from this study are discussed in relation to findings from the previous meta-analysis by Osborn et al. (2016).

Prevalence of GAD

The prevalence of GAD following TBI ranged from 4% to 34% across different post-injury intervals. The lowest prevalence rate was observed in the six months following injury with a 4% prevalence rate, compared with a prevalence rate of 10% found by Osborn et al. during this time. From six months to two years post-injury, the prevalence of GAD increased to 20% in this study, whereas Osborn et al. saw a decrease to 6%. Beyond five years the

prevalence rate was 8% in this review and 5% in the previous. Data was not available between two-and five years post-injury in this study, however Osborn et al. found that GAD was most prevalent during this time, at 17%. The highest prevalence rate emerged in this study where the time since injury was unknown (34%). Across all time-points, the prevalence of GAD in the literature from 2014 is 15%, compared with 11% in Osborn et al. Whilst the findings from Osborne et al. do not fully correlate with this study, it is important to note that the previous incorporated studies that assessed GAD using clinical interview methods only, whereas this study has also incorporated self-report methods which may account for some of the discrepancy across findings.

Prevalence of anxiety

The prevalence of anxiety following TBI was consistent across post-injury intervals, ranging from 19% to 29%. During the six months post-injury period, a 23% prevalence rate was observed. In Osborn et al. it was 28% during this time. From six months to two years post-injury the prevalence of anxiety was 21% in this study but rose to 37% in Osborne et al. From the period of two to five years a rate of 19% was seen here, compared to its highest rate of 39% in Osborn et al. Beyond five years the prevalence rate was at 28% in this study, whilst remaining higher in Osborn et al. at 36%. As with results for GAD, the highest prevalence rate emerged in this study where the time since injury was unknown (29%). Across all time-points, the prevalence of anxiety in the literature from 2014 is 23%, compared with 37% in Osborn et al. Again, there was some discrepancy between findings across meta-analyses, although a much higher prevalence of anxiety compared to GAD was seen in both. A summary of the comparative findings for GAD and anxiety prevalence across studies is shown in table 10.

Table 10

Comparison of GAD and anxiety prevalence across studies

	Prevalence rate across studies published until May 2014 (Oborn et al., 2016)	Prevalence rate across studies published from June 2014 to July 2023
GAD		
≤6 months post-injury	10%	4%
6 months to 2 years	6%	20%
2 to 5 years	17%	-
>5 years	5%	8%
Unspecified	-	34%
Overall	11%	15%
Clinically significant anxiety		
≤6 months post-injury	28%	23%
6 months to 2 years	37%	21%
2 to 5 years	39%	19%
>5 years	36%	28%
Unspecified	-	29%
Overall	37%	23%

Whilst prevalence rates across post-injury intervals are inconsistent between studies, the past decade of research has seen a reduction in the overall prevalence rate of clinically significant anxiety from over a third of all people with TBI, to just under a quarter. However, this study incorporated three times more data than Osborn et al. which reported a combined sample of 4,210 participants. This may account for why prevalence rates for clinically significant anxiety were found to be more conservative here. Overall prevalence rates for GAD are similar between studies but have substantial variation across timepoints.

Method of assessment

A statistically significant difference was observed across method of assessment for GAD. A lower prevalence rate of 3% was seen where clinical interview was used to

determine the threshold for disorder, compared with a 26% prevalence rate for self-reporting of symptoms. These findings are consistent with literature that has identified higher rates of psychological distress through self-administered methods compared with clinical interviewing (Moum, 1998). Whilst some self-report measures of anxiety include symptoms that could be related to TBI and therefore inflate prevalence rates (Osborn et al., 2016), the GAD-7 does not include items related to memory, concentration, or sleep. The discrepancy between methods may be explained by the fact that to measure GAD by diagnostic interview, symptoms must be present for several months, whereas the self-report measure (i.e. the GAD-7) only requires symptoms to be present for two weeks. Osborne et al. (2016) only analysed clinical interview methods and found a higher rate of 11% than the 3% observed here. Anxiety prevalence was unaffected by method of assessment, with rates comparable across self-report (22%) and clinical interviewing (24%). Whilst clinical interviewing is generally seen as a gold standard approach to assessing psychological disorders, the consistency amongst prevalence rates of clinically significant anxiety suggests that the use of well validated self-report measures is a robust alternative to clinical interviewing.

Injury severity

The severity of TBI appeared to have no effect on the overall prevalence rate of GAD. However, data was limited and comparisons between mild TBI and moderate to severe TBI were restricted. One study with data for moderate to severe GAD revealed an 8% prevalence rate compared with 16% across studies in the mild TBI group. This is consistent with literature that has found mild TBI patients report higher rates of psychiatric disorders than those with moderate or severe injuries over time (Zgaljardic et al., 2015), though more data is required to strengthen these findings. The impact of injury severity on anxiety was unclear due to an influential and discrepant study, with prevalence rates ranging from 16% to 71%.

However, when this study was excluded, the findings became non-significant and ranged from 16% to 25%.

Psychiatric history

When individuals with a known psychiatric history were included in studies, a 12% prevalence rate of GAD was observed. Only one study provided data where participants with a history were excluded, which resulted in 2% prevalence. However, numerous studies did not state whether those with a psychiatric history were included or not which contributed to an overall prevalence rate amongst them of 21%. This makes it difficult to ascertain the impact of this factor upon GAD prevalence post-injury. A deficit of this information was also observed in the previous meta-analysis. Data for anxiety prevalence was more abundant, with a rate of 22% for participants with a psychiatric history and 25% for those without. As with rates for GAD, the highest prevalence rate was observed in the group where psychiatric history had not been examined or reported (30%), though no significant differences were observed amongst groups.

Limitations and recommendations for future research

Whilst there was some consistency with findings from the original meta-analysis, there was substantial discrepancy between prevalence rates of clinically significant anxiety between the two meta-analyses, with Osborn et al. (2016) reporting an overall much higher prevalence rate of 37% compared to 23% in this study which might be influenced by the size difference of the studies. The overall prevalence rate of GAD was similar across both studies (15% compared with 11% in Osborn et al.), but the inclusion of self-report assessment measures in this study accounted for this increase. Had self-report measures been excluded, a

3% prevalence rate of GAD would have been observed, again highlighting disparity with Osborn et al.

Numerous studies did not provide adequate information about participants, and this made it difficult to interpret the extent of the influence of some factors, particularly in the case of time since injury and pre-existing psychiatric history. This emphasises the importance of capturing and defining sample characteristics and methodical variables in future studies. Another limitation in this study was the lack of comparable control groups to contextualise findings. Future research that incorporates control groups such as orthopaedic trauma patients would enhance understanding about what factors impact TBI patients and whether this is unique to their presentation or experienced more broadly across trauma survivors.

Conclusion

In conclusion, the past decade of research has shown that following TBI, both GAD and clinically significant anxiety are observed at much higher rates than the general population base rate of 4%. Overall prevalence rates for GAD are almost four times higher than in the general population following TBI, though rates are sensitive to factors including time since injury, assessment method, and psychiatric history and vary substantially. Overall prevalence rates for anxiety are almost six times higher than the general population following TBI and remain largely consistent across factors. Findings support the literature that both GAD and anxiety may endure for many years post-injury, with anxiety prevalence rates at their highest over five years post-injury. These findings underscore the significant role clinicians may have in supporting individuals to cope with the long-term psychological impact of their injuries following TBI.

References

Al-Kader, D. A., Onyechi, C. I., Ikedum, I. V, Fattah, A., Zafar, S., Bhat, S., Malik, M. A., Bheesham, N., Qadar, L. T., & Cheema, M. S. (2022). Depression and anxiety in patients with a history of traumatic brain injury: A case-control study. *Cureus Journal of Medical Science*, *14*(8). https://doi.org/10.7759/cureus.27971

Albicini, M., & McKinlay, A. (2018). Anxiety disorders in adults with childhood traumatic brain injury: Evidence of difficulties more than 10 years post injury. *Journal of Head Trauma Rehabilitation*, *33*(3), 191–199. https://doi.org/10.1097/HTR.0000000000000312

Albrecht, J. S., Abariga, S. A., Rao, V., & Wickwire, E. M. (2020). Incidence of new neuropsychiatric disorder diagnoses following traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *35*(4), E352–E360. https://doi.org/10.1097/HTR.0000000000000551

Alway, Y., Gould, K. R., Johnston, L., McKenzie, D., & Ponsford, J. (2016). A prospective examination of Axis I psychiatric disorders in the first 5 years following moderate to severe traumatic brain injury. *Psychological Medicine*, *46*(6), 1331–1341. https://doi.org/10.1017/S0033291715002986

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Andelic, N., Soberg, H. L., Berntsen, S., Sigurdardottir, S., & Roe, C. (2014). Self-perceived health care needs and delivery of health care services 5 years after moderate-to-severe traumatic brain injury. *PM&R*, *6*(11), 1013–1021; quiz 1021. https://doi.org/10.1016/j.pmrj.2014.05.005

Anke, A., Andelic, N., Skandsen, T., Knoph, R., Ader, T., Manskow, U., Sigurdardottir, S., & Røe, C. (2015). Functional recovery and life satisfaction in the first year after severe traumatic brain injury: A prospective multicenter study of a Norwegian national cohort. *Journal of Head Trauma Rehabilitation*, *30*(4), E38–E49. https://doi.org/10.1097/HTR.0000000000000080

Auclair-Pilote, J., Lalande, D., Tinawi, S., Feyz, M., & de Guise, E. (2021). Satisfaction of basic psychological needs following a mild traumatic brain injury and relationships with post-concussion symptoms, anxiety, and depression. *Disability and Rehabilitation*, *43*(4), 507–515. https://doi.org/https://dx.doi.org/10.1080/09638288.2019.1630858

Baujat, B., Mahe, C., Pignon, J. P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials. *Statistics in Medicine*, *21*(18), 2641–2652.

Bodley-Scott, S. E. M., & Riley, G. (2015). How partners experience personality change after traumatic brain injury - Its impact on their emotions and their relationship. *Brain Impairment*, *16*(3), 205–220. https://doi.org/10.1017/BrImp.2015.22

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Complex data

structures. In *Introduction to meta-analysis*. Wiley-Blackwell Publishing Ltd.

Chaurasiya, A., Pandey, N., Ranjan, J. K., & Asthana, H. S. (2021). Neurocognitive and affective sequelae following complicated mild and moderate traumatic brain injury: A case series. *NEUROLOGY INDIA*, *69*(1), 56–61. https://doi.org/10.4103/0028-3886.310110

Corrigan, J. D., Whiteneck, G., & Mellick, D. (2004). Perceived needs following traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *19*(3).

Crawford, J. R., Garthwaite, P. H., Sutherland, D., & Borland, N. (2011). Some supplementary methods for the analysis of the Delis-Kaplan Executive Function System. *Psychological Assessment*, *23*(4), 888–898. https://doi.org/10.1037/a0023712

Curvis, W., Simpson, J., & Hampson, N. (2018). Social anxiety following traumatic brain injury: An exploration of associated factors. *Neuropsychological Rehabilitation*, *28*(4), 527–547. https://doi.org/https://dx.doi.org/10.1080/09602011.2016.1175359

Dahm, J., & Ponsford, J. (2015a). Comparison of long-term outcomes following traumatic injury: what is the unique experience for those with brain injury compared with orthopaedic injury?. *Injury*, *46*(1), 142–149. https://doi.org/https://dx.doi.org/10.1016/j.injury.2014.07.012

Dahm, J., & Ponsford, J. (2015b). Comparison of long-term outcomes following traumatic injury: What is the unique experience for those with brain injury compared with orthopaedic injury? *Injury*, *46*(1), 142–149. https://doi.org/10.1016/j.injury.2014.07.012

de Koning, M. E., Gareb, B., El Moumni, M., Scheenen, M. E., van der Horn, H. J., Timmerman, M. E., Spikman, J. M., & van der Naalt, J. (2016). Subacute posttraumatic complaints and psychological distress in trauma patients with or without mild traumatic brain injury. *Injury*, *47*(9), 2041–2047. https://doi.org/10.1016/j.injury.2016.04.036

Delmonico, R. L., Theodore, B. R., Sandel, M. E., Armstrong, M. A., & Camicia, M. (2022). Prevalence of depression and anxiety disorders following mild traumatic brain injury. *PM&R*, *14*(7), 753–763. https://doi.org/10.1002/pmrj.12657

Derogatis, L. R. (1975). *Brief Symptom Inventory*. Clinical Psychometric Research.

