**ACCEPTED VERSION**

Scott, Julia E.; Mathias, Jane Leanne; Kneebone, Anthony C.
Postoperative cognitive dysfunction after total joint replacement in the elderly: A meta-analysis
Journal of Arthroplasty, 2013; InPress

Crown Copyright © 2013

**NOTICE**: this is the author's version of a work that was accepted for publication in *Journal of Arthroplasty*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Arthroplasty, 2013; InPress.

DOI: 10.1016/j.arth.2013.06.007

**PERMISSIONS**

http://www.elsevier.com/journal-authors/policies/open-access-policies/article-posting-policy#accepted-author-manuscript

**Elsevier's AAM Policy:** Authors retain the right to use the accepted author manuscript for personal use, internal institutional use and for permitted scholarly posting provided that these are not for purposes of **commercial use** or **systematic distribution**.

Elsevier believes that individual authors should be able to distribute their AAMs for their personal voluntary needs and interests, e.g. posting to their websites or their institution's repository, e-mailing to colleagues. However, our policies differ regarding the systematic aggregation or distribution of AAMs to ensure the sustainability of the journals to which AAMs are submitted. Therefore, deposit in, or posting to, subject-oriented or centralized repositories (such as PubMed Central), or institutional repositories with systematic posting mandates is permitted only under specific agreements between Elsevier and the repository, agency or institution, and only consistent with the publisher's policies concerning such repositories.

29 October 2013

http://hdl.handle.net/2440/80500

# Postoperative Cognitive Dysfunction after Total Joint Replacement in the Elderly: A Meta-Analysis

Julia E. Scott, B Psych (hons)[1]

Jane L. Mathias, BA, BA (hons), PhD, MAPS [1]

Anthony C. Kneebone, BA, Dip App Psych, MA, PhD, FAPS [2]

1. School of Psychology, University of Adelaide, Adelaide, South Australia, Australia, 5005.

2. Department of Clinical Psychology, Flinders Medical Centre, School of Medicine, Bedford Park, South Australia, Australia, 5042.

Please address all correspondence to:
Julia Scott
School of Psychology
University of Adelaide
Adelaide, South Australia
AUSTRALIA, 5005
Phone: +61 8 8313 5266
Email: julia.scott@adelaide.edu.au

# Postoperative cognitive dysfunction after total joint replacement in the elderly: A meta-analysis

## *Abstract*

This meta-analysis consolidated the research on post-operative cognitive dysfunction (POCD) following total joint arthroplasty (TJA). Data from 17 studies that assessed cognition pre- and post-surgery in TJA patients alone (15 studies) or matched TJA and control groups (2 studies) were analysed. Results were grouped by cognitive domain (memory, attention, language, speed, general cognition) and follow-up interval (pre-discharge, 3-6 months post-surgery). The TJA data revealed small declines in reaction time and general cognition pre-discharge, but no evidence of decline 3-6 months post-surgery. Very limited TJA and Control data indicated no group differences in the changes to performance over time; however, the TJA group was cognitively compromised pre- and post-surgery compared to Controls. Further appropriately controlled research is required to clarify whether POCD commonly occurs after TJA.

*Keywords:* Total joint replacement, post-operative cognitive dysfunction, elderly, meta-analysis, outcomes

## Introduction

Total joint arthroplasty of the hip and knee (TJA) is among the most common major surgeries performed on older adults [1]. The number of TJAs performed each year has increased substantially over the last decade [2] and this trend is predicted to continue as our population ages. TJAs are usually performed to treat damage caused by osteoarthritis [2], which is common among older adults, and typically yield good surgical outcomes as they markedly improve pain, physical function, and have few medical complications [3]. Although surgically successful, patients may still experience short- or long-term post-operative cognitive dysfunction (POCD) [4-6], which is a subtle form of cognitive decline that can develop after surgery and affect multiple cognitive domains, particularly in the elderly [7].

There are multiple theories regarding the cause of POCD. One is that it results from intra-operative microemboli that travel to the brain. These emboli are thought to be released when the artificial prosthesis is inserted or the tourniquet that is used during surgery is removed [8]. Other potential causes of POCD include general anaesthesia and/or postoperative analgesia [9]; although research has consistently failed to find a relationship between general anaesthesia and POCD after TJA [5, 10, 11]. Furthermore, a systematic review that investigated the potential role of analgesia in post-operative cognitive problems confirmed that POCD was not related to either the type of analgesia or its method of administration [9]. Thus, the underlying cause of POCD has yet to be established.

The actual incidence of POCD after TJA is presently unclear, with some studies reporting substantial rates [4, 5, 12-14] and others failing to find evidence of cognitive dysfunction [11, 15]. Moreover, the incidence rates reported by those studies that did find evidence of POCD following TJA are highly variable, ranging from 16% to 45% [5, 13], with both rapid recovery [16] and chronic dysfunction [4] also noted.

Some of the variability in these research findings may result from between-study methodological differences. For example, the existing POCD studies differ in terms of their mean age, sample sizes, sample composition (i.e. TJA patients only versus partial/revision procedures plus TJA), research design (i.e. TJA group only versus TJA and Control groups), follow-up interval, and the definition of clinically significant change (i.e. use of reliable change indices versus cut-off scores). There are also differences in how cognition is measured, as both detailed cognitive batteries [4, 5, 14] and basic screening tools, such as the Mini Mental Status Examination (MMSE) [17, 18], have been used. Moreover, differences in the ability of individual tests to detect subtle cognitive dysfunction may impact on the reported incidence, severity, and duration of POCD.

Importantly, research suggests that POCD may negatively impact on the post-surgical quality of life of patients and their families. While this area remains under-researched in TJA patients, there is comparable evidence in cardiac and general hospital inpatients suggesting that POCD is associated with enduring negative effects, even after controlling for potential confounding factors, such as age and comorbid medical conditions [19, 20]. For example, cardiac patients who developed cognitive problems within six weeks of their surgery also experienced a range of other problems one year later [19]. Specifically, they had reduced functional capacity, were more limited by their symptoms (e.g. shortness of breath interfered more with their daily activities), and reported more cognitive difficulties [19]. Based on research conducted with general surgery patients, POCD may also have significant financial implications for both TJA patients and the wider community [21, 22]. For instance, patients who developed POCD are reportedly more likely to leave the labor market prematurely and to spend more time on welfare [21], and require more assistance with their activities of daily living [22]. While research of this type has yet to be conducted with TJA patients, it might be expected that TJA patients with POCD would be similarly affected.

