Pergamon

Archives of Clinical Neuropsychology
19 (2004) 437–454

Archives of CLINICAL NEUROPSYCHOLOGY

Persistence of cognitive effects after withdrawal from long-term benzodiazepine use: a meta-analysis

Melinda J. Barker, Kenneth M. Greenwood, Martin Jackson, Simon F. Crowe*

School of Psychological Science, La Trobe University, Vic. 3086, Australia

Accepted 19 May 2003

Abstract

Despite the widespread prescribing of benzodiazepines, uncertainty still surrounds the potential for cognitive impairment following their long-term use. Furthermore, the degree of recovery that may take place after withdrawal or the level of residual impairment, if any, that is maintained in long-term benzodiazepine users is also unclear. The current paper employed meta-analytic techniques to address two questions: (1) Does the cognitive function of long-term benzodiazepine users improve following withdrawal? (2) Are previous long-term benzodiazepine users still impaired at follow-up compared to controls or normative data? Results of the meta-analyses indicated that long-term benzodiazepine users do show recovery of function in many areas after withdrawal. However, there remains a significant impairment in most areas of cognition in comparison to controls or normative data. The findings of this study highlight the problems associated with long-term benzodiazepine therapy and suggest that previous benzodiazepine users would be likely to experience the benefit of improved cognitive functioning after withdrawal. However, the reviewed data did not support full restitution of function, at least in the first 6 months following cessation and suggest that there may be some permanent deficits or deficits that take longer than 6 months to completely recover.

Keywords: Benzodiazepines; Tranquilizing drugs; Meta-analysis; Neuropsychological assessment; Cognitive ability

* Corresponding author. Tel.: +61-3-9479-1380; fax: +61-3-9479-1956.
E-mail address: s.crowe@latrobe.edu.au (S.F. Crowe).

0887-6177/$ – see front matter © 2003 National Academy of Neuropsychology.
doi:10.1016/S0887-6177(03)00096-9

Benzodiazepines are the most frequently used class of drugs in the treatment of anxiety disorders. An estimated past year prevalence of use in the USA has been reported at 12.9%, with 14.2% of this group taking the drug for a period of 12 months or more (Balter, Manheimer, Mellinger, & Uhlenhuth, 1984). Despite such widespread use, little is known regarding the potential long-term effects of benzodiazepines. It is well accepted that, even at therapeutic doses, benzodiazepines are capable of causing physiological and pharmacological dependence leading to a withdrawal syndrome after cessation of use (Ashton, 1986; Chen, 1990; Lader, 1982). However, research into the residual cognitive effects after long-term use yields conflicting results.

Previous research supports ongoing impairments after withdrawal from long-term benzodiazepine use in the areas of memory (Birzele, 1992; Curran, 1992; Curran et al., 1994; Tata, Rollings, Collins, Pickering, & Jacobson, 1994), attention and concentration (Birzele, 1992; Golombok, 1989; Golombok, Moodley, & Lader, 1988; Petursson, Gudjonsson, & Lader, 1983), visuospatial skills (Bergman, Borg, & Holm, 1980; Golombok et al., 1988; Sakol & Power, 1988; Tata et al., 1994) and numerous other cognitive domains (Aranko, Mattila, & Seppala, 1983; Bergman, Borg, Engelbrektson, & Vikander, 1989; Bergman et al., 1980; Birzele, 1992; Brosan, Broadbent, Nutt, & Broadbent, 1986; Gorenstein, Bernik, & Pompeia, 1994; Lucki, Rickels, & Geller, 1986; Petursson et al., 1983; Sakol & Power, 1988; Tata et al., 1994; Toenne et al., 1995). In contrast, some researchers claim little or no effects on cognition after withdrawal from long-term use (Chaves, Bianchin, Peccin, & Rotta, 1993; Lucki & Rickels, 1986; Lucki, Rickels, & Geller, 1985; Petursson et al., 1983).

The growing body of research investigating cognitive improvement after discontinuation of chronic use of benzodiazepines generates equally conflicting results. Sakol and Power (1988) studied 12 long-term benzodiazepine users (mean use 9 years) before and after graded withdrawal and observed improvements in tasks assessing attention, vigilance, and speed of information processing. Although only 7 of the original 12 subjects were assessed at the 4-week follow-up stage, the authors suggest that these results support a rapid recovery of function. Considering that the 4-week withdrawal program was relatively short and not all patients were drug free at this point, these conclusions seem premature and the possibility of practice effects due to weekly assessments should be considered.

Rickels, Lucki, Schweizer, Garcia-Espana, and Case (1999) compared two groups of long-term benzodiazepine users (mean duration of use 8 years), those who were successful in their withdrawal program and those who were not. Successful taper patients displayed improved performance on a digit-symbol-substitution test and a symbol copy task, and furthermore, were observed to be more alert, more relaxed and less anxious.

Salzman, Fisher, Nobel, and Glassman (1992) reported that cognitive impairment observed in their samples of 13 nursing home residents was reversed after withdrawal as compared to the 12 residents who did not withdraw. In addition, the authors reported that significant improvements observed in short-term memory (as measured by digit span and the vigilance test paradigm) were readily apparent to staff and family members, with patients who had withdrawn appearing brighter, more energetic, less dysphoric, and more intellectually alert. Other areas which have been reported to improve following discontinuation of long-term benzodiazepine use include non-verbal visual memory (Birzele, 1992), explicit memory (Kilic, 1999) and general neuropsychological function (measured by the synonyms, reasoning and block design

(SRB) battery and five subtests from the Halstead–Reitan battery) (Bergman et al., 1989; Toenne et al., 1995).