Derogatis, L. R. (2001). *Brief Symptom Inventory 18*. Johns Hopkins University.

Draper, K., Ponsford, J., & Schonberger, M. (2007). Psychosocial and emotional outcomes 10 years following traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *22*(5), 278–287.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I): Clinician version user's guide*. American Psychiatric Publishing, Inc.

Giustini, M., Longo, E., Azicnuda, E., Silvestro, D., D'Ippolito, M., Rigon, J., Cedri, C., Bivona, U., Barba, C., & Formisano, R. (2014). Health-related quality of life after traumatic brain injury: Italian validation of the QOLIBRI. *Functional Neurology*, *29*(3), 167–176.

Goldberg, D., Bridges, K., Duncan-Jones, P., & Grayson, D. (1988). Detecting anxiety and depression in general medical settings. *British Medical Journal*, *297*, 897–899.

Gould, K. R., Ponsford, J. L., Johnston, L., & Schönberger, M. (2011). The nature, frequency and course of psychiatric disorders in the first year after traumatic brain injury: A prospective study. *Psychological Medicine*, *41*(10), 2099–2109. https://doi.org/10.1017/S003329171100033X

Hart, T., Benn, E. K. T., Bagiella, E., Arenth, P., Dikmen, S., Hesdorffer, D. C., Novack, T. A., Ricker, J. H., & Zafonte, R. (2014). Early trajectory of psychiatric symptoms after traumatic brain injury: Relationship to patient and injury characteristics. *Journal of Neurotrauma*, *31*(7), 610–617. https://doi.org/10.1089/neu.2013.3041

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

Julien, J., Tinawi, S., Anderson, K., Frenette, L. C., Audrit, H., Ferland, M. C., Feyz, M., & De Guise, E. (2017). Highlighting the differences in post-traumatic symptoms between patients with complicated and uncomplicated mild traumatic brain injury and injured controls. *Brain Injury*, *31*(13–14), 1846–1855. https://doi.org/10.1080/02699052.2017.1346289

Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H. V. (1998). The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research*, *7*(4), 171–185. https://doi.org/https://doi.org/10.1002%2Fmpr.47

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H. J., Sheen, S. S., Hahn, S., Jang, B. H., & Son, H. J. V. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, *66*(4), 408–414. https://doi.org/10.1016/j.jclinepi.2012.09.016

Kirkham, J. J., Dwan, K. M., Altman, D. G., Gamble, C., Dodd, S., Smyth, R., & Williamson, P. R. (2010). The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*, *340*. https://doi.org/10.1136/bmj.c365

Lamontagne, G., Belleville, G., Beaulieu-Bonneau, S., Souesme, G., Savard, J., Sirois, M.-J., Giguere, M., Tessier, D., Le Sage, N., & Ouellet, M.-C. (2022). Anxiety symptoms and disorders in the first year after sustaining mild traumatic brain injury. *Rehabilitation Psychology*, *67*(1), 90–99. https://doi.org/10.1037/rep0000422

Leong Bin Abdullah, M. F. I., Ng, Y. P., & Sidi, H. Bin. (2018). Depression and anxiety among traumatic brain injury patients in Malaysia. *Asian Journal of Psychiatry*, *37*(July),

67–70. https://doi.org/10.1016/j.ajp.2018.08.017

Maestas, K., Sander, A., Clark, A., van Veldhoven, L., Struchen, M., Sherer, M., & Hannay, H. (2014). Preinjury coping, emotional functioning, and quality of life following uncomplicated and complicated mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *29*(5), 407–417. https://doi.org/10.1097/HTR.0b013e31828654b4

Mallya, S., Sutherland, J., Pongracic, S., Mainland, B., & Ornstein, T. J. (2015). The manifestation of anxiety disorders after traumatic brain injury: A review. *Journal of Neurotrauma*, *32*(7), 411–421. https://doi.org/10.1089/neu.2014.3504

Marinkovic, I., Isokuortti, H., Huovinen, A., Marinkovic, D. T., Maki, K., Nybo, T., Korvenoja, A., Rahul, R., Vataja, R., & Melkas, S. (2020). Prognosis after mild traumatic brain injury: Influence of psychiatric disorders. *Brain Sciences*, *10*(12). https://doi.org/10.3390/brainsci10120916

Mascialino, G., Canadas, V., Valdiviezo-Ona, J., Rodriguez-Lorenzana, A., Arango-Lasprilla, J. C., & Paz, C. (2022). Self-concept 6 months after traumatic brain injury and its relationship with emotional functioning. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.995436

Mikolic, A., van Klaveren, D., Groeniger, G. O., Wiegers, E. J. A., Lingsma, H. F., Zeldovich, M., von Steinbuchel, N., Maas, A. I. R., Roeters van Lennep, J. E., & Polinder, S. (2021). Differences between men and women in treatment and outcome after traumatic brain injury. *Journal of Neurotrauma*, *38*(2), 235–251. https://doi.org/10.1089/neu.2020.7228

Moum, T. (1998). Mode of administration and interviewer effects in self-reported symptoms of anxiety and depression. *Social Indicators Research*, *45*(1–3), 279–318. https://doi.org/10.1023/a:1006958100504

O'Donnell, M. L., Alkemade, N., Creamer, M. C., McFarlane, A. C., Silove, D., Bryant, R. A., & Forbes, D. (2016). The long-term psychiatric sequelae of severe injury: a 6-year follow-up study. *The Journal of Clinical Psychiatry*, *77*(4), e473-9. https://doi.org/https://dx.doi.org/10.4088/JCP.14m09721

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, *8*, 157–159.

Osborn, A. J., Mathias, J. L., & Fairweather-Schmidt, A. K. (2016). Prevalence of anxiety following adult traumatic brain injury: A meta-analysis comparing measures, samples and postinjury intervals. *Neuropsychology*, *30*(2), 247–261. https://doi.org/10.1037/neu0000221

Osborn, A. J., Mathias, J. L., Fairweather-Schmidt, A. K., & Anstey, K. J. (2017). Anxiety and comorbid depression following traumatic brain injury in a community-based sample of young, middle-aged and older adults. *Journal of Affective Disorders*, *213*, 214–221. https://doi.org/10.1016/j.jad.2016.09.045

Ouellet, M. C., Sirois, M. J., & Lavoie, A. (2009). Perceived mental health and needs for mental health services following trauma with and without brain injury. *Journal of*

Rehabilitation Medicine, *41*(3), 179–186. https://doi.org/10.2340/16501977-0306

Ponsford, J., Alway, Y., & Gould, K. R. (2018). Epidemiology and natural history of psychiatric disorders after TBI. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *30*(4), 262–270. https://doi.org/10.1176/appi.neuropsych.18040093

Ponsford, J., Nguyen, S., Downing, M., Bosch, M., McKenzie, J. E., Turner, S., Chau, M., Mortimer, D., Gruen, R. L., Knott, J., & Green, S. (2019). Factors associated with persistent post-concussion symptoms following mild traumatic brain injury in adults. *Journal of Rehabilitation Medicine*, *51*(1), 32–39. https://doi.org/10.2340/16501977-2492

Popov, N., Mercier, L. J., King, R., Fung, T., & Debert, C. T. (2022). Factors associated with quality of life in adults with persistent post-concussion symptoms. *Canadian Journal of Neurological Sciences*, *49*(1), 109–117. https://doi.org/10.1017/cjn.2021.53

Reger, M. L., Poulos, A. M., Buen, F., Giza, C. C., Hovda, D. A., & Fanselow, M. S. (2012). Concussive brain injury enhances fear learning and excitatory processes in the amygdala. *Biological Psychiatry*, *71*(4), 335–343. https://doi.org/10.1016/j.biopsych.2011.11.007

Roebuck-Spencer, T., & Cernich, A. (2014). Epidemiology and societal impact of traumatic brain injury. In M. Sherer & A. M. Sander (Eds.), *Handbook on the neuropsychology of traumatic brain injury* (pp. 3–23). Springer. https://doi.org/10.1007/978-1-4939-0784-7_1

Roozenbeek, B., Maas, A. I. R., & Menon, D. K. (2013). Changing patterns in the epidemiology of traumatic brain injury. *Nature Reviews Neurology*, *9*(4), 231–236.

Ruet, A., Bayen, E., Jourdan, C., Ghout, I., Meaude, L., Lalanne, A., Pradat-Diehl, P., Nelson, G., Charanton, J., Aegerter, P., Vallat-Azouvi, C., & Azouvi, P. (2019). A detailed overview of long-term outcomes in severe traumatic brain injury eight years post-injury. *Frontiers in Neurology*, *10*(FEB). https://doi.org/10.3389/fneur.2019.00120

Savitsky, B., Givon, A., Rozenfeld, M., Radomislensky, I., & Peleg, K. (2016). Traumatic brain injury: It is all about definition. *Brain Injury*, *30*(10), 1194–1201. https://doi.org/10.1080/02699052.2016.1187290

Schönberger, M., Ponsford, J., Gould, K. R., & Johnston, L. (2011). The temporal relationship between depression, anxiety, and functional status after traumatic brain injury: A cross-lagged analysis. *Journal of the International Neuropsychological Society*, *17*(5), 781–787. https://doi.org/10.1017/S1355617711000701

Scicutella, A. (2019). Neuropsychiatry and traumatic brain injury. *Acquired Brain Injury: An Integrative Neuro-Rehabilitation Approach., 2nd Ed.*, 227–301. https://doi.org/https://dx.doi.org/10.1007/978-3-030-16613-7_10

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *50*(Suppl 20), 22–33.

Shields, C., Ownsworth, T., O'Donovan, A., & Fleming, J. (2016). A transdiagnostic investigation of emotional distress after traumatic brain injury. *Neuropsychological Rehabilitation*, *26*(3), 410–415. https://doi.org/10.1080/09602011.2015.1037772

Silverberg, N. D., Panenka, W. J., & Iverson, G. L. (2018). Work productivity loss after mild traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *99*(2), 250–256. https://doi.org/10.1016/j.apmr.2017.07.006

Simon, K. C., Reams, N., Beltran, E., Wang, C., Hadsell, B., Maurer, D., Hillman, L., Tideman, S., Garduno, L., Meyers, S., Frigerio, R., Maraganore, D. M., Tidemanb, S., Garduno, L., Meyers, S., Frigerio, R., & Maraganore, D. M. (2020). Optimizing the electronic medical record to improve patient care and conduct quality improvement initiatives in a concussion specialty clinic. *Brain Injury*, *34*(1), 62–67. https://doi.org/10.1080/02699052.2019.1680867

Singh, R., Choudhri, K., Sinha, S., Mason, S., Lecky, F., & Dawson, J. (2019). Global outcome after traumatic brain injury in a prospective cohort. *Clinical Neurology and Neurosurgery*, *186*. https://doi.org/10.1016/j.clineuro.2019.105526

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. V. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Bernd, L. L. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. In *Archives of Internal Medicine* (Vol. 166, Issue 10, pp. 1092–1098). American Medical Association. https://doi.org/10.1001/archinte.166.10.1092

Stenberg, M., Godbolt, A. K., Nygren De Boussard, C., Levi, R., & Stålnacke, B. M. (2015). Cognitive impairment after severe traumatic brain injury, clinical course and impact on outcome: A Swedish-icelandic study. *Behavioural Neurology*, *2015*. https://doi.org/10.1155/2015/680308

Stenberg, M., Stalnacke, B. M., & Saveman, B. I. (2022). Health and well-being of persons of working age up to seven years after severe traumatic brain injury in northern Sweden: A mixed method study. *Journal of Clinical Medicine*, *11*(5). https://doi.org/https://dx.doi.org/10.3390/jcm11051306

Stocchetti, N., & Zanier, E. R. (2016). Chronic impact of traumatic brain injury on outcome and quality of life: a narrative review. *Critical Care*, *20*(1), 148. https://doi.org/10.1186/s13054-016-1318-1

Theadom, A., Parag, V., Dowell, T., McPherson, K., Starkey, N., Barker-Collo, S., Jones, K., Ameratunga, S., & Feigin, V. L. (2016). Persistent problems 1 year after mild traumatic brain injury: A longitudinal population study in New Zealand. *British Journal of General Practice*, *66*(642), e16–e23. https://doi.org/10.3399/bjgp16X683161

Tölli, A., Höybye, C., Bellander, B.-M., Johansson, F., & Borg, J. (2018). The effect of time on cognitive impairments after non-traumatic subarachnoid haemorrhage and after traumatic brain injury. *Brain Injury*, *32*(12), 1465–1476.

https://doi.org/https://dx.doi.org/10.1080/02699052.2018.1497203

Van Praag, D. L. G., Cnossen, M. C., Polinder, S., Wilson, L., & Maas, A. I. R. (2019). Post-traumatic stress disorder after civilian traumatic brain injury: A systematic review and meta-analysis of prevalence rates. *Journal of Neurotrauma*, *36*(23), 3220–3232. https://doi.org/https://dx.doi.org/10.1089/neu.2018.5759

Vikane, E., Frøyland, K., Næss, H. L., Aßmus, J., Skouen, J. S., Froyland, K., Naess, H. L., Assmus, J., & Skoue, J. S. (2019). Predictors for psychological distress 2 months after mild traumatic brain injury. *Frontiers in Neurology*, *10*. https://doi.org/10.3389/fneur.2019.00639

Villa, D., Causer, H., & Riley, G. (2020). Experiences that challenge self-identity following traumatic brain injury: A meta-synthesis of qualitative research. *Disability and Rehabilitation*, *43*(23), 3298–3314. https://doi.org/10.1080/09638288.2020.1743773

Williams, N. (2014). The GAD-7 questionnaire. *Occupational Medicine*, *64*(3), 224. https://doi.org/10.1093/occmed/kqt161

Williamson, C., & Venkatakrishna, R. (2024). Traumatic brain injury: Epidemiology, classification, and pathophysiology. In M. J. A. & R. P. Goddeau (Ed.), *UpToDate*. Wolters Kluwer.