Despite the high incidence of TJA in the older population and the potential burden associated with the development of POCD, the nature and extent of POCD after TJA remains poorly understood. There is little consistency in the research findings, and the literature specifically relating to POCD after TJA has not been reviewed, either qualitatively or quantitatively. Rather, current reviews of post-surgical cognitive outcomes have either focused on cardiac surgery [23, 24] or have combined data from patients that have undergone different forms of surgery [25, 26]. The absence of an over-arching analysis of existing research on POCD following TJA represents a major obstacle to our understanding of the incidence and severity of these problems. The current meta-analysis synthesised the available research data in order to provide this information.

## Methods

*Search strategy and selection criteria.*

Comprehensive searches of the PubMed, PsycINFO, Embase and Scopus electronic databases were undertaken to locate all studies that assessed cognition among older adults after TJA that were published between January 1980 and August 2012. A complete list of the search terms is provided in Appendix A.

To be eligible for inclusion, studies had to have: (1) included a surgical group who underwent TJA of the hip or knee; (2) examined participants who were over 50 years of age (or mean age minus 1 SD ≥ 50 years); (3) assessed cognition using standardised neuropsychological tests (excludes self-report measures and clinician ratings); (4) performed pre- and post-surgical cognitive assessments of either one (TJA) or two (TJA, Control) samples; (5) completed at least one post-surgical assessment 24 hours or more after surgery; (6) provided data that would allow for the computation of an effect size (e.g. proportions, means and SDs, or exact *t*-values); (7) assessed participants who were not reported to have had a neurological (e.g. dementia) or medical condition that may have impacted on cognition; (8) a sample size that was greater than one (excludes case studies); and (9) been published in a journal in English.

Studies were deemed eligible for inclusion if they included patients who had ‘elective’ hip and knee surgery. Although this term can be used to refer to partial and revision procedures (in addition to TJA), these procedures usually only constitute a small number of elective surgery patients [2]. Therefore, where TJA data were not reported separately, studies that assessed samples of elective surgery patients were assumed to consist primarily of TJA patients.

If a study that was published within the previous decade did not provide the necessary data, but was otherwise eligible, the corresponding author was contacted by email to request these data. The authors of eleven studies were contacted for this purpose [4, 5, 12-14, 27-31], five of whom provided the requisite data [12-14, 27, 29].

The original literature search was kept broad in order to capture the maximum number of potentially relevant papers and identified 1,312 studies (excluding duplicates). An examination of the titles and abstracts of these papers revealed that approximately 65% were not relevant to either TJA or POCD. A further 25% were relevant to TJA only, while approximately 5% addressed POCD, but not in a TJA sample. The full-texts of the remaining 5% were screened using the inclusion criteria to determine their eligibility, with 19 studies being eligible for inclusion. The most common reasons for the exclusion of studies were a failure either to provide the data for TJA patients separately from that of other surgical patients or to conduct a post-surgical cognitive assessment (i.e. only pre-surgical assessments completed). Meta-analyses assume that the data from different studies are independent and, therefore, that each sample only contributes once to the calculation of a mean effect-size [32]. If sample independence could not be established through the information provided in the publication, the corresponding authors were contacted by email for confirmation. When samples overlapped, the respective studies were combined and treated as one. To this end, the data from two studies by Evered et al [5, 33] were combined, as were the data from two studies by Dupplis and Wikblad [34, 35]. Therefore, the data from a total of 17 independent studies were analysed in this study.

### *Research design and data preparation*

Four of the 17 studies used an experimental design to examine whether different types of anaesthesia resulted in different cognitive outcomes following TJA [6, 10, 11, 36]. Only the ‘standard care’ patients (the control group) from Cheng et al. [36] could be included in the current meta-analysis, as the experimental group received non-standard treatment. In contrast, both the experimental and surgical control groups from Jones et al. [11], Nielson et al. [10] and Williams-Russo et al. [6] received standard care TJA; consequently these groups were combined for current purposes, and means and SDs for the total TJA sample calculated. In addition, two studies provided mean (and SDs) cognitive scores for specific subgroups; namely the presence/absence of POCD [37] or post-surgical delirium [27]. As these subgroups were not required for the current meta-analysis, their data were combined to provide an overall mean and SD for the entire sample. Thus, the data that was extracted from these six studies equated to that of a single-sample pre- and post-surgery design, and hereafter will be referred to as such.

Of the 17 studies that assessed cognition after TJA, 15 used a single-sample pre- and post-surgery design [6, 10-13, 15-18, 27, 29, 35-38] and only two used a two-sample (TJA and Controls) pre- and post-surgery design [5, 14]. Given that very few studies used the latter design, the TJA data from these two studies were additionally treated as if they came from a single-sample pre- and post-surgery design (i.e. TJA data analysed, control group data excluded) and tabled with the other studies that used this design for comparative purposes.

Cognition was assessed using a wide variety of different tests, and many studies used either the same test or a close variant (e.g. Grooved Pegboard Task and the Purdue Pegboard Task). When tests were deemed to measure the same construct, they were analysed together and given a more generic label (e.g. motor speed). For reporting purposes, all tests were grouped into five broad cognitive domains, based on those identified by Lezak et al. [39]: memory, attention, language, motor and processing speed, and general cognition.

Studies also varied considerably with regard to the interval that elapsed between the TJA surgery and the follow-up cognitive assessment; ranging from one day to 12 months. For current purposes, all assessments were combined into one of two follow-up intervals: pre-discharge (within one week of surgery) and longer follow-up (3 to 6 months post-surgery). With two exceptions, these groupings captured all of the available data, with only the six-week follow-up from Patel et al. [15], and one-year follow-up from Stockton et al. [38] being excluded. Both of these studies conducted two cognitive assessments that fell within the longer follow-up interval that was used here (3-6 months post-surgery), but each study can only contribute one effect to the calculation of a mean. Therefore, only the three-month data from Patel et al. [15] and the 6-month follow-up data from Stockton et al. [38] were used, as these assessments were more comparable to the follow-up intervals of the other studies.

Surgical procedures can vary in terms of the type of implant that is used (i.e. cemented or uncemented/press-fixed) and whether or not a tourniquet is used in total knee replacements, which may contribute to surgical outcomes. Information relating to these two variables was obtained from each study. Unfortunately, nine studies did not report this information [5, 10, 11, 16, 17, 27, 35, 38, 40], three used a combination of methods but did not provide separate data for each [13, 15, 18], and the remaining five studies reported using only one of these subgroups (cemented implants [6, 12, 29, 37], tourniquets [12, 14, 37]). Therefore, it was not possible to examine the impact of these two variables on cognitive outcome.