Conversely, some studies support a lack of cognitive improvement after withdrawal from long-term benzodiazepine use and suggest that a variety of impairments persist some time after discontinuation.

Gorenstein, Bernik, Pompeia, and Marcourakis (1995) followed up 18 previous low-dose benzodiazepine users of more than 5 years. Performance did not significantly improve after an average of 10 months abstinence on tests such as digit-symbol-substitution, digit span and immediate and delayed recall. These authors note the significance of such a finding in users of low therapeutic doses of benzodiazepines as it contradicts suggestions that impairments are unlikely in those with a low intake of these drugs. Tata et al.'s (1994) group of chronic benzodiazepine users followed up 6 months after successful discontinuation displayed significant impairments in verbal learning and memory, psychomotor, visuomotor, and visuoconceptual abilities compared to controls. Although some degree of recovery was observed, their patients performed significantly worse than controls, which the authors suggest indicates persistent cognitive deficits related to long-term benzodiazepine use.

A number of questions therefore remain. First, does long-term benzodiazepine use lead to cognitive impairment? Second, what degree of recovery takes place after withdrawal, and third, what is the level of permanent impairment in these patients.

With regard to the rate of recovery, some support for rapid recovery after only 4 weeks withdrawal has been found (Sakol & Power, 1988), while others have reported impairments in patients up to 1 year (Toenne et al., 1995) or 4–6 years later (Bergman et al., 1989). These findings suggest that in addition to pre-morbid differences, considerable individual variation may influence the rate and level of recovery.

The current paper addressed two questions: (1) Does the cognitive function of long-term benzodiazepine users improve following withdrawal? (2) Are previous long-term benzodiazepine users still impaired at follow-up compared to controls or normative data? In order to examine the questions of recovery of function and persistence of effects, two meta-analyses were undertaken.

1. Method

1.1. Selection of studies

A comprehensive search of the computerized databases MEDLINE, PsychINFO, and the Cochrane Controlled Trials Register was conducted to identify articles addressing long-term use of benzodiazepines published between 1980 and 2000. Key search terms used included "benzodiazepine," "benzodiazepines," "hypnotics," and "sedatives" paired with "long-term," "chronic," "effects," "cognitive," and "deficits." Only articles written in English and published in peer-reviewed journals were included. Relevant articles were obtained and the bibliographies examined for additional relevant articles not identified through the computer-based searches. These articles were then obtained and their reference lists scanned and so on.

1.2. Criteria for inclusion

For an article to be included in the meta-analysis, it was necessary for it to meet the following criteria:

1. published between 1980 and 2000;
2. published in peer-reviewed journal;
3. English language;
4. control group or within-subjects design used;
5. cognitive assessment conducted;
6. minimum period of benzodiazepine use of at least 1 year;
7. results reported sufficient to allow calculation of effect sizes; and,
8. original results or results not reported elsewhere.

Of the 34 papers identified from the searches, 19 were not eligible for inclusion for the following reasons:

- one did not use any objective test measures;
- five used computed axial brain tomography and no psychometric tests;
- nine did not reach the minimum of 1 year of benzodiazepine use;
- two were reporting preliminary results of later studies included in the meta-analysis;
- one reported a summary of previous results included in the meta-analysis; and,
- one reported results as correlations between effects and dosage, and it was not possible to transform these data in such a way as to allow for calculation of effect sizes.

On two occasions, pairs of articles were combined as they reported on the same participant group using different tests. This resulted in a sample of 13 independent studies. For inclusion in the current meta-analysis, studies needed to have conducted long-term follow-up assessments. These occurred in 10 of the 13 independent studies. Each test and its corresponding area of cognitive function measured was grouped into 1 of 12 categories corresponding to the broad cognitive area measured as determined by two neuropsychological compendia (Lezak, 1995; Spreen & Strauss, 1998). The 12 categories are listed in Table 1. A full list of the 41 assessment tools used and the cognitive areas measured is included in Appendix A.

Table 1
Cognitive categories to which tests were assigned

Sensory processing	Psychomotor speed
Non-verbal memory	Visuospatial
Speed of processing	Problem-solving
Attention/concentration	Verbal memory
General intelligence	Motor control/performance
Working memory	Verbal reasoning

1.3. Coding of study characteristics

Each of the studies that met inclusion criteria was coded according to certain study attributes. The following variables were extracted and recorded:

Study attributes

1. publication year;
2. journal;
3. country in which the study was carried out;

Subject attributes

4. number of long-term benzodiazepine users in each group;
5. number of controls in each group;
6. type of control group used (i.e., anxious or normal);
7. number of males and females in each group;
8. age, mean, S.D., and range;
9. mean level of education;
10. source of participants (i.e., General Practitioner or primary physician population, hospital patients);
11. length of benzodiazepine use (mean, S.D., and range);
12. type of benzodiazepine used;
13. dosage of benzodiazepine;
14. matching of participants;
15. presence of alcohol or other drug use;
16. condition benzodiazepines prescribed for (i.e., anxiety, depression, insomnia);
17. post-withdrawal psychiatric symptoms
18. definition of dependence
19. time since last dose (if recorded);
20. length of follow-up;

Test information

21. test used in each study/cognitive areas measured;
22. category of cognitive area tested (see Table 1);

Outcome measures

23. exact statistics, means, and standard deviation;
24. results of statistical analysis (i.e., *t*, *P*, and *F* values); and,
25. significance levels.