Wise, E. K., Mathews-Dalton, C., Dikmen, S., Temkin, N., Machamer, J., Bell, K., & Powell, J. M. (2010). Impact of traumatic brain injury on participation in leisure activities. *Archives of Physical Medicine and Rehabilitation*, *91*(9), 1357–1362. https://doi.org/10.1016/j.apmr.2010.06.009

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/.

World Health Organization. (2023). *Anxiety disorders*. https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

Yilmaz, T., Roks, G., de Koning, M., Scheenen, M., van der Horn, H., Plas, G., Hageman, G., Schoonman, G., Spikman, J., & van der Naalt, J. (2017). Risk factors and outcomes associated with post-traumatic headache after mild traumatic brain injury. *Emergency Medicine Journal*, *34*(12), 800–805. https://doi.org/10.1136/emermed-2015-205429

Zahniser, E., Nelson, L. D., Dikmen, S. S., Machamer, J. E., Stein, M. B., Yuh, E., Manley, G. T., & Temkin, N. R. (2019). The temporal relationship of mental health problems and functional limitations following mTBI: A TRACK-TBI and TED Study. *Journal of Neurotrauma*, *36*(11), 1786–1793. https://doi.org/10.1089/neu.2018.6172

Zgaljardic, D. J., Seale, G. S., Schaefer, L. A., Temple, R. O., Foreman, J., & Elliott, T. R. (2015). Psychiatric disease and post-acute traumatic brain injury. *Journal of Neurotrauma*, *32*(23), 1911–1925. https://doi.org/https://dx.doi.org/10.1089/neu.2014.3569

Zhu, Y., Jin, W., Liu, H., Peng, D., Ding, Z., Tang, Z., Zhu, L., & Yu, Y. (2016). Effects of electromagnetic fields from mobile phones on depression and anxiety after titanium

mesh cranioplasty among patients with traumatic brain injury. *Brain Injury*, *30*(1), 66–73. https://doi.org/https://dx.doi.org/10.3109/02699052.2015.1089594

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370.

Zung, W. W. (1971). A rating instrument for anxiety disorders. *Psychosomatics*, *12*(6), 371–379.

Chapter 2: Detecting Feigned Impairment with the Denver Attention Test

Abstract

Background This study explores the utility of a newly developed performance validity measure (PVT) called the Denver attention Test (DAT). The DAT is a computerized PVT that evaluates the validity of responses based on accuracy and speed.

Methods A simulator design was used to examine the DATs ability to discriminate between two groups of participants who were either instructed to do their best on a neuropsychological test battery or instructed to feign cognitive impairment. The sensitivity and specificity of the DAT was validated against three well established PVTs, and the magnitude of effect of DAT failure on measures of cognitive functioning was explored.

Results Forty-two participants were randomly allocated to either the control (n = 21) or experimental (n = 21) group. Both the DAT Total Correct (TC) and Total Time (TT) domains showed excellent classification accuracy in relation to the criterion of failure on two or more established PVTs. Failure on either of the TC and TT domains was associated with a significant suppression of cognitive performance across measures assessing executive functioning, memory, attention, processing speed, and visual-motor coordination skills, with medium to very large effect sizes observed.

Conclusion The DAT is a rapid, easy to administer PVT that provides a robust measure of performance validity. It demonstrates excellent ability to detect feigned cognitive impairment in a simulator sample, though further investigation is needed with clinical groups. Providing an accurate representation of an individual's neurocognitive functioning is crucial to ensure that appropriate care is provided in clinical settings.

Introduction

Neuropsychological tests provide valuable insight to brain functioning and fulfil a critical role in diagnosis, treatment planning, and intervention following brain injury. However, neuropsychological tests are inherently dependent on the active participation of the individual being assessed and if they do not fully engage with the assessment process, it can produce results that do not accurately reflect their true cognitive abilities (Beetar & Williams, 1995). Suboptimal performance on neuropsychological tests can lead to a significant underestimation of an individual's cognitive abilities and an overestimation of cognitive impairments (Green et al., 2001). The effect of litigation and financial incentive has also been found to impact cognitive performance; in a study involving three groups of clinical patients including a mild head trauma group seeking financial compensation, a brain dysfunction group seeking compensation, and a brain dysfunction group not seeking compensation, the two compensation-seeking groups performed worse on the cognitive measures compared to the non-compensation group (Binder, 1993). Such misinterpretations about cognitive ability may affect diagnosis, referrals, intervention, and could even cause iatrogenic harm to the individual (Lippa, 2018). It may also lead to avoidable economic burdens on the healthcare system and society, while also diverting resources from others who may need them (British Psychological Society, 2021).

Various tests have been developed to help clinicians assess whether individuals are performing to their true ability. These can be standalone tests, such as the Test of Malingering Memory (TOMM; Tombaugh, 1996) or Word Choice Test (WCT; Holdnack & Drozdick, 2009), or embedded within cognitive ability measures, such as Reliable Digit Span (RDS; Boone, 2007; Greiffenstein et al., 1994). There are two types of tests, namely performance validity tests (PVTs) and symptom validity tests (SVTs). PVTs assess the credibility of an individual's performance on objective cognitive measures, whereas SVTs assess the accuracy,

or ‘truthfulness’ of an individual’s self-report symptoms and as such, are more subjective (Larrabee, 2012). PVT cutoff criteria must be carefully established to avoid incorrect assumptions about applied effort. A false positive emerges when test results indicate that an individual has not given their best effort, when in fact they have. Conversely, a false negative arises when a test indicates that the individual has given their best effort, when they have not. It is crucial that the occurrence of potential errors, particularly false positive errors, is minimised to ensure that results are accurate and reliable.

A pivotal study by Green et al. (2001) illustrated the essential role of evaluating performance levels in cognitive testing. Their research focussed on 470 individuals who were seeking compensation following brain injury and found that without the inclusion of validity testing, the severity of head injuries did not correlate with neuropsychological test outcomes, that is, there was no dose response relationship. This might be interpreted as either mild traumatic brain injury (TBI) having a similar outcome to severe TBI (Moss et al., 2003), or that the tests of cognitive ability might be insensitive to cognitive impairment following brain injury. Examination of the results, however, indicated that those who failed on the established PVT, the Word Memory Test (WMT; Green et al., 1996), performed significantly worse (more than one standard deviation lower) on cognitive function measures compared to those who passed. This was regardless of the severity of their head injuries and skewed the relationship between TBI severity and acquired cognitive impairment. When the analysis controlled for effort by removing those who failed the WMT, the anticipated relationship between head injury severity and cognitive function was found, indicating that more severe head injuries were associated with greater cognitive impairment. Notably, the data showed that suboptimal effort accounted for over four times the variation in cognitive test scores than the severity of the head injury itself. These findings demonstrate the relative effect of PVT failure on the rest of the tests of cognitive ability. There is an overlap but some degree of

dissociation between cognitive underperformance and symptom over-reporting. A study examining how PVTs and SVTs can help determine the credibility of reported symptoms and test results for mild traumatic brain injury, found that failing a PVT was linked to lower scores on cognitive tests, suggesting possible exaggeration of cognitive problems. Whereas failing an SVT was associated with higher levels of reported symptoms, indicating possible exaggeration of symptoms (Sabelli et al., 2021). Authors emphasised the importance of using both PVTs and SVTs to obtain a comprehensive understanding of an individual's cognitive and emotional state following mild TBI.

When individuals purposely do not apply full effort to neuropsychological testing, it may reflect malingering for primary gain or possibly factitious or conversion disorder. Whilst malingering is not classified as a mental disorder, it is listed among conditions that may require clinical focus in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association, 2013). The eleventh revision of the International Classification of Diseases (ICD-11; World Health Organization, 2019) classifies malingering within factors which can influence health status or contact with health services. Both the DSM-5 and ICD-11 characterize malingering as the deliberate fabrication or exaggeration of physical or psychological difficulties. They state that the reasons for malingering are typically external, such as seeking to avoid duty or work, pursuing financial compensation such as disability or personal injury claims, evading criminal proceedings, or obtaining medication. Factitious disorder on the other hand is a classified mental disorder within the DSM-5 and ICD-11, characterised by the intentional feigning, falsifying, or aggravating of physical or psychological symptoms in oneself or in another person. Contrasting with malingering, it appears without any obvious external incentives such as financial gain or avoidance of duty and rather is assumed to satisfy an internally motivated desire to be perceived as ill (World Health Organization, 2019). Conversion disorder, or

functional neurologic symptom disorder, is also classified as a mental disorder by the DSM-5 and is characterised by neurological symptoms that are inconsistent with any known neurological disease. Symptoms may affect voluntary motor or sensory functions and cannot be explained by another medical or mental disorder. The ICD-11 classifies this as a dissociative disorder related to a disconnection and disruption of normal psychological functioning, whereas the DSM-5 considers it a somatoform disorder linked to physical symptoms and health anxiety.

Failure on PVTs is not uncommon, particularly where there is an external incentive to appear impaired. This is not limited to litigation. A study by Chafetz (2008) sampled 196 adult and 96 child consecutive referrals to the United States Disability Determinations Service (DDS) who mostly alleged low cognitive functioning. The DDS Malingering Rating Scale and a validated PVT, the Test of Malingering Memory (TOMM; Tombaugh, 1996), were used to assess performance effort. Results revealed that 67.8% of adults and 60% percent of children failed at least one of the TOMM or DDS Malingering Rating Scale and 45.8% of adults failed on both measures. In another study which surveyed 131 neuropsychologists in the USA, probable suboptimal performance effort occurred among 29% of personal injury cases, 30% of disability cases, 19% of criminal cases and 8% of medical cases (Mittenberg et al., 2002). Other studies exploring performance effort have reported similar findings (Binder, 1993; Green et al., 2001; Greiffenstein et al., 1994; Griffin et al., 1996).