### *Effect size calculations and analyses*

Contrary to expectation, it was not possible to examine the incidence of POCD because the studies that provided these data used different criteria to define dysfunction, which would significantly impact on incidence rates [41]. Effectively, the incidence rates from these studies were not comparable and could not meaningfully be meta-analysed [24, 32].

Group data (means and SDs) were examined using Cohen's *d* effect sizes. When a study used a single-sample pre- and post-surgery design, a variant of Cohen's *d* ($d_{\mathrm{RM}}$) was calculated [32], with a negative $d_{\mathrm{RM}}$ indicating a decline in cognitive performance between baseline and follow-up. When a two-sample (TJA and Controls) pre- and post-surgery design was used, an independent-groups repeated measures Cohen's *d* ($d_{\mathrm{IGRM}}$) was calculated [42]. A negative $d_{\mathrm{IGRM}}$ indicates that there was a greater decline in cognitive performance between baseline and follow-up in the TJA group, compared to the controls or, alternatively, the cognitive performance of the Control group improved more than the TJA group when they were re-tested. A $d_{\mathrm{RM}}$ or $d_{\mathrm{IGRM}}$ of .2, .5 and .8 equates to a small, medium and large effect size, respectively [43].

The reliability of an effect is influenced by its sample size, consequently it is recommended that the effect sizes from individual studies be weighted prior to calculating a mean effect [32]. The inverse variance method, which is the inverse of the squared standard error, is often used for this purpose but requires the correlation ($r$) between the baseline and follow-up scores [32, 42]. Unfortunately, no study reported these data. Three alternatives to $r$ were identified: (1) published test-retest reliability coefficients [32]; (2) an estimate of $r$, based on the results of $t$-tests comparing pre- and post-surgery scores [42]; or (3) a single test-retest reliability coefficient for all tests, based on the minimum acceptable level of reliability (Lipsey, personal communication, 2012). The first two options were not possible because the data were not consistently available, either in the literature or individual studies, respectively; leaving only the third option. To this end, a test-retest reliability coefficient of .7, which is considered the minimum acceptable reliability coefficient for published psychological tests [44], was used when calculating the weights for individual effect sizes.

Weighted mean effect sizes and ninety-five percent confidence intervals (95% CIs) were then calculated. If the 95% CIs include zero, it suggests that there is no significant difference between the cognitive performance of the TJA group pre- and post-surgery ($d_{\mathrm{RM}}$) or that there is no significant difference between the pre- and post-surgery test score changes of the TJA and controls groups ($d_{\mathrm{IGRM}}$).

A random-effects model was used because there was heterogeneity among individual results and it was likely that a range of uncontrolled variables would impact on the effect sizes (e.g. demographic variables, specific surgical techniques, length of hospital admission) [32]. A random-effects model assumes that the effect sizes for individual studies vary due to both random error (unidentified sources of error) and normal sampling error, and weights individual effect sizes to counteract these two sources of error.

One limitation of meta-analyses is their susceptibility to publication bias, whereby significant results are more likely to be published, potentially skewing the findings of a meta-analysis [32]. Fail-safe *N* statistics ($N_{\mathrm{fs}}$), which estimate the number of unpublished studies with non-significant results that would be required to reduce an effect to an inconsequential size ($d = .1$ for current purposes), were calculated to address this problem [32]. When the $N_{\mathrm{fs}}$ for an effect size was higher than $N_{\mathrm{studies}}$, it was considered unlikely that there would be sufficient unpublished studies with non-significant findings to draw the current finding into question.

Effect sizes are reported for all tests that were used by at least two studies, either at the early (pre-discharge) or later (3-6 months post-surgery) follow-up interval, and were interpreted to suggest that POCD occurred following TJA if: (1) $d_{\mathrm{RM}}$ or $d_{\mathrm{IGRM}} \leq -.2$ (i.e., at least a small negative effect, indicating decline); (2) the 95% CIs did not include zero (i.e., statistically significant effect); and *(c)* the $N_{\mathrm{fs}} > N_{\mathrm{studies}}$ (i.e., it was unlikely that there would be this number of unpublished studies with very small effects, relative to the number that had been published). As the $d_{\mathrm{RM}}$ measures change in the cognitive performance of TJA patients over time and the $d_{\mathrm{IGRM}}$ measures whether the cognitive changes over time differed between patients and controls, the data for these two effect sizes are reported separately, as are the data for the pre-discharge and longer post-operative outcomes.

## Results

*Participants.*

The cognitive outcomes of 1,089 TJA patients and 89 healthy controls were assessed by 17 studies that were included in this meta-analysis. Summary demographic and surgical data for these samples are provided in Table 1, where it can be seen that the majority of participants were females, in their late-60s to mid-70s, who underwent total hip replacement. Unfortunately, too few studies provided data on marital status ($N_{\text{studies}} = 6$), mean education ($N_{\text{studies}} = 8$), pre-morbid IQ ($N_{\text{studies}} = 3$), and co-morbid medical conditions ($N_{\text{studies}} = 7$) to reliably report these sample characteristics.

*(Table 1 – Summary demographic and surgical data for TJA and control groups)*

*Early post-operative outcomes.*

In total, 13 studies ($N_{\text{TJA}} = 807$) examined cognitive functioning prior to being discharged from hospital (mean interval = 5.1 days, *SD* = 1.9) using a single-sample (TJA) pre- and post-surgery design. As seen in Table 2, the cognitive tests that were most frequently used were tests of immediate verbal recall ($N_{\text{studies}} = 7$), Trail Making Tasks ($N_{\text{studies}} = 6$), Controlled Oral Word Association Test ($N_{\text{studies}} = 6$), and the Mini Mental Status Examination ($N_{\text{studies}} = 6$).

*(Table 2 – Pre-discharge cognitive outcomes after TJA surgery)*

Prior to discharge, TJA patients showed a small but significant negative effect for Choice Reaction Time performance, which suggests that they were slower to respond (refer to Table 2). However, this result is derived from only one study and should therefore be interpreted with caution. A moderate and significant negative effect was also evident for the Mini Mental Status Examination, indicating that patients typically performed more poorly on a commonly used cognitive screen shortly after surgery. In contrast, a moderate and significant improvement was found for delayed visual recall; however, this result was also only based on one study and may therefore be less reliable.