With regard to the coding of significance levels, a conservative approach was adopted. If the study stated there was "no significant difference" or there was "no difference between the groups," the effect size for that test was set at zero. Similarly, if a study stated that the significance level was, for example, "less than .05" or "less than .01," the *P* value used to calculate the effect size was set only marginally lower at .049 or .0099, respectively.

Separate effect sizes were calculated for each of the 12 cognitive categories. Each study was only allowed to contribute one effect size to each cognitive category by averaging together the effect sizes from the same study if more than one type of test was used to measure a

particular category. This strategy resulted in equal weight being given to each study per category regardless of the number of tests in that category.

The average daily dose of benzodiazepine was converted into a diazepam equivalent dose using tables from Cooper (1982).

1.4. Calculation of effect sizes

Effect sizes were calculated following the method set out by Rosenthal (1991) using Cohen's *d* as the effect size index. Effect sizes addressing the first question regarding improvement (meta-analysis A) represent the difference between the patient group at initial assessment and at follow-up assessment, divided by the pooled standard deviation ($S.D._p$). Therefore, a positive effect size indicates improvement of function. The effect sizes addressing the second question regarding persistence (meta-analysis B) represent the difference between the patients at follow-up assessment and controls at follow-up assessment or normative data. Therefore, a negative effect size represented poorer performance of long-term benzodiazepine users compared to controls or normative data.

Where means and standard deviations were not available, *P* values were converted to *Z* scores that were then used to calculate *r*. To maintain consistency, *r* was converted to *d* using the procedures described by Rosenthal (1991). Effect sizes were weighted on the basis of their sample size.

The relationships between the study characteristics coded and effect size were examined using *t* tests where the variable was categorical, and Pearson correlations where the variable was continuous. The variables examined in the moderator analyses were publication year, number of long-term benzodiazepine users, type of control group (anxious or normal), sex, age, education, length of use, type of benzodiazepine used, dosage, matching method, time since last dose, and length of follow-up period.

2. Results

Publication details of the 10 independent studies selected for the meta-analysis are listed in Table 2. The articles used in meta-analysis A are marked in the reference list with an asterisk.

One study was excluded from meta-analysis B as the follow-up results were not suitably presented to enable calculation of effect sizes. More specifically, it was not possible to obtain follow-up data on the withdrawn patients in order to compare to normative data as only mean change scores were reported without significance levels. In addition, an unsuitable comparison group was used which consisted of non-withdrawn benzodiazepine users. The articles used in meta-analysis B are marked in the reference list with a symbol (†). Results from each meta-analysis are therefore presented separately.

2.1. Meta-analysis A: Does the cognitive function of long-term benzodiazepine users improve following withdrawal?

The median number of benzodiazepine users in the studies was 20.5. The mean and standard deviation was 29.7 (26.6) (range = 10–96). Overall, 37.1% of the 297 subjects were male.

Table 2
Publication details of the studies used in the meta-analysis

Publication year	Country of origin	Publication source
1980 and 1989[a]	Sweden	American Journal of Psychiatry and British Journal of Addiction
1992	Germany	European Review of Applied Psychology
1992	England	Journal of Psychopharmacology
1999	USA	Journal of Clinical Psychopharmacology
1994	England	Psychological Medicine
1995	Sweden	Acta Psychiatrica Scandinavica
1988	Scotland	Psychopharmacology
1994 and 1995[a]	Brazil	International Clinical Psychopharmacology and Journal of Psychopharmacology
1983	England	Psychopharmacology
1992[b]	USA	International Journal of Psychiatry

[a] Denotes where pairs of studies reporting on the same patient group were combined.
[b] Denotes article excluded from meta-analysis B.

The mean age of participants was 47.1 years with a range of 21–75. Nine studies (90%) recruited subjects from those admitted to a hospital or clinic for the purposes of withdrawal or investigation of drug dependence. One study recruited nursing home residents. With regard to the definition of dependence, only two studies specified whether their subjects met any formal criteria, which were DSM-III-R criteria in one instance and WHO (1964) in another. Four studies simply described patients as "long-term benzodiazepine users" or "benzodiazepine dependant." The remaining four studies used the terms "low therapeutic" or "normal dose" users. Eight studies specified the following benzodiazepines, listed in order of decreasing usage frequency, as those used by participants: diazepam, lorazepam, alprazolam, oxazepam, triazolam, bromazepam, chlorodiazepoxide, temazepam, clonazapam, flunitrazepam, clobazam, and nitrazepam. The average daily dose expressed as a diazepam equivalent was 15.3 mg (S.D. = 18.9).

The mean length of benzodiazepine use, specified by all but one study, was 10 years. The range of length of use, specified by seven of the studies, was between 1 and 29 years. Nine studies specified when initial psychometric testing was carried out in relation to time since last dose. In five of the studies testing took place at least one day since the patient's last dose. In the remaining four studies, testing took place either just prior to the normal daily dose, or not within 4 h of taking a normal dose. In those studies that examined the effect of a normal daily dose, the pre-dose data were used in the meta-analysis. The median length of time between initial and post-withdrawal assessment was 3 months (range 1–65). The mean and standard deviation was 10.6 months (19.6).