Research has also explored prevalence rates of PVT failure amongst clinical populations. A review by McWhirter et al. (2020) found rates to be comparable across functional disorders and other clinical conditions, raising the issue of potential false positives in bona fide clinical patients. McWhirter at al. noted widespread PVT failure across clinical conditions even when there was no clear incentive to underperform, including brain injury, psychiatric disorders, intellectual disability, degenerative brain disease, functional disorders,

and epilepsy. However, the work has been heavily criticised by neuropsychological experts in the United States and United Kingdom, and has brought into question the study's methodological rigour, lack of neuropsychological expertise, and the validity of its comparisons (Kemp & Kapur, 2020; Larrabee et al., 2020). Furthermore, findings of a recent meta-analysis of PVT failure rates in clinical populations (Roor et al., 2024) are contrary to the propositions of McWhirter et al. (2020). The review by Roor et al. found very low failure rates in patients with neurological conditions and higher failure rates in patients with mild TBI who would not be expected to demonstrate cognitive impairment, as noted by others (Green, 2011). This indicates PVT failure is not due to cognitive impairment except in extreme cases, such as demetia, and suggests there may be hidden incentives to underperfom within clinical populaitons. Additionally, Roor et al. were unable to fully exclude participants with external incentives to underperform, for example those seeking compensation, so it is likely that the false positive rate was even lower than reported. Elsewhere, reviews of the use of PVTs in clinical populations have demonstrated that other issues such as mood are not credible explanations for PVT failure (Marshall & Schroeder, 2022). Above chance failure on PVTs can also be seen in participants who may be disinterested or disengaged, for example in people completing research to obtain course credits (An et al., 2017; Roye et al., 2019).

Irrespective of the reason for failure on a PVT, an individual's cognitive ability will be underestimated. If a person scores statistically below chance, the only interpretation is that they knew the correct answers but chose to give the wrong one and appear more impaired than is the case. Below chance scores do not occur even in genuinely severely impaired groups (Olsen et al., 2019) and are therefore typically considered to be reflective of deliberate underperformance. However, because below chance scores are rare, their use lacks sensitivity to detection of suboptimal performance. Hence it has also been argued that the use of multiple PVTs provides an equivalent level of confidence in concluding deliberate intent, for example,

above chance failure on six or more PVTs has demonstrated a similar suppression of scores on measures of cognitive ability to below chance responding (Rohling et al., 2023) and as such appears indictive of deliberate underperformance. It should also be noted that whilst most individuals may provide adequate effort during testing, those who pass a PVT should not be assumed to have applied their full, complete, or best effort during the assessment (Iverson, 2006). This is because PVTs are designed to be simple, allowing even those with significant cognitive impairments to pass and prioritise specificity over sensitivity. However, interpreting scores just above chance as valid sets a very low bar for test performance. Research shows that valid PVT scores are typically much higher than chance, suggesting that higher cutoffs are necessary for accurate assessment (Erdodi, 2023).

There are various limitations within the current field of PVTs. Commercially available tests exclusively relate to verbal and visual memory and as such, there is an emerging need for a broader range of performance validity tests which focus on other areas of cognitive functioning that may further strengthen the validity of neuropsychological testing. PVTs are also vulnerable to coaching (Brennan et al., 2009), which further highlights the need for test security. Furthermore, people are often evaluated more than once, and so prior test exposure needs to be controlled for. Lastly, as engagement may vary throughout neuropsychological assessment, multiple PVTs are required to accommodate any fluctuation in performance (Boone, 2009).

PVT validation studies are typically conducted through two different approaches. Known group designs compare PVT performances between groups who are either performing optimally, as indicated by no failures on other PVT criteria, or responding suboptimally according to the same criteria. Although this approach yields findings that are more generalisable, it has a significant limitation in that there is no universally accepted ‘gold standard’ for identifying malingering. Consequently, the calibration of a new PVT depends on

the psychometric properties of the existing PVTs chosen for the study (Schroeder et al., 2021). Simulator, or analogue malingerer designs are conducted when healthy participants are instructed by researchers to act as malingerers by faking symptoms of cognitive impairment. Although simulator designs lack the external validity of real-life clinical presentations, they demonstrate strong internal validity (Rogers, 2008).

Whilst neuropsychological tests and test manuals are commercially available along with extensive test validity studies, the proper and responsible use of tests ultimately lies with the clinician (Iverson, 2006). The gravity of this is embedded within the British Psychological Society's code of ethics and conduct (British Psychological Society, 2018). The potential invalidity of cognitive test performances needs to be carefully examined as a crucial aspect of test interpretation and in light of this, the use of standardised tests to assess performance validity is recognised as an important component of neuropsychological evaluations by the British Psychological Society, the American Academy of Clinical Neuropsychology, the National Academy of Neuropsychology, and the Interorganizational Practice Committee (IOPC, a joint committee of major neuropsychological organizations which aims to improve neuropsychology practice). Subsequently, there has been a significant expansion in empirical research related to PVTs over the past couple of decades and their use has become a routine part of clinical practice for cognitive testing and evaluation. The use of multiple PVTs reduces false positive rates and increases sensitivity (Larrabee, 2003; Larrabee et al., 2020). Thus, by using multiple PVTs, clinicians can obtain a more comprehensive understanding of an individual's effort and performance reliability (Heilbronner et al., 2009). Therefore, it is considered best practice to incorporate at least two effort measures when assessing cognitive functioning (British Psychological Society, 2021; Pearson, 2009; Sherman et al., 2020).

The Denver Attention Test (DAT) is a relatively new, computerized PVT that evaluates the validity of responses based on their accuracy and speed. A preliminary

validation study in a mixed neurological sample (Reilly et al., 2021) demonstrated that the DAT possesses good specificity and moderate sensitivity for identifying suboptimal performance in neuropsychological assessment, compared to an established PVT benchmark, the Word Memory Test (Green et al., 2002). To further validate the DAT, additional testing across various clinical groups is necessary, and its effectiveness specifically for TBI patients remains to be assessed. Furthermore, cross-validation with other established PVTs will enhance the understanding of the DAT's psychometric properties.

The purpose of this study is to validate the DAT using a simulator design. This will include an experimental group in which participants will be given a scenario setting the scene for underperformance with instructions to feign cognitive impairment, and a control group where participants will be instructed to perform to the best of their ability on a neuropsychological test battery. The aims of the study are to determine the sensitivity and specificity of the DAT against three well validated PVTs, in this case the TOMM (Trial 1), Medical Symptom Validity Test (MSVT; Green, 2004), and Reliable Digit Span (RDS; Boone, 2007; Greiffenstein et al., 1994), and to establish the magnitude of effect of DAT failure on other measures of cognitive functioning. Data pertaining to mood will also be captured and examined, as previous research has found that performance on some neuropsychological measures is correlated with emotional distress (Løvstad et al., 2016; Shwartz et al., 2020). It is hoped that supporting the development of the DAT in this study will help to expand the range of PVTs which can determine performance on psychometric tests and strengthen the validity of neuropsychological assessments. Providing an accurate representation of an individual's neurocognitive functioning is crucial to ensure that appropriate care is provided when applied in clinical settings.

Methods

A neuropsychological assessment battery was administered to control and experimental group participants. Both groups received a vignette with instructions on how to perform on the test battery (see appendix 3). The vignette for the experimental condition was adapted from instructions by Hacker and Jones (2009) which asked participants to imagine they had been involved in a road traffic accident and sustained a head injury 12 months previous. The vignette outlined that since the accident they have made a good recovery, however, they are now involved in legal proceedings against the driver of the other vehicle and instructed that if they can successfully convince the examiner that they have ongoing symptoms of brain injury, they may receive a large financial sum in compensation. It warned participants not to be too obvious because if the examiner were to suspect they are not applying full effort or are exaggerating, this would jeopardise their compensation claim. The vignette detailed common symptoms of traumatic brain injury which are readily accessible through internet searches. It highlighted problems with memory, difficulty with attention and concentration, and affected speed of processing information. It also stated that difficulties with organising, planning, or completing tasks may be present. The vignette for the control group also asked participants to imagine they had sustained a head injury in a road traffic accident 12 months earlier and since made a good recovery. They were then instructed to perform to the best of their ability, with the incentive of a promotion at work. The assessment battery was administered in the same manner to both groups.

Participants

Adults over the age of 18 were eligible to take part in the study if they met the following criteria: They do not have a diagnosis of a neurological condition or learning disability; they are English speakers to a sufficient standard that would not invalidate the

standard administration of the test; and they are able to give informed consent. Recruitment took place via word of mouth and through announcements in the Psychology Department of the University of Birmingham. The largest uptake in recruitment came through the researcher's social network which aimed to capture more diversity in age and academic background. Individuals interested in taking part were provided with information about the study and were given time to consider this prior to providing consent. The researcher contacted them after 24 hours to arrange the assessment. At the point of recruitment, participants were told of their right to withdraw from the study and signed a consent form detailing their rights. They were informed if they chose to withdraw from the study, their data would be destroyed but that once the analysis had begun, it would no longer be possible to withdraw their data from the study. Demographic data captured age at testing, gender, ethnicity as defined by the participant and grouped according to 2021 Census data (Office for National Statistics, 2021), and years of education.

Materials

The following battery of measures and tests were undertaken to ensure that the DAT was administered in a manner consistent with routine neuropsychological assessment:

Criterion PVTs

Denver Attention Test (DAT; undergoing development)

The DAT is a stand-alone, forced choice PVT that consists of 16 items which are used to evaluate attention and concentration over three trials, with a total score of 48 correct responses possible. The DAT beings with an instruction phase, followed by a presentation phase in which visual target items are highlighted, and lastly an assessment phase. During the presentation phase an on-screen visual 'distractor' appears as a bouncing soccer ball at the

sides of the computer screen. Verbal instructions, administration, and scoring are completed in under four minutes when a participant provides valid responses, and it can take up to 20 minutes in cases of extreme negative response bias (Reilly et al. 2021). There are two domains which the DAT measures performance effort against including Total Correct (TC) responses and Total Time (TT) taken to complete the test. Both domains have three categorises of effort. A TT score of 97 seconds and less indicates good performance speed; 98 to 197 seconds indicates acceptable performance speed; and a cutoff of 198 seconds or more indicates delayed responding bias, with high specificity of 95% and limited sensitivity of 34%. In this study, a binary cutoff for pass or failure was obtained by recording all acceptable and good responses as Pass, and all negative or delayed responses as Fail.

Test of Malingering Memory (TOMM; Tombaugh, 1996)

The TOMM is a standalone PVT with 50 on-screen picture items that are used to assess visual recognition and whether an individual is deliberately exaggerating or falsely displaying memory problems. Findings support Trial 1 of the TOMM as a standalone PVT (Schroeder et al., 2013; Webber et al., 2018). A pass score of 42 and above in Trial 1 has demonstrated sensitivity values ranging from 62% to 66%, with 93% specificity across healthy controls, neurocognitive and psychiatric samples (Martin et al., 2020).

Medical Symptom Validity Test (MSVT; Green, 2004)

The MSVT assesses auditory memory as well as response consistency. One subtest comprising a list of 10-word pairs is used in this study to assess immediate and delayed memory recall. A score for consistency is also calculated. A participant fails the WSVT if they score below 85% on either of the immediate recall, delayed recall or consistency

domains. A pass-fail distinction across the three domains on has shown 88% sensitivity and 91% specificity (Green, 2004).

Reliable Digit Span (RDS; Boone, 2007; Greiffenstein et al., 1994)

RDS is an embedded effort measure within Digit Span of the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008). Whilst scores of 7 are cited as an indicator of potentially unreliable test results by the WAIS-IV manual, research suggests that this threshold does not consistently meet the desired 90% specificity across different patient groups (Schroeder et al., 2012). By lowering the RDS cutoff score to 6 or below, specificity remains high, but sensitivity may decrease. In a review by Schroeder et al. (2012) which included data about TBI patients, a cutoff score of 6 or below resulted in specificity 97% compared with 82% for cutoff score of 7 or below. A pass cutoff score of 7 and above is used in this study.

Mood

Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001)

The PHQ-9 is brief questionnaire which measures symptoms of depression over the past two weeks. Scores range from 0 to 27, with higher scores indicating a greater presence of depression. The PHQ-9 displays good internal consistency in an English TBI sample with a coefficient alpha of 0.88, and excellent test-retest reliability with a correlation of 0.90 (von Steinbuechel et al., 2021).

Generalised Anxiety Disorder Assessment (GAD-7; Spitzer et al., 2006)

The GAD-7 is brief questionnaire measuring symptoms of generalised anxiety over the past two weeks. Scores range from 0 to 21, with higher scores reflecting greater anxiety.

The GAD-7 demonstrates excellent internal consistency in an English TBI sample with a coefficient alpha of 0.90, and excellent test-retest reliability with a correlation of 0.91 (von Steinbuechel et al., 2021).