Interestingly, the majority of tests that were used at the pre-discharge assessment showed no discernable change in performance from the baseline testing that was performed prior to surgery. Specifically, within the memory domain, there was no change in immediate verbal recall, recognition memory, or delayed visual recall. In addition, no changes were observed for any of the attention or language tasks. Lastly, no significant change was noted for the tests of motor speed or digit substitution.

In contrast, there were only two studies ($N_{TJA}$ = 220, $N_{\text{controls}}$ = 89) that assessed cognitive function prior to hospital discharge (mean interval = 5.5 days, *SD* = 2.1) using a two-sample (TJA and Controls) pre- and post-surgery design, with a verbal recall task being the one test that was used by both studies (refer to Table 2). A small, negative effect was found for this test, which suggests that the verbal recall of TJA patients improved less than the control participants between baseline (pre-surgery) and follow-up. However, the $N_{\text{fs}}$ is low, suggesting that this finding may overestimate the true effect if publication bias has occurred. Although not evident from the data provided in Table 2, the raw data for these studies indicate that the TJA patients performed more poorly at both baseline and follow-up, compared to controls. Thus, the TJA patients had compromised verbal recall performance *prior* to surgery and this disparity increased in the early period after their surgery.

*Longer-term post-operative outcomes.*

A total of 12 studies ($N_{\text{TJA}}$ = 970, $N_{\text{controls}}$ = 89) assessed cognitive function at longer follow-up (mean follow-up interval = 122.5 days, SD = 48.9) using a single-sample (TJA) pre- and post-surgery design. As seen in Table 3, the cognitive tests that were most commonly used by these studies were tests of immediate verbal recall ($N_{\text{studies}}$ = 8), Trail Making Tasks ($N_{\text{studies}}$ = 8), the Controlled Oral Word Association Test ($N_{\text{studies}}$ = 5), and digit substitution tasks ($N_{\text{studies}}$ = 5).

*(Table 3 – 3-6 month cognitive outcomes after TJA surgery)*

The TJA sample did not show evidence of deteriorating cognitive performance on any of the tests that were used; however, significant improvements were noted for three tests (refer to Table 3). Specifically, low-to-moderate positive effect sizes were found for the tests of immediate verbal and delayed visual recall, which suggests that the memory performance of the TJA patients improved when they were followed-up 3 to 6 months after their surgery. A small positive effect was also found for substitution task performance, indicating that processing speed improved.

As was seen for the pre-discharge results, there was no substantial change in performance on the majority of tests between baseline and the longer follow-up, as indicated by negligible effect sizes. More specifically, performance on specific tests of memory (delayed verbal recall, recognition memory, immediate visual recall), attention (Digit Span, Wechsler Memory Scale Attention and Concentration Index), language (Controlled Oral Word Association Test, Category Fluency, Boston Naming Test), speed (Choice Reaction Time, motor speed) and general cognition (Mini Mental Status Examination) all remained relatively constant.

Once again, two studies ($N_{TJA}$ = 220, $N_{\text{controls}}$ = 89) assessed cognition at longer follow-up (3 months) using a two-sample pre- and post-surgery design, and tests of immediate verbal recall and Trail Making. No substantial effects were found for either of these measures, which suggest that any changes in performance between baseline and follow-up were comparable for the TJA patients and healthy controls (Table 3). However, consistent with the pre-discharge data, TJA patients performed significantly more poorly than controls on all tests at the baseline and longer follow-up interval, suggesting that the TJA patients were more cognitively compromised prior to surgery and remained so for three months post-surgery.

## Discussion

The current meta-analysis investigated cognitive outcomes after TJA using data from 17 studies ($N_{patients}$ = 1,089), only two of which used a control group ($N_{controls}$ = 89). When followed-up prior to hospital discharge, the data for the TJA sample revealed small deficits in reaction time and general cognitive performance, compared to their pre-surgery performance. However, the majority of tests showed no change in performance, suggesting that TJA had minimal impact on cognitive performance within one week of surgery. Although the data were limited, when both TJA and control groups were assessed at baseline and pre-discharge, it was found that the immediate verbal recall of the TJA group was significantly poorer than that of the controls; a difference that was present prior to surgery but increased at the time of the pre-discharge follow-up. Pain and the use of opioid analgesics may partly explain the poorer pre-surgery performance of the TJA group [16, 45]. In addition, a small number of ‘elective’ surgeries may have involved revision procedures that were required due to infection, which may also impair cognition [8]. That fact that the TJA patients showed evidence of poorer cognitive performance soon after surgery was not surprising, as they were likely to be experiencing high levels of pain and many more would be taking opioid analgesics, both of which impact on cognition [16, 45].

Data collected 3 to 6 months after TJA surgery may therefore be more informative with regard to POCD, as patients were likely to be experiencing less pain and unlikely to be taking high doses of opioid analgesics. However, based on the available data, there was no evidence of cognitive decline following TJA; rather, there were small-to-moderate improvements in immediate verbal recall, delayed visual recall and processing speed. Consistent with pre-discharge assessments, there were no changes in performance on the majority of tasks. Finally, the two studies that assessed both TJA and control groups found that there were no differences in the changes to the immediate verbal recall and Trail Making performance of these groups over time. Thus, although the TJA patients performed more poorly at baseline, they remained equally compromised relative to their healthy peers three months after surgery. The poorer baseline performance of the TJA patients is likely to be the result of joint-related pain and the use of analgesia prior to undergoing surgery. However this does not explain why the differences continued when pain levels and the use of analgesics are likely to have decreased.

The large number of small and non-significant results in this study may, in part, be explained by the fact that a single conservative estimate of test-retest reliability ($r$ = .7) was used to weight all effect sizes. This value is likely to have underestimated the reliability of some tests [39] and, consequently, increased the size of the CIs and the likelihood that they would span zero, leading to the conclusion that there was no effect [46]. Non-significant findings may also have resulted from the use of tests that are insensitive to subtle levels of cognitive decline, such as the MMSE [47].

Alternatively, it is possible that TJA does not affect cognition. However, the failure to observe improvements in test scores due to repeated testing may itself indicate that there has been some cognitive deterioration that has been masked by practice effects [39, 41]. While Control groups normally provide the means by which practice effects are measured and statistically controlled, there were only two studies that used control groups; neither of which showed evidence of improved performance with repeated testing. Unfortunately, these studies only used a small number of comparable measures and so it is not known whether practice effects counteracted any decline on other measures.