The majority of studies (80%) excluded patients with a history of heavy alcohol or other drug use, one study did not specify and one study stated that 23% of patients had a history of alcohol use in excess of four standard drinks per day. Six studies specified the condition for which subjects used benzodiazepines. The patients in five studies used benzodiazepines to treat anxiety or depression and in one study the patients had used benzodiazepines to treat insomnia. Seven studies specified pre- and post-withdrawal psychiatric symptoms

Table 3
Summary statistics for each cognitive category including number of effect sizes (*n*), *d*, weighted *d*, and standard deviation of weighted *d*, listed in order of decreasing, weighted effect size, meta-analysis A

Category	*n*	*d*	Weighted *d*	S.D. of weighted *d*
Visuospatial	2	0.67	0.70	0.12
Attention/concentration	8	0.69	0.69	0.48
Problem-solving	1	0.64	0.64	n/a
General intelligence	2	0.58	0.62	0.21
Psychomotor speed	4	0.51	0.50	0.15
Sensory processing	2	0.47	0.37	0.46
Non-verbal memory	3	0.42	0.34	0.29
Speed of processing	6	0.35	0.32	0.44
Motor control/performance	4	0.28	0.21	0.27
Verbal memory	5	0.44	0.36	0.46
Working memory	4	0.19	0.15	0.41
Verbal reasoning	3	0.08	0.06	0.07
Overall	44	0.42	0.41	0.22

using a variety of subjective mood scales such as the Spielberger State-Trait Anxiety Inventory, the Hamilton Anxiety Scale, the Hamilton Depression Scale, and the Beck Depression Inventory. Mean scale scores in four of these would be considered in the clinical range and three in the non-clinical range. All seven studies reported that mean scores either remained the same following withdrawal (five studies) or significantly improved (two studies).

The most frequently used test was the digit-symbol-substitution test which was used in 50% of studies, followed by symbol copy and tapping rate, which were each used in 40% of studies. The most frequently measured category was attention/concentration contributing to 18% of the overall effect sizes, followed by speed of processing (14%).

From 10 studies covering 12 categories (a maximum of 120 possible effect sizes), 44 effect sizes were obtained. The mean weighted effect size was 0.41 (median = 0.37) with a standard deviation of 0.22. Unweighted mean effect sizes, weighted mean effect sizes, and the standard deviations of weighted mean effect sizes are reported in Table 3 for each cognitive category.

Compared to their initial assessment, previous long-term benzodiazepine users appeared to improve across all cognitive categories examined at follow-up assessment. All of the effect sizes were positive and ranged in magnitude from 0.06 to 0.70. It is clear from Fig. 1 that the 95% confidence intervals do not span zero for 5 of 11 effect sizes and therefore, that these effects are significant and different from zero (it was not possible to calculate effect sizes for all 12 categories as only one test within the problem-solving category was used).

Despite the very small sample size, an analysis conducted on moderator variables revealed one significant correlation between mean age and attention/concentration, $r\ (8) = -.72$, $P < .05$, suggesting that as age increases, post-withdrawal recovery decreases on tasks of attention/concentration.

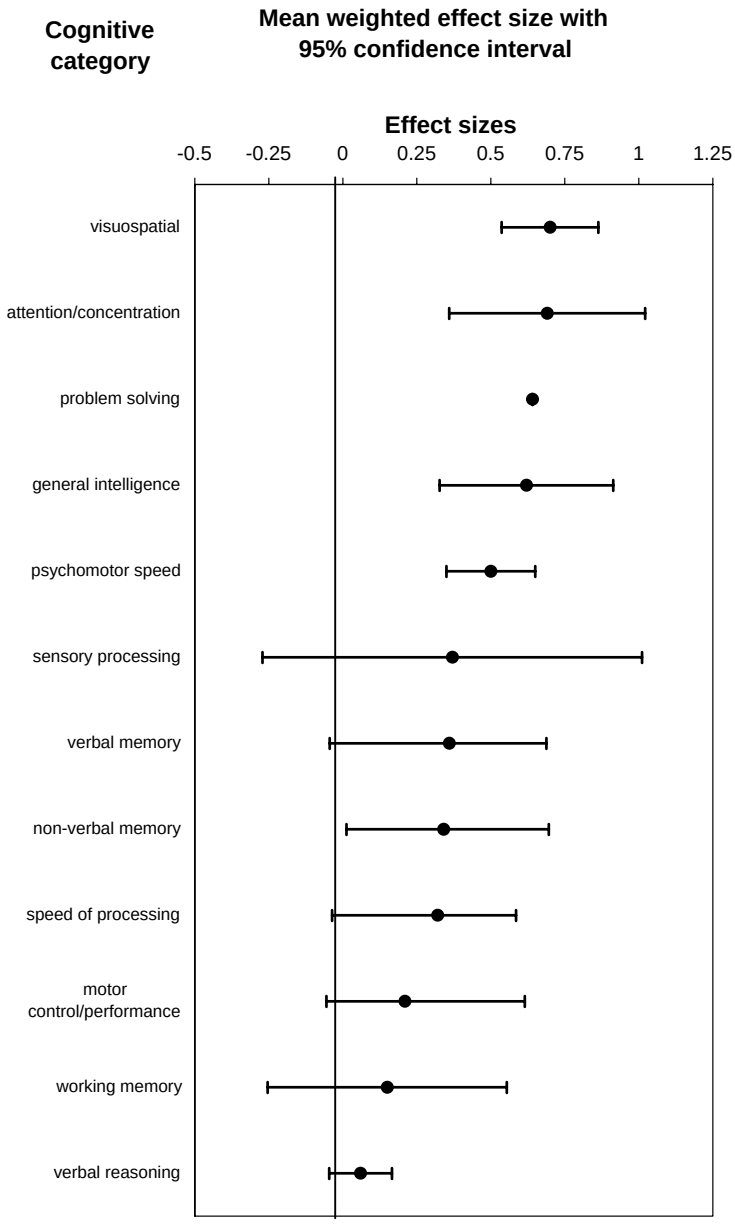

Fig. 1. Weighted mean effect sizes and 95% confidence intervals for each cognitive category, meta-analysis A.