Measures of cognitive ability

Test of Premorbid Functioning - UK Version (TOPF-UK; Wechsler, 2011)

The TOPF-UK comprises a list of 70 words that are characterised by atypical mappings between their written and spoken forms. It supports clinicians to estimate an individual's cognitive and memory abilities prior to any injury or health condition and is co-normed with WAIS-IV and WMS-IV. Internal consistency for the TOPF is high across all age groups and clinical groups with coefficient alpha values ranging from of 0.96 to 0.99 and 0.97 to 0.99 respectively. Test-retest reliability is high to very high, ranging from 0.89 to 0.95 across age groups (Wechsler, 2011).

Logical Memory Subtest of Welscher Memory Scale (WMS-IV; Wechsler, 2009)

The logical memory subtest is used clinically to assess verbal memory through recollection of a brief story that has been read aloud to the participant. The participant is instructed to recall details about the story immediately after hearing it and again after a delay of approximately 20 minutes. Story B was used in this study. Test-retest reliability correlations for the Logical Memory subtest are good, with average correlations of 0.82 for immediate recall and 0.85 for delayed recall across ages 16 to 69 (Wechsler, 2009). The WMS-IV has demonstrated sensitivity to the effects of TBI (Carlozzi et al., 2013).

Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008)

The WAIS-IV is a measure of cognitive abilities which can identify an individual's cognitive strengths and difficulties against normed data scores. Three subtests of WAIS-IV are used in this study, including Digit Span, Symbol Search, and Coding. Digit Span is used to assess short-term and working memory through number sequence recall. Symbol Search is concerned with processing speed and an individual's ability to process nonverbal visual information quickly and accurately. Coding also measures processing speed and an individual's ability to process visual information with speed and precision. A Processing Speed Index (PSI) score is calculated from scores on Symbol Search and Coding which measures visual and motor speed. Internal consistency ranges from 0.94 to 0.98 across subtests and test-retest reliability correlations range from mid-0.70s to upper 0.80s (Wechsler, 2008). The WAIS-IV has demonstrated sensitivity to the effects of TBI (Carlozzi et al., 2015).

Delis Kaplan Executive Function System (D-KEFS; Delis et al., 2001)

The D-KEFS consists of nine individual tests which target specific aspects of verbal and non-verbal executive function. Three tests of the D-KEFS are used in this study, including Colour Word Interference, Trail Making Test, and Verbal Fluency. Colour Word Interference measures inhibition through the ability to supress automatic verbal responses. The Trail Making Test measures planning skills and cognitive flexibility during a visual-motor sequencing task. Verbal Fluency assesses letter fluency, category fluency, and category switching. The number of errors a participant makes on Colour Word Interference and Trail Making are also recorded. An Executive Functioning Index (EFI) score is also obtained from scores on all the executive tasks on the D-KEFS. The method used to calculate the EFI is described in detail by Crawford et al. (2011). Reliability coefficients generally indicate moderate to high reliability for different subtests and age groups (Delis et al., 2001). Test-retest reliability for D-KEFS variables vary significantly across tests and conditions, but

higher correlations have been reported for Trail Making (0.77), Verbal Fluency (0.80), and Colour-Word Interference (0.62–0.76) (Homack et al., 2005). The D-KEFS has demonstrated sensitivity to the effects of TBI (Hacker et al., 2024).

Procedure

Data collection took place in participants' homes and at the University of Birmingham. At the start of the experiment, participants were randomly allocated to the control or experimental condition by the flip of a coin until either group reached saturation at 21 participants. At this point all remaining participants were allocated to the remaining condition. Participants completed the assessment battery as per their respective vignette instructions. Results were not shared with participants, but they were asked if they would like a copy of the overall results once the study had concluded.

Results

Participants

Forty-two participants were randomly allocated to either the control (n = 21) or experimental (n = 21) group (descriptive statistics of participants are shown in table 11). The mean age of the control group was 39.38 years (*SD* = 12.17) and 71.4% were female. The ethnicity of the participants in the control group was 14.3% Black, Black British, Caribbean or African (Black British Caribbean, and Black British African), 4.8% Mixed or multiple ethnic groups (Mixed Black Caribbean and White), and 81% White (British, and other: European). The mean years of education for the control group was 15.19 (*SD* = 2.58).

The mean age of the experimental group was 37.86 (*SD* = 10.87) and 59.5% were female. The ethnicity of the participants in the experimental group was 9.5% Asian or Asian British (Pakistani), 9.5% Black, Black British, Caribbean or African (Black British, and Black

British Caribbean), 4.8% Mixed or multiple ethnic groups (Mixed Black Caribbean and White), and 76.2% White (British, and other: European). The mean years of education for the experimental group was 15.86 (*SD* = 2.39).

Table 11

Descriptive statistics of participants

Variable	Control group	Experimental group	Total
Gender (n, %)			
Female	15 (71.4)	10 (47.6)	25 (59.5)
Male	6 (28.6)	11 (52.4)	17 (40.5)
Age at testing (Mean, SD)	39.38 (12.17)	37.86 (10.87)	38.62 (11.42)
Ethnicity (n, %)			
Asian or Asian British	0	2 (9.5)	2 (4.8)
Black, Black British, Caribbean or African	3 (14.3)	2 (9.5)	5 (11.9)
Mixed or multiple ethnic groups	1 (4.8)	1 (4.8)	2 (4.8)
White	17 (81)	16 (76.2)	32 (78.6)
Years of education (Mean, SD)	15.19 (2.58)	15.86 (2.39)	15.52 (2.48)

There were no statistically significant differences between the control or experimental group in gender ($X^2 = 2.56$, $p = 0.11$), age at testing ($t = 0.45$, $p = 0.66$), ethnicity ($X^2 = 2.23$, $p = 0.53$), or years of education ($t = 1.09$, $p = 0.28$).

Mood

Regarding measures of mood, mean scores on the PHQ-9 and GAD-7 were calculated for the control and experimental groups. Both groups scored within the non-clinical range (i.e. none to minimal symptoms) for depression (PHQ-9) and anxiety (GAD-7), as shown in table 12. There were no significant differences between the control or experimental group on the PHQ-9 ($t = 0.52$, $p = 0.61$) or the GAD-7 ($t = 0.41$, $p = 0.68$).

Table 12

Group scores on mood measures

Mood measure	Control group		Experimental group	
	Mean	Standard Deviation	Mean	Standard Deviation
PHQ-9	4.05	2.93	4.70	4.82
GAD-7	4.15	3.05	4.65	4.49

Compliance with instruction set

To establish the participant's compliance with their respective instruction set, the respondents' classification on the established PVTs was cross tabulated with the experimental instructions that they had received. It would be expected that participants receiving control instructions should score above the cutoff on the criterion PVTs and those receiving the experimental instruction set should perform below the cutoff level. Results from the cross-tabulation analyses are shown in table 13. A false positive result would represent a score below the cutoff in a participant who has been told to perform optimally. A false negative is defined as an above cutoff score in the experimental (feigning) group.

Table 13

Compliance with instruction set relative to different criterion PVTs

Criterion	Cutoff	Instruction Set		Compliance		
		Control (best ability)	Experimental (feign impairment)	False Positive	False Negative	Overall Accuracy
	Above cutoff	*TP*	*FN*	*(FP)/(TP+FP)*	*(FN)/(TN+FN)*	*(TP+TN)/(TP+TN+FN+FP)*
	Below cutoff	*FP*	*TN*			
Fail on criterion PVT						
TOMM	Above	20	2	4.7%	9.5%	92.8%
	Below	1	19			
MSVT	Above	18	1	14.2%	4.8%	90.5%
	Below	3	20			
RDS	Above	21	5	0.0%	23.8%	88.1%
	Below	0	16			
Fail on any one PVT	Above	18	1	14.2%	4.8%	90.5%
	Below	3	20			
Fail on any two PVTs	Above	20	1	4.8%	4.8%	95.2%
	Below	1	20			
Fail on all three PVTs	Above	21	6	0.0%	28.5%	85.7%
	Below	0	15			

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

Compliance with instruction set showed 92.8% convergence with TOMM classification, 90.5% convergence with MSVT classification, and 88.1% convergence with the RDS. When compared to the criteria of failure on any one measure, an overall compliance of 90.5% was observed. When compared to the criteria of failure on any two measures, an overall compliance of 95.2% was observed and when compared to the criteria of failure on all three measures, an overall compliance of 85.7% was observed. These results indicate that the criteria of failure on any two measures provides the best balance of false positive and false negative classification errors with the highest overall accuracy and therefore will be employed as the criterion for compliance in further analysis. Subsequently, the two participants whose compliance with instruction set show discrepancy with PVT expectations were removed from the subsequent validation of the DAT.

Convergent validity of the DAT relative to failure on PVT

The two performance validity measures of the DAT, namely the Total Correct (TC) score and the Total Time (TT) scores were each examined. Table 14 shows the sensitivity and specificity of the TC scores relative to instruction set, failure on each PVT, and failure of any two PVTs. In addition, the area under the receiver operator characteristic curve (AUC) and the probability of the AUC are reported. The AUC is a measure of the DAT's ability to distinguish between Pass and Fail classifications against another PVT criterion variable. The AUC may be interpreted as representing the probability that a randomly selected Pass score will be ranked higher than a randomly selected Fail score by the model. AUC ranges from 0 to 1, with a higher value indicating better discriminatory power (Hanley & McNeil, 1982).

Table 14

Convergent validity of the DAT Total Correct Score with instruction set and various indices of PVT performance

Criterion		DAT Total Correct Score					*p* (AUC)
		Pass	Fail	Sensitivity to feigning	Specificity	Area under the curve (95% CI)	
	Best ability *Feign*	*A* *C*	*B* *D*	*D/(C+D)*	*A/(A+B)*		
Instruction Set	Best ability	19	1	75.0%	95.0%	0.984 (0.95	<0.001
	Feign	5	15			to 0.99)	
Failure on PVT							
TOMM	Pass	20	0	95.0%	100.0%	0.981 (0.95	<0.001
	Fail	1	19			to 0.99)	
MSVT	Pass	17	1	72.7%	94.4%	0.956 (0.89	<0.001
	Fail	6	16			to 0.99)	
RDS	Pass	21	3	87.5%	87.5%	0.943 (0.87	<0.001
	Fail	2	14			to 0.99)	
Fail on any two PVTs	Pass	19	1	80.0%	95.0%	0.984 (0.95 to 0.99)	<0.001
	Fail	4	16				

As shown above, the TC score has a near perfect ability to distinguish between Pass and Fail classifications against all three PVT criterion variables and against failure on any two PVTs.

Table 15 describes the sensitivity and specificity of the TT scores relative to instruction set, failure on each PVT and failure of any two PVTs.

Table 15

Convergent validity of the DAT Total Time score with instruction set and various indices of PVT performance

Criterion		DAT Total Time Score					*p* (AUC)
		Pass	Fail	Sensitivity to feigning	Specificity	Area under the curve (95% CI)	
	Best ability *Feign*	*A* *C*	*B* *D*	*D/(C+D)*	*A/(A+B)*		
Instruction Set	Best ability	20	0	30.0%	100%	0.875 (0.77 to 0.98)	<0.001
	Feign	14	6				
Failure on PVT Test							
TOMM	Pass	21	0	31.6%	100%	0.865 (0.75 to 0.98)	<0.001
	Fail	13	6				
MSVT	Pass	18	0	27.2%	100%	0.904 (0.81 to 0.99)	<0.001
	Fail	16	6				
RDS	Pass	23	1	31.2%	95.8%	0.862 (0.74 to 0.99)	<0.001
	Fail	11	5				
Fail on any two measures	Pass	20	0	30.0%	100%	0.875 (0.77 to 0.98)	<0.001
	Fail	14	6				

As seen above, the TT score demonstrates acceptable discrimination between Pass and Fail classifications against TOMM and RDS PVT criterion variables, and against failure on any two PVTs. The TT score shows highest discriminatory ability with the MSVT.

The Sensitivity and Specificity of different DAT cutoff values relative to failure on any two PVTs

Total Correct score

The relationship between sensitivity, specificity, and cutoff value on the TC score relative to the criterion of failure on any two PVTs is described in Table 16.