It is also possible that cognitive decline occurred in a subset of individuals but was masked when group data was analysed [41, 48, 49]. This is illustrated in a recent meta-analysis of cognitive outcomes after cardiac surgery [24] which, based on group data, concluded that there was either no change or small improvements in performance following surgery. However, the prevalence rates of those studies that provided group data, indicate that up to one third of cardiac patients showed cognitive decline; suggesting that there may be a subset of people who experience poor outcomes. This also appears to be the case for the current meta-analysis, with the group data indicating that there are either no persisting cognitive changes or small improvements after TJA surgery, and the prevalence rates for POCD varying between 16% and 45% [5, 12-14, 37]. For this reason, the *Statement of Consensus on Assessment of Neurobehavioral Outcomes After Cardiac Surgery* [48] advocates monitoring the outcomes of individuals, as well as groups, when evaluating surgical interventions.

Finally, it is important to consider some of the limitations of this study. First, the data from different tests were combined for present purposes, which may have meant that tests that were subtly different (i.e. in terms of their sensitivity to cognitive deficits) were combined. Second, some studies assessed participants twice and others three times (pre-surgery + pre-discharge and/or 3-6 months), which may have affected the extent to which practice effects would be expected. Third, patients who were assessed between three and six months post-surgery were combined. Lastly, it was necessary to exclude three studies [4, 28, 30] because they did not report the requisite data and attempts to obtain it were unsuccessful. This is particularly regrettable in the case of the Ancelin et al. [4] study because it used a sizeable cognitive test battery and a matched control group. The remaining two studies [28, 30] did not use a control group and only assessed cognition using the MMSE, and were therefore unlikely to add significantly to the findings.

Importantly, this meta-analysis highlights the need for additional good quality research, which uses an appropriate control group and clear criteria for identifying cases of POCD, in order to evaluate the frequency with which POCD occurs following TJA [48]. Demographically-matched healthy controls are likely to be the most viable option because they are readily available and control both for normal changes in performance over time and practice effects [41]. However, a surgical control group is needed in order to determine whether TJA, as opposed to surgery itself, is responsible for any observed cognitive decline, provided that the surgical control and TJA groups are matched on important variables (e.g., demographic: age, education; surgical: type of and time under anaesthesia; and disability: pain, physical disability). Another option would be to use a control group consisting either of persons with osteoarthritis or who patients who are wait-listed for TJA surgery, provided they are comparable in terms of pain and disability [41].

Future research should also report both group and individual data. POCD must additionally be identified using sound theoretical and statistical methods, such as standardised regression-based methodology [50], which takes time-related confounds into consideration (e.g. practice effects, test-retest reliability, regression to the mean); therefore measuring ‘true’ change in individual patients. Lastly, it is important that researchers provide detailed information regarding other variables that may independently contribute to poor cognitive performance or other outcomes, such as the surgical method used (implant type, tourniquet vs no tourniquet), anaesthetic (type, amount), medication (type, dosage, duration), level of pain, and depression.

In summary, prior to discharge, TJA patients showed no change in their performance on the majority of cognitive tasks from their pre-surgical assessment but there were small deficits in reaction time and general cognition. However, these assessments are likely to have been affected by post-operative pain and opioid analgesia. At the 3 to 6 month interval, the TJA group showed no change from baseline on the majority of tasks, although there were small improvements in immediate verbal recall, delayed visual recall, and speed of processing. Unfortunately, in the absence of adequate control data, it is not possible to more definitively determine whether TJA is associated with cognitive decline, as it is not known whether practice effects masked any such decline. Moreover, in the two studies where control groups were used, it appeared that TJA patients were cognitively compromised even before surgery, which further complicates the picture. Additional methodologically rigorous research into POCD after TJA is needed if we are to gain a better understanding of the incidence, cause and impact of POCD.

## Supplementary Material

*Search terms used*

| Total Joint Replacement | Cognition |
|---|---|
| “joint replacement” | “cogniti*” |
| “hip replacement” | “neurocog*” |
| “knee replacement” | “neuropsycholo*” |
| “joint surgery” | “neurobehavioural” |
| “hip surgery” | “mental performance” |
| “knee surgery” | “mental capacity” |
| “joint procedure” | “mental function” |
| “hip procedure” | “mental competenc*” |
| “knee procedure” | “mental ability*” |
| “joint prosthesis” | “psychologic*” |
| “hip prosthesis” | |
| “knee prosthesis” | |
| “arthroplasty” | |
| “TJR” | |
| “THR” | |
| “TKR” | |
| “THA” | |
| “TKA” | |