2.2. Meta-analysis B: Are previous long-term benzodiazepine users still impaired at follow-up compared to controls or normative data?

Where studies employed a within-subjects design, published test norms were used for comparison in order to calculate effect sizes. This was necessary on three occasions (details

provided in Appendix B). Of the seven studies that used a control group, six were recruited from the general population with no history of anxiety and one study used both a healthy control group and an anxious control group. In this instance, the healthy control group was used in the calculation of effect sizes. Five of the studies matched controls on at least age and sex. Of those studies, two also matched on education, one on social class, one on marital status and type of work, and another on a pre-morbid IQ estimation (NART) score. In one study the methods used to match controls were not specified and one study did not use any method of matching.

The median number of benzodiazepine users in the studies was 21. The mean and standard deviation was 31.6 (27.4) (range = 10–96). Overall, 38.7% of the 284 subjects were male.

The mean age of participants was 42.7 years with a range of 21–75. All nine studies recruited subjects from those admitted to a hospital or clinic for the purposes of withdrawal or investigation of drug dependence. One study used DSM-III-R criteria to define dependence in their sample, and other used WHO (1964). Four studies described patients as either "long-term benzodiazepine users" or "benzodiazepine dependant." The remaining three studies used the terms "low therapeutic" or "normal dose" users. Seven studies specified the following benzodiazepines, listed in order of decreasing usage frequency, as those used by participants: diazepam, lorazepam, alprazolam, bromazepam, chlorodiazepoxide, oxazepam, flunitrazepam, clobazam, triazolam, and nitrazepam. The average daily dose expressed as a diazepam equivalent was 16.7 mg (S.D. = 21.1).

The mean length of benzodiazepine use, specified by all but one study, was 8.9 years. The range of length of use, specified by seven of the studies, was between 1 and 29 years. The median length of time assessments were carried out post-withdrawal was 3 months (range 1–65). The mean and standard deviation was 11.6 months (20.2).

Eight studies excluded patients with a history of heavy alcohol or other drug use and one study stated that 23% of patients had a history of alcohol use in excess of four standard drinks per day. Five studies specified the conditions for which subjects used benzodiazepines and for all five the conditions were anxiety or depression. Six studies specified pre- and post-withdrawal psychiatric symptoms using a variety of subjective mood scales including the Spielberger State-Trait Anxiety Inventory, the Hamilton Anxiety Scale, the Hamilton Depression Scale, and the Beck Depression Inventory. Mean scale scores in four of these would be considered in the clinical range and two in the non-clinical range. All six studies reported that mean scores either remained the same following withdrawal (four studies) or significantly improved (two studies).

Again the most frequently used test was the digit-symbol-substitution test (55% of studies) followed by symbol copy and tapping rate (44% of studies). The most frequently measured categories were speed of processing and psychomotor speed, each contributing to 16% of the overall effect sizes, followed by attention/concentration (14%).

From nine studies covering 12 categories (a maximum of 108 possible effect sizes), 37 effect sizes were obtained. The mean weighted effect size was −0.48 (median = 0.48) with a standard deviation of 0.45. Unweighted mean effect sizes, weighted mean effect sizes, and the standard deviations of weighted mean effect sizes are reported in Table 4 for each cognitive category.

Table 4
Summary statistics for each cognitive category including number of effect sizes (*n*), *d*, weighted *d*, and standard deviation of weighted *d*, listed in order of decreasing, weighted effect size, meta-analysis B

Category	*n*	*d*	Weighted *d*	S.D. of weighted *d*
Verbal memory	3	−0.89	−1.50	0.88
Psychomotor speed	6	−0.87	−0.78	0.60
Speed of processing	6	−0.59	−0.76	0.55
Motor control/performance	2	−0.65	−0.62	0.04
Working memory	2	−0.46	−0.58	0.42
Visuospatial	2	−0.30	−0.49	0.45
General intelligence	2	−0.34	−0.47	0.23
Attention/concentration	5	−0.63	−0.43	0.41
Non-verbal memory	3	−0.36	−0.26	0.10
Problem-solving	1	−0.11	−0.11	n/a
Verbal reasoning	3	−0.05	−0.02	0.05
Sensory processing	2	0.48	0.26	0.34
Overall	37	−0.41	−0.48	0.45

Compared to controls or norms, previous long-term benzodiazepine users performed more poorly across all cognitive categories, except sensory processing, when examined at follow-up assessment. Effect sizes ranged in magnitude from −1.50 to 0.26. It is clear from Fig. 2 that the 95% confidence intervals do not span zero for 8 of the 11 effect sizes and therefore, that most of these effects are significant and different to zero (it was not possible to calculate effect sizes for all 12 categories as only one test within the problem-solving category was used).