Table 16

The relationship between sensitivity, specificity, and cutoff value on the DAT Total Correct score relative to the criterion of failure on any two PVTs

Positive if Less Than or Equal To	Sensitivity	Specificity
7	0	1
11	0.1	1
17.5	0.15	1
22	0.2	1
23.5	0.3	1
24.5	0.45	1
27	0.5	1
29.5	0.65	1
30.5	0.7	1
33.5	0.75	1
37	0.8	0.95
39	0.85	0.95
40.5	0.9	0.95
42	0.95	0.95
43.5	0.95	0.9
44.5	0.95	0.85
45.5	1	0.85
47	1	0.75
49	1	0

To reduce the false positive error rate, it is typical to set the cutoff for PVTs to a minimum 90% specificity. As seen in the table above, based on the DAT's current lower boundary cutoff of 37 correct responses, 95% specificity and 80% sensitivity can be achieved. With the DAT's upper boundary cutoff of 40 and above correct responses, 95% specificity

and 90% sensitivity can be achieved. The optimal pass cutoff score is 42 and above, with 95% specificity and 95% sensitivity.

Total Time score

The relationship between sensitivity, specificity, and cutoff value on the TT score, relative to the criterion of failure on any two PVTs, is described in table 17.

Table 17

The relationship between sensitivity, specificity, and cutoff value on the DAT Total Time score relative to the criterion of failure on any two PVTs

Positive if Greater Than or Equal To	Sensitivity	Specificity	Positive if Greater Than or Equal To	Sensitivity	Specificity
43	1	0	78.5	0.8	0.85
45	1	0.05	79.5	0.8	0.9
46.5	1	0.1	81.5	0.75	0.9
49.5	1	0.15	85	0.7	0.9
52.5	1	0.2	91	0.65	0.9
54.5	1	0.25	101.5	0.6	0.9
56.5	1	0.3	112.5	0.6	0.95
58.5	1	0.35	119	0.55	0.95
60.5	1	0.4	121.5	0.5	0.95
61.5	1	0.45	122.5	0.5	1
62.5	0.9	0.45	129	0.4	1
63.5	0.9	0.5	138	0.35	1
64.5	0.85	0.5	164.5	0.3	1
68	0.85	0.55	222	0.25	1
71.5	0.8	0.6	257.5	0.2	1
72.5	0.8	0.65	274	0.15	1
73.5	0.8	0.75	347.5	0.1	1
76	0.8	0.8	522	0.05	1

In accordance with the DAT's upper threshold pass score of 97 seconds and less, the above data demonstrates 95% specificity, and 60% sensitivity can be achieved. The DATs lower threshold score of 198 seconds or more has high specificity but lacks sensitivity at around 25%. The optimal pass cutoff score is 79.5 seconds which is associated with a substantial increase in sensitivity of 80% with 90% specificity.

Differences in cognitive performance between PVT passers and failures

Scaled scores were calculated for participants' performance on the cognitive measures and indices as well as an index score for overall executive functioning performance. Three sets of analyses were carried out exploring the cognitive performance of participants passing and failing PVTs according to the three criteria: failure on any two PVTs; pass or failure on DAT TC score; and pass or failure on DAT TT score. Table 18 shows the comparative scores between groups based on failure of any two PVTs.

Table 18

Comparative scores between groups based on failure of any two PVTs

	Failure on any two PVTs				
	Pass		Fail		
	Mean	Standard Deviation	Mean	Standard Deviation	Cohen's d
Logical Memory Immediate	10.95_a	2.80	6.85_b	2.89	1.46
Logical Memory Delayed	12.10_a	2.65	6.30_b	2.94	2.19
Digit Span Backwards	10.80_a	3.17	4.50_b	2.84	1.99
Symbol Search	11.80_a	3.22	3.25_b	2.49	2.66
Coding	11.75_a	2.45	3.35_b	2.60	3.43
PSI	109.65_a	14.26	63.45_b	13.62	3.24
Colour Word Inhibition	11.00_a	2.08	2.10_b	2.07	4.28
Colour Word Errors	11.45_a	1.19	1.25_b	1.12	8.57
Trails Letter sequencing	10.15_a	3.22	1.95_b	2.06	2.55

	Failure on any two PVTs				
	Pass		Fail		
	Mean	Standard Deviation	Mean	Standard Deviation	Cohen's d
Trails Switching	11.30_a	1.87	2.75_b	2.51	4.57
Trails Errors	11.05_a	1.15	5.30_b	3.48	5.00
Phonemic Fluency	12.15_a	2.48	5.65_b	2.58	2.62
EFI	110.05_a	9.16	56.15_b	12.39	4.35

Note: Values in the same row which do not share the same subscript ($_{a, b}$) are significantly different at $p < .05$ in the two-sided test of equality for column means. Tests assume equal variances and are adjusted for all pairwise comparisons using the Benjamini-Hochberg correction.

As seen above, based on failure on any two PVTs, statistically significant differences were found between groups across all cognitive measures and indices, with very large effect sizes observed. The Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) is a statistical method used to control the false discovery rate when conducting multiple hypothesis tests by adjusting significance thresholds based on the number of tests performed.

Table 19 shows the comparative scores between groups based on the TC score.

Table 19

Comparative scores between groups based on failure of DAT Total Correct score

	DAT Total Correct Score				
	Pass		Fail		
	Mean	Standard Deviation	Mean	Standard Deviation	Cohen's d
Logical Memory Immediate	10.96_a	2.65	6.12_b	2.39	1.83
Logical Memory Delayed	11.74_a	2.54	5.76_b	2.99	2.35
Digit Span Backwards	10.35_a	3.24	4.00_b	2.69	1.96
Symbol Search	10.61_a	4.24	3.35_b	2.91	1.71
Coding	10.74_a	3.18	3.24_b	3.31	2.36
PSI	103.78_a	19.19	63.24_b	16.88	2.11
Colour Word Inhibition	9.65_a	3.68	2.35_b	2.94	1.98
Colour Word Errors	9.83_a	3.83	1.65_b	2.67	2.14
Trails Letter sequencing	9.13_a	4.05	1.88_b	2.18	1.79
Trails Switching	10.13_a	3.58	2.82_b	2.65	2.04
Trails Errors	10.65_a	1.70	4.82_b	3.45	3.43
Phonemic Fluency	11.61_a	2.92	5.24_b	2.25	2.18
EFI	103.17_a	18.87	55.94_b	15.65	3.02

Note: Values in the same row which do not share the same subscript ($_{a, b}$) are significantly different at $p < .05$ in the two-sided test of equality for column means. Tests assume equal variances and are adjusted for all pairwise comparisons using the Benjamini-Hochberg correction.

Based on pass or failure of the TC score, statistically significant differences with very large effect sizes observed between groups across all cognitive measures and indices.

Table 20 shows the comparative scores between groups based on TT score.

Table 20

Comparative scores between groups based on failure of DAT Total Time score

	DAT Total Time Score				
	Pass		Fail		
	Mean	Standard Deviation	Mean	Standard Deviation	Cohen's d
Logical Memory Immediate	9.21_a	3.56	7.17_a	2.71	0.57

	DAT Total Time Score				
	Pass		Fail		
	Mean	Standard Deviation	Mean	Standard Deviation	Cohen's d
Logical Memory Delayed	9.62_a	4.02	6.83_a	3.49	0.69
Digit Span Backwards	8.24_a	4.24	4.33_b	3.78	0.92
Symbol Search	8.50_a	5.01	2.00_b	.63	1.30
Coding	8.38_a	4.80	2.83_b	2.48	1.16
PSI	91.50_a	26.28	58.50_b	8.78	1.26
Colour Word Inhibition	7.26_a	4.93	2.50_b	2.81	0.97
Colour Word Errors	7.15_a	5.30	1.83_b	2.04	1.00
Trails Letter sequencing	6.74_a	4.93	2.17_b	2.86	0.93
Trails Switching	7.94_a	4.62	1.83_b	2.04	1.32
Trails Errors	8.88_a	3.64	4.17_b	2.64	1.29
Phonemic Fluency	9.53_a	4.05	5.33_b	2.66	1.04
EFI	88.24_a	28.27	54.00_b	15.63	2.19

Note: Values in the same row which do not share the same subscript ($_{a, b}$) are significantly different at $p < .05$ in the two-sided test of equality for column means. Tests assume equal variances and are adjusted for all pairwise comparisons using the Benjamini-Hochberg correction.

As can be seen above, based on pass or failure on the TT score, no statistically significant differences were observed between groups on either of the Logical Memory subtests, with medium effect sizes observed. Across all other cognitive tests and indices, there were statistically significant differences with large effect sizes between groups, though these effect sizes were not as large as those on the failure on any two PVTs, or TC score criterion.

Discussion

This study set out to investigate the utility of the Denver Attention Test (DAT), a recently developed PVT, in detecting feigned cognitive impairment. The sensitivity and specificity of the DAT was evaluated against three well validated PVTs; the TOMM (Trial 1), the MSVT, and RDS. Participants compliance with instruction set was established and results

indicated that failure on any two PVTs provides an optimal balance of sensitivity and specificity, which is consistent with current professional guidelines and practice in neuropsychological testing. The magnitude of effect of DAT failure on other measures of cognitive functioning was also examined. Failure on either of the DAT Total Correct and Total Time domains was associated with a significant suppression of cognitive performance across measures assessing executive functioning, memory, attention, processing ability and speed, and visual-motor coordination skills, with medium to very large effect sizes observed. These findings are concordant with other simulator studies that have found suppressed scores on cognitive ability measures (Kanser et al., 2017; Tombaugh, 1997).

DAT Total Correct domain

The DAT Total Correct (TC) score shows excellent classification accuracy in relation to the criterion of failure on two or more established PVTs. In terms of the utility of the DAT in detecting feigned cognitive impairment, a TC pass score of 40 was associated with specificity of 95% and sensitivity of 90%. This is in line with the TC cutoff score proposed by Reilly and colleagues (2021), however, the cutoff for a pass in this sample could be increased to a score of 42 and above and still achieve the minimum acceptable specificity of 95% whilst achieving an impressive sensitivity of 95% relative to the criterion PVTs. This suggests that the cutoff of 40 recommended by Reilly and colleagues is appropriate even when cross validated against a different set of criterion PVTs, at least in this simulator study. It is possible that further studies might indicate a higher cutoff is appropriate and can still maintain specificity, but it is hard to generalise from simulator to clinical participants. These findings also suggest that the lower cutoffs suggested by Reilly and colleagues for ‘acceptable performance’ (i.e. scores of 38 to 39) should raise significant concerns over invalid responding.

DAT Total Time domain

The DAT Total Time (TT) score has shown excellent classification accuracy in relation to the criterion of failure on two or more established PVTs. A TT cutoff score of greater than 79.5 seconds was associated with specificity of 90% and sensitivity of 80%. This suggests that the higher cutoff of 198 seconds as recommended by Reily and colleagues may be highly specific, but this may be at the cost of optimal sensitivity. The current data suggests that the original cutoff may be overly conservative and further validation with clinical data, particularly those with a high risk of reduced mental processing speed (for example, those with severe TBI) may be required to optimise the sensitivity of this PVT domain.

Utility of DAT Total Correct and Total Time domains

Although the TC and TT scores are effectively providing PVT measures that appear to participants to be measures of memory and processing speed, the question arises as to whether these a) show differential effects of domain specific cognitive performance, and b) whether they each provide additional information over the consideration of only one DAT score. In the former case, the effects sizes on memory and processing speed when considering TC and TT are both very large in this sample and suggest a general suppression of cognitive scores across domains rather than a domain specific effect. In other words, the group failing the memory-based PVT (TC) also showed poor performance on measures of speed and other cognitive domains, not just memory. There was a trend for the group failing TT to show a greater suppression on timed tasks relating to processing speed and executive functioning but non-significant effects on memory. It is possible that the time measure may be somewhat more sensitive to feigning of slowed processing, but this requires further investigation. The finding of a non-specific suppression of cognition for those failing TC is consistent with other papers

showing a general suppression across the test battery when PVTs are failed (Erdodi, 2023; Green et al., 2001).

In terms of redundancy of the DAT domains, of the DAT failures, three participants failed the TC domain but did not fail the TT domain, whilst 3 participants failed the TT domain but not the TC. Five of these participants failed at least two criterion PVTs. The remaining participant who was a false positive according to the criterion PVTs, only failed the TT domain. However, this participant did fail one criterion PVT (MSVT) and notably, despite being in the control group, scored poorly on some of the cognitive tests (for example, logical memory). Their MSVT scores were not consistent with an interpretation of failure due to severe cognitive impairment according to the test criteria (Green, 2004) with the difference between the mean of the easy and hard test components being 17.5%. Therefore, the evidence from the current data suggests that the TC and TT domains potentially provide non redundant information regarding performance validity. Upholding a separate TT domain supports other simulator design research that observed inconsistent and slowed responding as a strategy for feigning cognitive impairment, thus underscoring the potential value of embedded measures in visual tests that are sensitive to response time (Kanser et al., 2017).