## References

1. Harris WH, Sledge CB. Total hip and total knee replacement. N Engl J Med 1990; 323(11): 725.
2. Australian Orthopaedic Association National Joint Replacement Registery. Hip and knee arthroplasty: Annual report. In. 2012
3. Labek G, Thaler M, Janda W, Agreiter M, Stockl B. Revision rates after total joint replacement: cumulative results from worldwide joint register datasets. J Bone Joint Surg Br 2011; 93(3): 293.
4. Ancelin ML, de Roquefeuil G, Scali J, Bonnel F, Adam JF, Cheminal JC, Cristol JP, Dupuy AM, Carriere I, Ritchie K. Long-term post-operative cognitive decline in the elderly: the effects of anesthesia type, apolipoprotein E genotype, and clinical antecedents. J Alzheimers Dis 2010; 22 Suppl 3: 105.
*5. Evered L, Scott DA, Silbert B, Maruff P. Postoperative cognitive dysfunction is independent of type of surgery and anesthetic. Anesth Analg 2011; 112(5): 1179.
*6. Williams-Russo P, Sharrock NE, Mattis S, Liguori GA, Mancuso C, Peterson MG, Hollenberg J, Ranawat C, Salvati E, Sculco T. Randomized trial of hypotensive epidural anesthesia in older adults. Anesthesiology 1999; 91(4): 926.
7. Deiner S, Silverstein JH. Postoperative delirium and cognitive dysfunction. Br J Anaesth 2009; 103 Suppl 1: i41.
8. Koessler MJ, Pitto RP. Fat embolism and cerebral function in total hip arthroplasty. Int Orthop 2002; 26(5): 259.
9. Fong HK, Sands LP, Leung JM. The role of postoperative analgesia in delirium and cognitive decline in elderly patients: a systematic review. Anesth Analg 2006; 102(4): 1255.
*10. Nielson WR, Gelb AW, Casey JE, Penny FJ, Merchant RN, Manninen PH. Long-term cognitive and social sequelae of general versus regional anesthesia during arthroplasty in the elderly. Anesthesiology 1990; 73(6): 1103.
*11. Jones MJ, Piggott SE, Vaughan RS, Bayer AJ, Newcombe RG, Twining TC, Pathy J, Rosen M. Cognitive and functional competence after anaesthesia in patients aged over 60: controlled trial of general and regional anaesthesia for elective hip or knee replacement. BMJ 1990; 300(6741): 1683.
*12. Deo H, West G, Butcher C, Lewis P. The prevalence of cognitive dysfunction after conventional and computer-assisted total knee replacement. Knee 2011; 18(2): 117.
*13. Koch S, Forteza A, Lavernia C, Romano JG, Campo-Bustillo I, Campo N, Gold S. Cerebral fat microembolism and cognitive decline after hip and knee replacement. Stroke 2007; 38(3): 1079.
*14. Salazar F, Donate M, Boget T, Bogdanovich A, Basora M, Torres F, Fabregas N. Intraoperative warming and post-operative cognitive dysfunction after total knee replacement. Acta Anaesthesiol Scand 2011; 55(2): 216.
*15. Patel RV, Stygall J, Harrington J, Newman SP, Haddad FS. Cerebral microembolization during primary total hip arthroplasty and neuropsychologic outcome: A pilot study. Clin Orthop Relat Res 2010; 468(6): 1621.
*16. Duggleby W, Lander J. Cognitive status and postoperative pain: Older adults. J Pain Symptom Manage 1994; 9(1): 19.
*17. McCaffrey R. The effect of music on acute confusion in older adults after hip or knee surgery. Appl Nurs Res 2009; 22(2): 107.
*18. Postler A, Neidel J, Günther KP, Kirschner S. Incidence of early postoperative cognitive dysfunction and other adverse events in elderly patients undergoing elective total hip replacement (THR). Arch Gerontol Geriatr 2011; 53(3): 328.
19. Phillips-Bute B, Mathew JP, Blumenthal JA, Grocott HP, Laskowitz DT, Jones RH, Mark DB, Newman MF. Association of neurocognitive function and quality of life 1 year after coronary artery bypass graft (CABG) surgery. Psychosom Med 2006; 68(3): 369.
20. Newman MF, Grocott HP, Matthew JP, White WD, Landolfo K, Reves JG, Laskowitz DT, Mark DB, Blumenthal JA. Report of the substudy assessing the impact of neurocognitive function on quality of life 5 years after cardiac surgery. Stroke 2001; 32: 2874.
21. Steinmetz J, Christensen KB, Lund T, Lohse N, Rasmussen LS. Long-term consequences of postoperative cognitive dysfunction. Anesthesiology 2009; 110(3): 548.

22. Moller JT, Cluitmans P, LS R. Long-term postoperative cognitive dysfunction in the elderly ISPOCD1 study. ISPOCD investigators. International Study of Post-Operative Cognitive Dysfunction. Lancet 1998; 351: 857.
23. Dunkel A, Kendel F, Lehmkuhl E, Babitsch B, Oertelt-Prigione S, Hetzer R, Regitz-Zagrosek V. Predictors of preoperative depressive risk in patients undergoing coronary artery bypass graft surgery. Clin Res Cardiol 2009; 98(10): 643.
24. Cormack F, Shipolini A, Awad WI, Richardson C, McCormack DJ, Colleoni L, Underwood M, Baldeweg T, Hogan AM. A meta-analysis of cognitive outcomes following coronary artery bypass graft surgery. Neurosci Biobehev Rev 2012; 36(9): 2118.
25. Guay J. General anaesthesia does not contribute to long-term post-operative cognitive dysfunction in adults: A meta-analysis. Indian J Anaesth 2011; 55(4): 358.
26. Newman S, Stygall J, Hirani S, Shaefi S, Maze M. Postoperative cognitive dysfunction after noncardiac surgery: a systematic review. Anesthesiology 2007; 106(3): 572.
*27. Lowery DP, Wesnes K, Brewster N, Ballard C. Quantifying the association between computerised measures of attention and confusion assessment method defined delirium: a prospective study of older orthopaedic surgical patients, free of dementia. Int J Geriatr Psychiatry 2008; 23(12): 1253.
28. Kudoh A, Takase H, Takazawa T. A comparison of anesthetic quality in propofol-spinal anesthesia and propofol-fentanyl anesthesia for total knee arthroplasty in elderly patients. J Clin Anesth 2004; 16(6): 405.
*29. Gray AC, Torrens L, Howie CR, Christie J, Robinson CM. Cognitive function and cerebral emboli after primary hip arthroplasty. Hip Int 2008; 18(1): 40.
30. Bickel H, Gradinger R, Kochs E, Forstl H. High risk of cognitive and functional decline after postoperative delirium: A three-year prospective study. Dement Geriatr Cogn Disord 2008; 26(1): 26.
31. Li YC, Xi CH, An YF, Dong WH, Zhou M. Perioperative inflammatory response and protein S-100β concentrations ? relationship with post-operative cognitive dysfunction in elderly patients. Acta Anaesthesiol Scand 2012; 56(5): 595.
32. Lipsey MW, Wilson DB. Practical Meta-Analysis. 2001 Thousand Oaks: Sage Publications
*33. Evered LA, Silbert BS, Scott DA, Maruff P, Ames D, Choong PF. Preexisting cognitive impairment and mild cognitive impairment in subjects presenting for total hip joint replacement. Anesthesiology 2011; 114(6): 1297.
*34. Duppils GS, Wikblad K. Acute confusional states in patients undergoing hip surgery. a prospective observation study. Gerontology 2000; 46(1): 36.
*35. Duppils GS, Wikblad K. Cognitive function and health-related quality of life after delirium in connection with hip surgery. A six-month follow-up. Orthop Nurse 2004; 23(3): 195.
*36. Cheng CM, Chiu MJ, Wang JH, Liu HC, Shyu YIL, Huang GH, Chen CCH. Cognitive stimulation during hospitalization improves global cognition of older Taiwanese undergoing elective total knee and hip replacement surgery. J Adv Nurs 2011.
*37. Rodriguez RA, Tellier A, Grabowski J, Fazekas A, Turek M, Miller D, Wherrett C, Villeneuve PJ, Giachino A. Cognitive dysfunction after total knee arthroplasty: Effects of intraoperative cerebral embolization and postoperative complications. J Arthroplasty 2005; 20(6): 763.
*38. Stockton P, Cohen-Mansfield J, Billig N. Mental status change in older surgical patients: Cognition, depression, and other comorbidity. Am J Geriatr Psychiatry 2000; 8(1): 40.
39. Lezak MD, Howieson DB, Bigler ED, Tranel D. Neuropsychological Assessment. 2012 New York: Oxford University Press
40. Chen X, Zhao M, White PF, Li S, Tang J, Wender RH, Sloninsky A, Naruse R, Kariger R, Webb T. The recovery of cognitive function after general anesthesia in elderly patients: a comparison of desflurane and sevoflurane. Anesth Analg 2001; 93(6): 1489.
41. Slade P, Sanchez P, Townes B, Aldea GS. The use of neurocognitive tests in evaluating the outcome of cardiac surgery: Some methodological considerations. J Cardiothorac Vasc Anesth 2001; 15(1): 4.
42. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups design. Psychol Methods 2002; 7: 105.