An analyses conducted on moderator variables did not reveal any significant relationships.

3. Discussion

In this study, meta-analytic techniques were used to integrate the available information on recovery and persistence of cognitive effects of long-term benzodiazepine use after withdrawal. In meta-analysis A, effect sizes greater than zero were found across all cognitive domains measured, suggesting that improvement in cognitive function does occur following withdrawal from long-term benzodiazepine use. Negative effect sizes were found across 11 of the 12 cognitive categories in meta-analysis B, suggesting that, except for sensory processing, patients who had withdrawn from long-term benzodiazepine use continued to perform more poorly than controls (or normative data) at follow-up.

According to Cohen (1988), effect sizes of $d = 0.20$, 0.50, and 0.80 are considered small, medium, and large in magnitude, respectively. The effect sizes found in this meta-analysis were substantial across most domains and many of the 95% confidence intervals did not span zero. It should also be noted that although a number of the 95% confidence intervals for the remaining effect sizes did span zero (most likely due to the small sample size), the direction and magnitude of these effect sizes should not be disregarded. In addition, the conservative approach to effect size calculation that was adopted, combined with the small sample of studies, may have resulted in an underestimation of the true effect size.

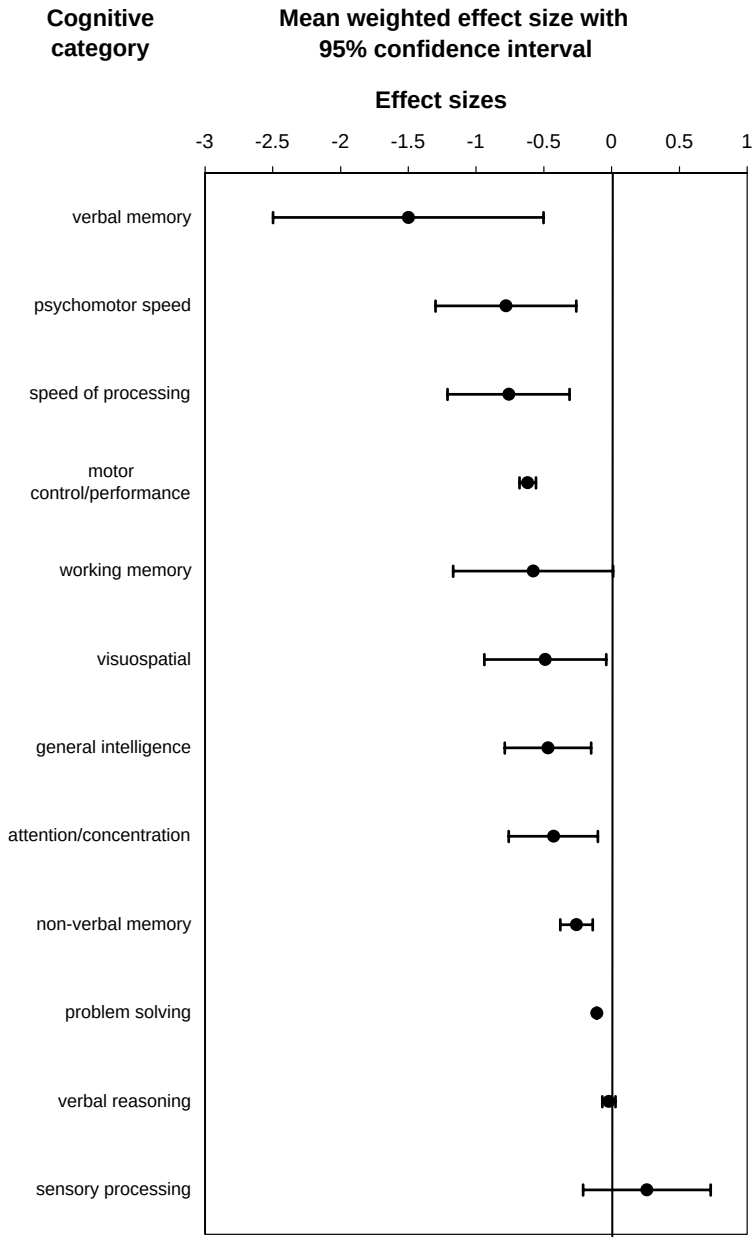

Fig. 2. Weighted mean effect sizes and 95% confidence intervals for each cognitive category, meta-analysis B.

The studies included in these meta-analyses were all published in peer-reviewed journals. Because studies that obtain non-significant results are less likely to be published, extracting data from only published results is likely to bias results in favor of a significant mean effect size (Rosenthal, 1991). This is known as *the file drawer problem* (i.e., unpublished studies with non-significant results tend to remain in file drawers).

To determine whether the results of a meta-analysis are susceptible to the file drawer threat, the number of additional unpublished or unretrieved studies that are likely to exist was estimated. Rosenthal has suggested that a conservative estimate for this tolerance level is $5k + 10$ where k is the number of studies retrieved. In the present meta-analyses, the tolerance level is estimated at $5(10) + 10 = 60$ and $5(9) + 10 = 55$ for meta-analyses A and B, respectively. The fail-safe n, which represents the number of studies obtaining null results that would need to be in existence to threaten the significant effect size found, was calculated to be four for meta-analysis A and five for meta-analysis B. Since the fail-safe n falls short of the tolerance level, the possibility of a file drawer problem should be considered for both meta-analyses, and the resulting effect sizes interpreted with caution.