Limitations and recommendations for future research

There are some limitations to this research that must be noted. Firstly, this study used a simulator design to generate data and whilst this ensured a sufficient base rate of PVT failures to assess the DAT by, the findings from this study cannot be readily generalised to real-world settings or clinical populations. Also, it is unclear from this study how the DAT may perform with ‘sophisticated malingerers’ or those who are coached on PVTs. This study also used a relatively small and opportunistic sample. Whilst there were no statistically significant demographic differences between the experimental and control groups, a larger

sample size may produce more diverse results both in terms of demographic details and in cognitive performance scores. Future research should focus on obtaining data from clinical settings, including those with varying degrees of cognitive impairment.

Conclusion

In conclusion, the DAT is a rapid, easy to administer PVT that provides a robust measure of performance validity. It has shown excellent ability to detect feigned cognitive impairment in a simulator sample and has elicited results consistent with other well validated PVTs, though further investigation is needed with clinical groups. The continued development of the DAT will help to expand the range of PVTs which can determine performance on psychometric tests and strengthen the validity of neuropsychological assessments. PVTs are essential to provide an accurate representation of an individual's neurocognitive functioning so that appropriate care can be provided in clinical settings.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association.

An, K. Y., Kaploun, K., Erdodi, L. A., & Abeare, C. A. (2017). Performance validity in undergraduate research participants: a comparison of failure rates across tests and cutoffs. *Clinical Neuropsychology*, *31*(1), 193–206. https://doi.org/10.1080/13854046.2016.1217046

Beetar, J. T., & Williams, J. M. (1995). Malingering response styles on the Memory Assessment Scales and symptom validity tests. In *Archives of Clinical Neuropsychology* (Vol. 10, Issue 1, pp. 57–72). Elsevier Science. https://doi.org/10.1016/0887-6177(94)E0005-A

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Binder, L. M. (1993). Assessment of malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology*, *15*(2),

170–182.

Boone, K. B. (2007). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. Guilford Press.

Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *Clin Neuropsychol*, *23*(4), 729–741. https://doi.org/10.1080/13854040802427803

Brennan, A. M., Meyer, S., David, E., Pella, R., Hill, B. D., & Gouvier, W. D. (2009). The vulnerability to coaching across measures of effort. *Clinical Neuropsychologist*, *23*(2), 314–328. https://doi.org/10.1080/13854040802054151

British Psychological Society. (2018). *Code of ethics and conduct*. British Psychological Society.

British Psychological Society. (2021). *Guidance on the assessment of performance validity in neuropsychological assessments*. https://www.bps.org.uk/guideline/guidance-assessment-performance-validity-neuropsychological-assessments

Carlozzi, N. E., Grech, J., & Tulsky, D. S. (2013). Memory functioning in individuals with traumatic brain injury: An examination of the Wechsler Memory Scale–Fourth Edition (WMS–IV). *Journal of Clinical and Experimental Neuropsychology*, *35*(9), 906–914. https://doi.org/10.1080/13803395.2013.833178

Carlozzi, N. E., Kirsch, N. L., Kisala, P. A., & Tulsky, D. S. (2015). An examination of the Wechsler Adult Intelligence Scales, fourth ddition (WAIS-IV) in individuals with complicated mild, moderate and severe traumatic brain injury (TBI). *Clinical Neuropsychologist*, *29*(1), 21–37. https://doi.org/10.1080/13854046.2015.1005677

Chafetz, M. D. (2008). Malingering on the social security disability consultative exam: Predictors and base rates. *Clinical Neuropsychologist*, *22*(3), 529–546. https://doi.org/10.1080/13854040701346104

Crawford, J. R., Garthwaite, P. H., Sutherland, D., & Borland, N. (2011). Some supplementary methods for the analysis of the Delis-Kaplan Executive Function System. *Psychological Assessment*, *23*(4), 888–898. https://doi.org/10.1037/a0023712

Delis, D. C., Kaplan, E., & Kramer, J. H. V. (2001). *Delis-Kaplan Executive Function System (D–KEFS) [Database record]*. PsycTESTS.

Erdodi, L. A. (2023). From “below chance” to “a single error is one too many”: Evaluating various thresholds for invalid performance on two forced choice recognition tests. *Behavioral Sciences and the Law*, *41*(5), 445–462. https://doi.org/10.1002/bsl.2609

Green, P. (2004). *Green’s Medical Symptom Validity Test (MSVT) for Microsoft windows user’s manual*. Green’s Publishing.

Green, P. (2011). Comparison between the test of memory malingering (TOMM) and the nonverbal medical symptom validity test (NV-MSVT) in adults with disability claims. *Applied Neuropsychology*, *18*(1), 18–26. https://doi.org/10.1080/09084282.2010.523365

Green, P., Allen, L., & Astner, K. (1996). *Manual for Computerised Word Memory Test*. CogniSyst.

Green, P., Lees-Haley, P. R., & Allen III, L. M. (2002). The Word Memory Test and the validity of neuropsychological test scores. In *Journal of Forensic Neuropsychology* (Vol. 2, Issues 3–4, pp. 97–124). Haworth Press. https://doi.org/10.1300/J151v02n03_05

Green, P., Rohling, M. L., Lees Haley, P. R., & Allen, L. M. V. (2001). Effort Has a Greater Effect on Test Scores Than Severe Brain Injury in Compensation Claimants. *Brain Injury*, *15*, 1045–1060. https://doi.org/http://dx.doi.org/10.1080/02699050110088254

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218–224.

Griffin, G. A. E., Normington, J., May, R., & Glassmire, D. (1996). Assessing dissimulation among social security disability income claimants. *Journal of Consulting and Clinical Psychology*, *64*(6), 1425–1430. https://doi.org/10.1037/0022-006X.64.6.1425

Hacker, D., Jones, A. C., Chan, Y. M., Yasin, E., Clowes, Z., Belli, A., Cooper, J., Bose, D., Hawkins, A., Davies, H., & Paton, E. (2024). Examining the validity of the Delis–Kaplan Executive Function System (D-KEFS) in traumatic brain injury. *Journal of Neuropsychology*, *18*(1), 81–99. https://doi.org/10.1111/jnp.12329

Hacker, V. L., & Jones, C. (2009). Detecting feigned impairment with the word list recognition of the Wechsler Memory Scale–3rd edition. *Brain Injury*, *23*(3), 243–249. https://doi.org/10.1080/02699050902748315

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29–36.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Participants1, C. (2009). American academy of clinical neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, *23*(7), 1093–1129. https://doi.org/10.1080/13854040903155063

Holdnack, J. A., & Drozdick, L. W. V. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Clinical and interpretive manual*. Pearson.

Homack, S., Lee, D., & Riccio, C. A. (2005). Test review: Delis-Kaplan executive function system. *Journal of Clinical and Experimental Neuropsychology*, *27*(5), 599–609. https://doi.org/10.1080/13803390490918444

Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, *13*(2), 77–90. https://doi.org/10.1207/s15324826an1302_3

Kanser, R. J., Rapport, L. J., Bashem, J. R., Billings, N. M., Hanks, R. A., Axelrod, B. N., & Miller, J. B. (2017). Strategies of successful and unsuccessful simulators coached to feign traumatic brain injury. *Clinical Neuropsychologist*, *31*(3), 644–653. https://doi.org/10.1080/13854046.2016.1278040

Kemp, S., & Kapur, N. (2020). Response to McWhirter et al. *Journal of Neurology, Neurosurgery & Psychiatry*.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*, 606–613.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *Clinical Neuropsychologist*, *17*(3), 410–425. https://doi.org/10.1076/clin.17.3.410.18089

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society*, *18*(4), 625–630. https://doi.org/10.1017/S1355617712000240

Larrabee, G. J., Boone, K. B., Bianchini, K. J., Rohling, M. L., & Sherman, E. M. S. (2020). Response to McWhirter et al. (2020). *Journal of Neurology, Neurosurgery, and Psychiatry*.

Lippa, S. M. (2018). Performance validity testing in neuropsychology: a clinical guide, critical review, and update on a rapidly evolving literature. *The Clinical Neuropsychologist*, *32*(3), 391–421. https://doi.org/10.1080/13854046.2017.1406146

Løvstad, M., Sigurdardottir, S., Andersson, S., Grane, V. A., Moberget, T., Stubberud, J., & Solbakk, A. K. (2016). Behavior rating inventory of executive function adult version in patients with neurological and neuropsychiatric conditions: Symptom levels and relationship to emotional distress. *Journal of the International Neuropsychological Society*, *22*(6), 682–694. https://doi.org/10.1017/S135561771600031X

Marshall, P. S., & Schroeder, R. W. (2022). Validity assessment in patients with psychiatric disorders. In R. W. Schroeder & P. K. Martin (Eds.), *Validity assessment in clinical neuropsychological practice: Evaluating and managing noncredible performance*. The Guilford Press.

Martin, P. K., Schroeder, R. W., Olsen, D. H., Maloy, H., Boettcher, A., Ernst, N., & Okut, H. (2020). A systematic review and meta-analysis of the Test of Memory Malingering in adults: Two decades of deception detection. *Clinical Neuropsychologist*, *34*(1), 88–119. https://doi.org/10.1080/13854046.2019.1637027

McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical population - a systematic review. *Journal of Neurology, Neurosurgery and Psychiatry*, *91*, 945–952.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*(8), 1094–1102. https://doi.org/10.1076/jcen.24.8.1094.8379

Moss, A., Jones, C., Fokias, D., & Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury*, *17*(5), 377–387. https://doi.org/10.1080/0269905031000070125

Office for National Statistics. (2021). *Ethnic group, England and Wales: Census 2021*.

Olsen, D., Schroeder, R., & Martin, P. (2019). Below Chance Performance of $p < .05$ or $p < .20$: Frequency of Statistically Below Chance Scores in Dementia. *Archives of Clinical Neuropsychology*, *34*(6), 835. https://doi.org/10.1093/arclin/acz035.03

Pearson. (2009). *Advanced Clinical Solutions for WAIS-IV and WMS-IV: Clinical and interpretive manual*. Pearson.

Reilly, K. J., Kalat, S. S., Richardson, A. H., & Armistead-Jehle, P. (2021). Preliminary investigation of the Denver Attention Test (DAT) in a mixed clinical sample. *Applied Neuropsychology: Adult*, *28*(2), 158–164. https://doi.org/https://doi.org/10.1080/23279095.2019.1607736

Rogers, R. (2008). *Clinical assessment of malingering and deception*. Guilford Press.

Rohling, M. L., Binder, L. M., Larrabee, G. J., & Langhinrichsen-Rohling, J. (2023). Forced choice test score of $p \leq .20$ and failures on $\geq$ six performance validity tests results in similar overall test battery means. *Clinical Neuropsychologist*, *38*(5), 1193–1209. https://doi.org/10.1080/13854046.2023.2284975

Roor, J. J., Peters, M. J. V., Dandachi-FitzGerald, B., & Ponds, R. W. H. M. (2024). Performance validity test failure in the clinical population: A systematic review and meta-analysis of prevalence rates. *Neuropsychology Review*, *34*(1), 299–319. https://doi.org/10.1007/s11065-023-09582-7

Roye, S., Calamia, M., Bernstein, J. P. K., De Vito, A. N., & Hill, B. D. (2019). A multi-study examination of performance validity in undergraduate research participants. *Clinical Neuropsychology*, *33*(6), 1138–1155. https://doi.org/10.1080/13854046.2018.1520303

Sabelli, A. G., Messa, I., Giromini, L., Lichtenstein, J. D., May, N., & Erdodi, L. A. (2021). Symptom versus performance validity in patients with mild TBI: Independent sources of non-credible responding. *Psychological Injury and Law*, *14*(1), 17–36. https://doi.org/10.1007/s12207-021-09400-6

Schroeder, R. W., Boone, K. B., & Larrabee, G. J. V. (2021). Design methods in neuropsychological performance validity, symptom validity, and malingering research. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (2nd ed., pp. 11–33). Guilford Press.