43. Cohen J. Statistical Power Analyses for the Behavioral Sciences. 1988 New Jersey: Lawrence Erlbaum Associates
44. Gregory R. Psychological Testing: History, Principles and Applications. 2011 Boston: Allyn & Bacon
45. Chapman SL, Byas-Smith MG, Reed BA. Effects of intermediate-and long-term use of opioids on cognition in patients with chronic pain. Clin J Pain 2002; 18(4): S83.
46. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. Psychol Methods 1996; 1(2): 170.
47. Friedman TW, Yelland GW, R RS. Subtle cognitive impairment in elders with Mini-Mental State Examination scores within the 'normal' range. Int J Geriatr Psychiatry 2011; 27: 463.
48. Murkin JM, Newman SP, Stump DA, Blumenthal JA. Statement of consensus on assessment of neurobehavioral outcomes after cardiac surgery. Ann Thorac Surg 1995; 59(5): 1289.
49. Newman SP. Analysis and interpretation of neuropsychologic tests in cardiac surgery. Ann Thorac Surg 1995; 59(5): 1351.
50. Sawrie SM, Chelune GJ, Naugle RI, Lüders HO. Empirical methods for assessing meaningful neuropsychological change following epilepsy surgery. J Int Neuropsychol Soc 1996; 2(6): 556.

Table 1

*Summary demographic and surgery data for the TJR and control groups*

| | **TJR group** | | | | | **Control group** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $N_{studies}$ | $N_{participants}$ | % | *M* | *SD* | $N_{studies}$ | $N_{participants}$ | % | *M* | *SD* |
| Sample size* | 17 | 1089 | | 64.1 | 54.6 | 2 | 89 | | 44.5 | 16.3 |
| Age (yrs) | 14 | 873 | | 71.6 | 3.4 | 2 | 89 | | 72.9 | 1.5 |
| Gender | 14 | 979 | | | | 2 | 89 | | | |
| Female | | | 63 | | | | | 57 | | |
| Male | | | 37 | | | | | 43 | | |
| Surgery Type | | | | | | | | | | |
| THR | 8 | 640 | 59 | | | | | | | |
| TKR | 4 | 177 | 16 | | | | | | | |
| THR/TKR | 5 | 272 | 25 | | | | | | | |

*Note.* THR = total hip replacement, TKR = total knee replacement, $N_{studies}$ = number of studies, $N_{participants}$ = number of participants.

*Participants with complete baseline and follow-up data only.

Table 2

*Pre-discharge cognitive outcomes after TJR surgery*

**TJR pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{RM}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | verbal - immediate recall | 7 | 572 | .10 | -.10 | .29 | 0 | [5, 6, 12-14, 29, 37] |
| | verbal - delayed recall | 5 | 236 | -.18 | -.39 | -.03 | 4 | [12-14, 29, 37] |
| | verbal - recognition | 2 | 238 | -.12 | -.38 | .12 | 1 | [6, 13] |
| | visual - immediate recall | 2 | 244 | .01 | -.08 | .10 | 0 | [6, 37] |
| | visual - delayed recall | 1 | 22 | .37* | .04 | .70 | 3 | [13] |
| Attention | Trail Making Task A | 6 | 509 | .03 | -.10 | .14 | 0 | [5, 6, 12, 13, 29, 37] |
| | Trail Making Task B | 6 | 509 | -.05 | -.20 | .09 | 0 | [5, 6, 12, 13, 29, 37] |
| | Digit Span | 4 | 321 | -.07 | -.23 | .08 | 0 | [6, 13, 14, 29] |
| | WMS Attention & Concentration | 1 | 28 | .09 | -.20 | .38 | 1 | [37] |
| Language | Controlled Oral Word Assoc. Test | 6 | 509 | .01 | -.05 | .08 | 0 | [5, 6, 12, 13, 29, 37] |
| | Category fluency | 3 | 251 | .02 | -.11 | .08 | 0 | [5, 12, 37] |
| | Boston Naming Test | 2 | 277 | .08 | .00 | .16 | 1 | [6, 12] |
| Motor and Processing Speed | Choice Reaction Time | 1 | 64 | -.24* | -.40 | -.08 | 1 | [27] |
| | Motor speed | 3 | 245 | -.11 | -.38 | .17 | 1 | [5, 12, 13] |
| | Substitution Task | 3 | 400 | -.26 | -.54 | .04 | 5 | [5, 6, 13] |
| General Cognitive Screen | Mini-Mental Status Examination | 6 | 216 | -.41* | -.61 | -.20 | 19 | [17, 18, 27, 35, 36, 51] |

**TJR and Controls pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{\text{IGRM}}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | verbal - immediate recall | 2 | 309[a] | -.20 | -.29 | -.12 | 2 | [5, 14] |

*Note.* $N_{\text{studies}}$ = number of studies, $N_{\text{participants}}$ = number of participants, $d_{RM}$ = weighted mean single-sample repeated measures Cohen's *d*, 95% CI = 95% confidence interval, $N_{fs}$ = fail-safe *N*, $d_{\text{IGRM}}$ = weighted mean independent-groups repeated measures Cohen's *d*, WMS = Wechsler Adult Memory

*satisfies criteria for change in performance ($d \geq .2$, CI $\neq 0$, $N_{fs} > N_{\text{studies}}$)