Weighting for sample size had little impact on the effect size (0.42 unweighted compared to 0.40 weighted, meta-analysis A; −0.41 unweighted compared to −0.48 weighted, meta-analysis B), suggesting that the inclusion of studies with small samples that yielded large effect sizes did not artificially increase the effect size. Due to the small number of studies that met inclusion criteria in this meta-analysis, and the limited information provided on relevant characteristics, an analysis of heterogeneity (Q) (Hedges & Olkin, 1985) was considered inappropriate.

The small number of studies included in the meta-analysis also resulted in insufficient data to conduct a thorough investigation of the contribution of moderator variables. The significant negative correlation, between age and overall effect size, suggests that improvement in attention/concentration after withdrawal decreases as age increases. This finding may be related to deficits in performance on tasks of attention and concentration observed to accompany normal aging (Lezak, 1995). However, it should be interpreted with caution due to the inflation in type I error rate associated with the number of analyses conducted.

As expected, no other variables examined in either meta-analysis were found to be significantly related to any of the individual cognitive domain effect sizes, nor to the overall effect size. One trend of interest was a relationship between publication year and overall effect size, such that studies published later than 1994 tended to have larger effect sizes. Closer examination of studies published later than 1994 did not reveal any obvious differences in any participant or study variables compared to earlier published studies. It may be the case that some real differences do exist that were undetectable given the small sample size.

In a previous meta-analysis that we have conducted looking at the long-term cognitive effects of benzodiazepine use (Barker, Greenwood, Jackson & Crowe, in press) we found moderate to large effect sizes across all cognitive areas studied indicating that long-term benzodiazepine users are impaired across many cognitive areas. It was further noted that although most investigations of the effects of long-term use of benzodiazepines tend to focus on one or two specific areas of cognition, integrating all of the available evidence indicates that long-term benzodiazepine users may be affected in a generalized rather than specific way, with some areas being more affected than others. The current meta-analyses go further to suggest that these patients do show significant recovery of function in many areas after withdrawal. However, there remains a significant impairment in most areas of cognition compared to controls or normative data.

A comparison of the degree of recovery and residual impairment can be made by comparing the effect sizes for each cognitive category between Table 3 (improvement) and Table 4 (resid-

ual impairment). For example, the degree of recovery that appears to take place in the domain of attention/concentration is significant with a moderate to large effect size ($d = 0.69$), and a significant, small to moderate effect size ($d = -0.43$) indicating the level of residual impairment.

One important consideration in interpreting these results relates to the possibility of reduced performance as a function of return of pre-morbid symptomology, particularly when comparing to healthy controls. As previously stated, the majority of patients in the analyzed studies had used benzodiazepines to treat anxiety or depression and therefore, may have baseline cognitive deficits unrelated to long-term benzodiazepine use. Although this explanation must be considered, results of previous research support the notion that the observed deficits may be unrelated to patient's psychiatric symptomology. Gorenstein et al.'s (1995) well-controlled comparison of neuropsychological test performance in anxious drug-free patients and normal controls found no significant difference in performance on a variety of neuropsychological tests except for a ball-bearing test measuring motor co-ordination. The anxious control group was also compared to chronic benzodiazepine users and it was concluded that, despite similar anxiety levels, poorer performance could not be attributed to the anxiety disorder itself. It is also unlikely that the observed deficits may be a function of increased anxiety levels caused by withdrawal of medication. Seven out of 10 studies reported pre- and post-withdrawal anxiety or depression levels and in all of these, patient's symptoms either improved significantly or remained the same.

The presence or absence of dementia in older subjects was not typically reported and the possibility of a dementing process as a contributing factor to the cognitive deficits observed cannot be ruled out. This seems unlikely, however, given that the mean age in nine of the studies was between 39 and 47 years, and the remaining study, which examined mainly older adults, did in fact screen for and exclude patients with dementia. Furthermore, the study in question was not included in meta-analysis B for reasons previously outlined in Section 1.

In order to more thoroughly investigate the cognitive impact of long-term benzodiazepine use, further research is clearly necessary. Conducting large-scale studies examining many areas of cognition is an unlikely scenario and therefore, a series of smaller, methodologically sound studies, which comprehensively investigate a small number of cognitive domains, is a more feasible possibility. Provided these data were presented in a manner amenable to meta-analyses, a thorough systematic and statistical evaluation of this area could then follow.

The results of this study support extreme caution in the use of long-term benzodiazepine therapy. Although these findings suggest that previous long-term benzodiazepine use may lead to impairments in cognition, some degree of improvement in cognitive function after withdrawal was observed, suggesting that previous benzodiazepine users are likely to experience the benefit of improved cognitive functioning after withdrawal.

However, data from this study do not support a full recovery, at least in the first 6 months following cessation and suggests that there may be some permanent deficits or deficits that take periods longer than 6 months to completely recover. Professionals and patients should be well aware of the potential for cognitive impairments prior to prescribing or taking benzodiazepine medication on a long-term basis. It is imperative that patients are provided with this information so they may then make an informed decision as to whether to use benzodiazepines, weighing up any potential treatment benefits against the likelihood of dependence and cognitive impairment.

Appendix A

Assessment tools used, cognitive areas measured and cognitive categories.