Schroeder, R. W., Buddin, W. H., Hargrave, D. D., VonDran, E. J., Campbell, E. B., Brockman, C. J., Heinrichs, R. J., & Baade, L. E. (2013). Efficacy of Test of Memory Malingering trial 1, trial 2, the retention Trial, and the Albany Consistency Index in a criterion group forensic neuropsychological sample. *Archives of Clinical Neuropsychology*, *28*(1), 21–29. https://doi.org/10.1093/arclin/acs094

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span: A Systematic Review and Cross-Validation Study. *Assessment*, *19*(1), 21–30. https://doi.org/10.1177/1073191111428764

Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, *35*(6),

735–764. https://doi.org/10.1093/arclin/acaa019

Shwartz, S. K., Roper, B. L., Arentsen, T. J., Crouse, E. M., & Adler, M. C. (2020). The behavior rating inventory of executive function® - adult version is related to emotional distress, not executive dysfunction, in a veteran sample. *Archives of Clinical Neuropsychology*, *35*(6), 701–716. https://doi.org/10.1093/arclin/acaa024

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Bernd, L. L. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. In *Archives of Internal Medicine* (Vol. 166, Issue 10, pp. 1092–1098). American Medical Association. https://doi.org/10.1001/archinte.166.10.1092

Tombaugh, T. N. (1996). *Test of memory malingering*. Multi-Health Systems Inc.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, *9*(3), 260–268. https://doi.org/https://doi.org/10.1037/1040-3590.9.3.260

von Steinbuechel, N., Rauen, K., Bockhop, F., Covic, A., Krenz, U., Plass, A. M., Cunitz, K., Polinder, S., Wilson, L., Steyerberg, E. W., Maas, A. I. R., Menon, D., Wu, Y. J., & Zeldovich, M. (2021). Psychometric characteristics of the patient-reported outcome measures applied in the center-tbi study. *Journal of Clinical Medicine*, *10*(11), 1–34. https://doi.org/10.3390/jcm10112396

Webber, T. A., Bailey, K. C., Alverson, W. A., Critchfield, E. A., Bain, K. M., Messerly, J. M., O'Rourke, J. J. F., Kirton, J. W., Fullen, C., Marceaux, J. C., & Soble, J. R. (2018). Further validation of the Test of Memory Malingering (TOMM) trial 1 performance validity index: Examination of false positives and convergent validity. *Psychological Injury and Law*, *11*(4), 325–335. https://doi.org/10.1007/s12207-018-9335-9

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV)*. PsycTESTS.

Wechsler, D. (2009). *Wechsler Memory Scale* (4th ed.). Pearson.

Wechsler, D. (2011). *Test of Premorbid Functioning - UK Version*. Pearson.

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/.

Chapter 3: Press Release for the Literature Review

A decade of research reveals enduring high prevalence of anxiety following traumatic brain injury

A review examining prevalence rates of anxiety following traumatic brain injury (TBI) has found that people who have sustained a TBI are almost six times more likely than the general population to experience clinically significant anxiety. This means anxiety that is so intense or persistent that it interferes with their daily life and well-being and may require support from a mental health professional. Furthermore, it found that anxiety following TBI may last for many years with levels especially high five years or more after the injury was sustained.

The review which formed part of a doctoral thesis at the University of Birmingham, collated data from studies which have been published over the past ten years. This included 33 different studies with a total of 12,063 adults who had experienced a TBI at some point in their lives. The huge amount of data collected in the study enabled researchers to look at trends across different factors that might influence anxiety, such as how much time has passed since the injury, how anxiety was assessed, the severity of the injury, and whether the person had a previous diagnosis of an anxiety disorder.

Lead author of the review, Emma Johnson, explained: “It is not surprising to learn that people feel anxious after they have sustained a head injury, as living with the effects of the injury or the circumstances in which the injury took place may understandably cause distress. What is interesting to learn is that anxiety remains just as prevalent as time goes on. You might think that as someone gets used to life after their injury, the effects of anxiety might be less apparent. This study shows that’s not the case, and that the psychological impact of their injury might endure for many years after.”

Findings from the review suggest that the high prevalence rate of anxiety following TBI is not affected by how severe a person’s injury is, be that mild, moderate, or severe. Nor

does it seem to be affected by whether a person had a previous diagnosis of anxiety disorder prior to their brain injury.

The review also looked at rates of generalised anxiety disorder (GAD), which is a specific anxiety disorder where worries about everyday things like family matters, health issues, money, or work, interfere with a person's life to such an extent that it makes it difficult to think about anything else (World Health Organisation, 2019). Whilst experts estimate that globally, around four in every one hundred people have an anxiety disorder (World Health Organization, 2023), results from this review found that the rates of GAD were quadrupled following TBI.

Emma Johnson added: "What's also interesting about this study is that self-report measures of anxiety produced similar results to those where people were assessed by a clinician. Self-report measures can save time and resources compared with clinical interviews, so this is encouraging to know because often, clinical interviews are viewed as superior. Self-report measures can be much easier for people to complete, whether that's over the phone, by post, in person, or online."

Findings from this review highlight the need for support following TBI that considers both the physical and psychological aspects of a person's injury. Emma Johnson explained: "Whilst it is well known that TBI may cause problems with brain functioning, including memory, attention, thinking speed, and decision-making skills, the emotional impact of TBI must also be considered. The assessment of a person's mental health needs following TBI is essential, and clinicians can help play a vital role in a person's recovery following traumatic brain injury."

References

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/.

World Health Organization. (2023). *Anxiety disorders*. https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

Chapter 4: Press Release for the Empirical Research Paper

New research reminds us why cognitive test results should not be taken at face value

Cognitive assessments are essential to understanding how a person's brain is functioning following injury through accident or illness. However, results from these tests are highly dependent on how much a person engages with the testing process. If a person, for whatever reason, is not performing to their true ability, the results of the cognitive assessments are difficult to interpret and may be invalid. For this reason, special types of tests called performance validity tests (PVTs) are used to help determine whether someone's performance on cognitive assessments is valid or not.

A collaborative research team at the University of Birmingham and the Queen Elizabth Hospital in Birmingham have been trialling a new PVT called the Denver Attention Test (DAT). The DAT is a computerised test that is rapid and easy to administer and measures the accuracy and speed of a person's performance on the test. The DAT is designed to be implemented as part of a series of cognitive assessments that assess brain functioning such as memory, attention, speed of processing information, and visual-motor skills. If the DAT indicates that a person is not performing to their true ability, their scores on the series of cognitive tests would then have to be interpreted with caution.

Whilst the DAT has been able to detect deliberate underperformance, it cannot provide information about a person's motives. There are many reasons a person might not perform to their true ability, from avoiding certain responsibilities to seeking financial compensation, or due to psychological difficulties. Clinicians using the DAT might therefore be able to detect when someone is underperforming, but they cannot conclusively state the reasons for their behaviour. To help understand a person's performance on the DAT and other cognitive measures, literature on brain functioning can be used by clinicians to help interpret their performance.

Lead author of the study, Emma Johnson said: “Without some measure of performance validity, we would be taking a person’s scores on cognitive measures on face value. If these are not truly reflective of their cognitive ability, it may have ramifications in terms of diagnosis and appropriate care post-injury. Performance validity tests like the DAT are crucial in supporting clinicians to build the most accurate picture possible about how someone is impacted by their injury and how clinicians may best support them.”

On its own, the DAT appears to be a good indicator of whether someone’s performance is valid, but professional guidelines around cognitive testing indicate that two of these types of PVTs are more reliable than one alone in determining whether performance is valid or not. As such, results from the DAT should be interpreted in combination with at least one other well-established PVT.

The study highlights the DAT’s potential as a rapid, easy-to-administer PVT that delivers a reliable measure of performance validity. Its ability to accurately detect deliberate underperformance may be valuable in settings where the outcomes of cognitive assessments can have significant implications. By ensuring that assessments of cognitive functioning are accurate, healthcare providers can offer more appropriate and tailored care to their patients. While the results from this study are promising, authors note that more studies are needed with different patient groups to confirm how well the DAT works across a variety of people and clinical settings.

Appendices

Appendix 1: Letter from ethics committee granting full ethical approval for the research

UNIVERSITY OF BIRMINGHAM

Dear Carl Krynicki

RE: Denver Attention Test study

Application for Ethical Review: ERN_1661-Dec2023

Thank you for your application for ethical review for the above project, which was reviewed by the Science, Technology, Engineering and Mathematics Committee.

On behalf of the Committee, I confirm that this study now has ethical approval.

Any adverse events occurring during the study should be promptly brought to the Committee's attention by the Principal Investigator and may necessitate further ethical review.

Please ensure that the relevant requirements within the University's Code of Practice for Research and the information and guidance provided on the University's ethics webpages (available at https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Links-and-Resources.aspx) are adhered to.

Please be aware that whilst Health and Safety (H&S) issues may be considered during the ethical review process, you are still required to follow the University's guidance on H&S and to ensure that H&S risk assessments have been carried out as appropriate. For further information about this, please contact your School H&S representative or the University's H&S Unit at healthandsafety@contacts.bham.ac.uk.

Kind regards,

The Co-Chairs of the Science, Technology, Engineering and Mathematics Committee

E-mail: ethics-queries@contacts.bham.ac.uk

Appendix 2: Overlapping samples combined into independent studies

Study authors in meta-analysis	Papers with overlapping samples	Study name (if applicable)	Notes
Alway et al. (2016)	Alway et al. (2016) Carmichael et al. (2023) McKensie et al. (2018)	Longitudinal head injury outcomes study. (Ponsford Group D in original meta-analysis)	Data reported at one year post injury in original meta-analysis so years 2+ included in update.
Anke et al. (2015)	Anke et al. (2015) Soberg et al. (2017)		
de Koning et al. (2016)	Abdulle et al. (2018) de Koning et al. (2016) van der Horn et al. (2016)	UPFRONT study	

Study authors in meta-analysis	Papers with overlapping samples	Study name (if applicable)	Notes
Mikolić et al. (2021)	Mikolić et al. (2021) Voormorlen et al. (2020) Wang (2021)	CENTER-TBI study	
Shields at al. (2016)	Nalder et al. (2016) Shields et al. (2016)	Longitudinal study of hospital transition outcomes.	
Silverberg et al. (2018)	Silverberg et al. (2018a) Silverberg et al. (2018b)	n/a	
Singh et al. (2019)	Singh et al. (2019a) Singh et al. (2019b)	SHEFBIT study	
Stenberg et al. (2022)	Stålnacke et al. (2019) Stenberg et al. (2015) Stenberg et al. (2022)	ProBrain study	
Theadom et al. (2016)	Barker-Collo et al. (2015) Starkey et al. (2022) Theadom et al. (2016)	BIONIC study	

Appendix 3: Vignettes with instructions on how to perform on the test battery

Vignette 1

I would like you to pretend that you were involved in a road traffic accident 12 months ago. In this accident, another driver collided with your vehicle and consequently you hit your head against a windscreen. Following this, you briefly lost consciousness and were taken to hospital. Gradually, over the next few months, you started to recover and think you have now made a good recovery.

Imagine that you are now involved in legal proceedings against the driver of the other vehicle. If the injuries sustained by you are found to have resulted from the accident, you will obtain a very large financial sum in compensation. As part of the assessment for compensation you are required to undertake psychological testing to identify the difficulties you may have suffered as a result of the accident in question.

You are concerned that, because of your recovery your claim will be negatively affected. You have, therefore, decided to perform below your actual level of ability. If you can successfully convince the examiner that you have ongoing symptoms of brain injury, you

are likely to get a very large sum of money, but you must try not to be too obvious because if the examiner suspects you are not applying full effort or are exaggerating this will jeopardise your compensation claim.

You have searched the internet for symptoms commonly associated with a traumatic injury to the head. You have found that these can include problems with memory, difficulty with attention and concentration, and that speed of processing information can also be affected. Difficulties with organising, planning, or completing tasks may also be present.

Vignette 2

I would like you to pretend that you were involved in a road traffic accident 12 months ago. In this accident, another driver collided with your vehicle and consequently you hit your head against a windscreen. Following this, you briefly lost consciousness and were taken to hospital. Gradually, over the next few months, you started to recover and think you have now made a good recovery.

Imagine that you are now applying for a promotion at work. As part of the recruitment process, you are required to undertake psychological testing to identify your strengths and abilities. You wish to demonstrate that you have not been adversely impacted by the accident and are keen to perform on the tests to the best of your abilities.