[a] includes both TJR and Control participants

Table 3

*3-6 month cognitive outcomes after TJR surgery*

**TJR pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{RM}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | verbal - immediate recall | 8 | 656 | .29* | .16 | .42 | 15 | [5, 6, 11-15, 37] |
| | verbal - delayed recall | 4 | 216 | .11 | -.23 | .45 | 1 | [12-14, 37] |
| | verbal - recognition | 2 | 301 | .17 | -.22 | .57 | 1 | [6, 13] |
| | visual - immediate recall | 3 | 307 | -.03 | -.48 | .43 | 0 | [6, 14, 37] |
| | visual - delayed recall | 2 | 85 | .42* | .25 | .59 | 7 | [13, 14] |
| Attention | Trail Making Task A | 8 | 653 | .20 | .11 | .29 | 8 | [5, 6, 10, 12-15, 37] |
| | Trail Making Task B | 8 | 653 | .18 | .04 | .32 | 6 | [5, 6, 10, 12-15, 37] |
| | Digit Span | 3 | 301 | .07 | -.08 | .21 | 0 | [6, 13, 14] |
| | WMS Attention & Concentration | 2 | 89 | .09 | -.06 | .23 | 0 | [10, 37] |
| Language | Controlled Oral Word Assoc. Test | 5 | 484 | .09 | -.01 | .19 | 0 | [5, 6, 12, 13, 37] |
| | Category fluency | 3 | 246 | .02 | -.08 | .12 | 0 | [5, 12, 37 ] |
| | Boston Naming Test | 2 | 277 | .16 | .11 | .21 | 1 | [6, 12] |
| Motor & Processing Speed | Choice Reaction Time | 2 | 174 | .15 | .08 | .21 | 1 | [11, 15] |
| | Motor speed | 4 | 283 | .08 | -.06 | .22 | 1 | [5, 12, 13, 15] |
| | Substitution Task | 5 | 569 | .28* | .16 | .40 | 9 | [5, 6, 11, 13, 15] |
| General Cognitive Screen | Mini Mental Status Examination | 3 | 123 | .00 | -.38 | .38 | 0 | [18, 36, 38] |

**TJR and Control pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{\text{IGRM}}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | Immediate verbal recall | 2 | 309[a] | .07 | -.08 | .22 | 0 | [5, 14] |
| Attention | Trail Making Test A | 2 | 309[a] | -.05 | -.42 | .32 | 0 | [5, 14] |
| | Trail Making Test B | 2 | 309[a] | .01 | -.10 | .12 | 0 | [5, 14] |

*Note.* $N_{\text{studies}}$ = number of studies, $N_{\text{participants}}$ = number of participants, $d_{RM}$ = weighted mean single-sample repeated measures Cohen's *d*, 95% CI = 95% confidence interval, $N_{fs}$ = fail-safe *N*, $d_{\text{IGRM}}$ = weighted mean independent-groups repeated measures Cohen's *d*, WMS = Wechsler Adult Memory

*satisfies criteria for change in performance ($d \geq .2$, CI ≠ 0, $N_{fs} > N_{\text{studies}}$)

[a] includes both TJR and Control participants

Table 3

*3-6 month cognitive outcomes after TJR surgery*

**TJR pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{RM}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | verbal - immediate recall | 8 | 656 | .29* | .16 | .42 | 15 | [1-8] |
| | verbal - delayed recall | 4 | 216 | .11 | -.23 | .45 | 1 | [1, 4, 6, 7] |
| | verbal - recognition memory | 2 | 301 | .17 | -.22 | .57 | 1 | [4, 8] |
| | visual - immediate recall | 3 | 307 | -.03 | -.48 | .43 | 0 | [6-8] |
| | visual - delayed recall | 2 | 85 | .42* | .25 | .59 | 7 | [4, 7] |
| Attention | Trail Making Task A | 8 | 653 | .20 | .11 | .29 | 8 | [1, 2, 4-9] |
| | Trail Making Task B | 8 | 653 | .18 | .04 | .32 | 6 | [1, 2, 4-9] |
| | Digit Span | 3 | 301 | .07 | -.08 | .21 | 0 | [4, 7, 8] |
| | WMS Attention & Concentration | 2 | 89 | .09 | -.06 | .23 | 0 | [6, 9] |
| Language | Controlled Oral Word Assoc. Test | 5 | 484 | .09 | -.01 | .19 | 0 | [1, 2, 4, 6, 8] |
| | Category fluency | 3 | 246 | .02 | -.08 | .12 | 0 | [1, 2, 6 ] |
| | Boston Naming Test | 2 | 277 | .16 | .11 | .21 | 1 | [1, 8] |
| Motor & Processing Speed | Choice Reaction Time | 2 | 174 | .15 | .08 | .21 | 1 | [3, 5] |
| | Motor speed | 4 | 283 | .08 | -.06 | .22 | 1 | [1, 2, 4, 5] |
| | Substitution Task | 5 | 569 | .28* | .16 | .40 | 9 | [2-5, 8] |
| General Cognitive Screen | Mini Mental Status Examination | 3 | 123 | .00 | -.38 | .38 | 0 | [10-12] |

**TJR and Control pre- and post-surgery data**

| *Domain* | *Test* | $N_{\text{studies}}$ | $N_{\text{participants}}$ | $d_{\text{IGRM}}$ | *95% CI* | | $N_{fs}$ | *References* |
|---|---|---|---|---|---|---|---|---|
| Memory | Immediate verbal recall | 2 | $309^a$ | .07 | -.08 | .22 | 0 | [2, 7] |
| Attention | Trail Making Test A | 2 | $309^a$ | -.05 | -.42 | .32 | 0 | [2, 7] |
| | Trail Making Test B | 2 | $309^a$ | .01 | -.10 | .12 | 0 | [2, 7] |

*Note.* $N_{\text{studies}}$ = number of studies, $N_{\text{participants}}$ = number of participants, $d_{RM}$ = weighted mean single-sample repeated measures Cohen's *d*, 95% CI = 95% confidence interval, $N_{fs}$ = fail-safe *N*, $d_{\text{IGRM}}$ = weighted mean independent-groups repeated measures Cohen's *d*, WMS = Wechsler Adult Memory
*satisfies criteria for change in performance ($d \geq .2$, CI $\neq 0$, $N_{fs} > N_{\text{studies}}$)
[a]includes both TJR and Control participants