	Test used	Area measured (Lezak, 1995; Spreen & Strauss, 1998)	Category
1	Seashore Rhythm Test	Auditory perception	Sensory processing
2	Witkins Rod and Frame Test	Field dependence	Sensory processing
3	Digit-Symbol-Substitution Test	Psychomotor speed	Psychomotor speed
4	Symbol Copy	Psychomotor speed	Psychomotor speed
5	Gollin Picture Completion Test	Visual memory	Non-verbal memory
6	Bender Gestalt	Visuoconstructional ability	Non-verbal memory
7	Tactual Performance Test	Tactile memory	Non-verbal memory
8	Spatial Recognition Task	Visual recognition	Non-verbal memory
9	Memory for Designs	Immediate visuospatial memory	Non-verbal memory
10	Visual Perceptual Analysis	Visual information processing	Visuospatial
11	Koh's Blocks/Block Design	Visuoconstructive skill	Visuospatial
12	Four Choice Reaction Time (FCRT)	Reaction time	Speed of processing
13	Leeds Psychomotor Test Apparatus (Critical Flicker)	Reaction time	Speed of processing
14	Reaction Time Test	Reaction time	Speed of processing
15	Trails B	Visual search	Speed of processing
16	Tower of Hanoi	Problem-solving ability	Problem-solving
17	WCST/Bexley–Maudsley Category Sorting Test	Problem-solving ability	Problem-solving
18	Category Test	Abstract concept formation	Problem-solving
19	Vigilance Test Paradigm	Attention/concentration	Attention/concentration
20	The d2 Test	Concentration	Attention/concentration
21	Cancellation Task	Visual attention	Attention/concentration
22	Trails A	Visual conceptual tracking	Attention/concentration
23	Sensory Threshold Detection Test (STD)	Visual attention	Attention/concentration
24	Selective Reminding Test	Immediate memory/verbal learning	Verbal memory
25	Word Lists	Immediate memory/verbal learning	Verbal memory

Appendix A. *Continued*

	Test used	Area measured (Lezak, 1995; Spreen & Strauss, 1998)	Category
26	Logical Memory/Prose Recall/Story Memory	Immediate/delayed recall	Verbal memory
27	Word Stem Completion/ Priming Task	Implicit memory	Verbal memory
28	Paired Associates	Associate learning	Verbal memory
29	Paired Associate Interference Task	Procedural learning	Verbal memory
30	Vocab/Info Score	General intelligence	General intelligence
31	WAIS Score/SRB Score	General intelligence	General intelligence
32	National Adult Reading Test	Pre-morbid IQ	General intelligence
33	Ball-Bearing Test	Motor control/ performance	Motor control/ performance
34	Finger Tapping/Tapping Rate	Motor control/ performance	Motor control/ performance
35	Recognition Test	Recognition memory	Working memory
36	Digit Span	Immediate verbal recall	Working memory
37	Digit Span Backwards	Working memory	Working memory
38	Digit Span Forwards	Working memory	Working memory
39	Thurstone Figure Classification Test	Reasoning	Verbal reasoning
40	Controlled Oral Word Association Test/Word Fluency	Verbal fluency	Verbal reasoning
41	Synonyms	Verbal understanding	Verbal reasoning

Appendix B

Articles used to determine $S.D._p$ or comparison norms where this information was not provided in the study.

Test	Cognitive category	Reference
Critical Flicker Fusion	Speed of processing	Curran, Hindmarch, Wattis, and Shillingford (1990)
Digit-Symbol-Substitution Test	Psychomotor speed	Wechsler (1981)

References

References marked with an asterisk (*) indicate studies included in meta-analysis A; references marked with a symbol (†) indicate studies included in meta-analysis.

Aranko, K., Mattila, M. J., & Seppala, T. (1983). Development of tolerance and cross-tolerance to the psychomotor actions of lorazepam and diazepam in man. *British Journal of Pharmacology, 15*, 545–552.

Ashton, H. (1986). Adverse effects of prolonged benzodiazepine use. *Adverse Drug Reaction Bulletin, 118*, 440–443.

Balter, M. B., Manheimer, D. I., Mellinger, G. D., & Uhlenhuth, E. H. (1984). A cross-national comparison of anti-anxiety/sedative drug use. *Current Medical Research and Opinion, 8*, 5–19.

Barker, M. J., Greenwood, K. M., Jackson, M., & Crowe, S. F. (in press). The cognitive effects of long-term benzodiazepine use: A meta-analysis. *CNS Drugs*.

*,†Bergman, H., Borg, S., Engelbrektson, K., & Vikander, B. (1989). Dependence on sedative-hypnotics: Neuropsychological impairment, field dependence and clinical course in a 5-year follow-up study. *British Journal of Addiction, 84*, 547–553.

*,†Bergman, H., Borg, S., & Holm, S. (1980). Neuropsychological impairment and exclusive abuse of sedatives or hypnotics. *American Journal of Psychiatry, 137*, 215–217.

*,†Birzele, H. J. (1992). Benzodiazepine induced amnesia after long-term medication and during withdrawal. *European Review of Applied Psychology, 42*, 277–282.

Brosan, L., Broadbent, D., Nutt, D., & Broadbent, M. (1986). Performance effects of diazepam during and after prolonged administration. *Psychological Medicine, 16*, 561–571.

Chaves, M. L., Bianchin, M., Peccin, S., & Rotta, F. (1993). Chronic use of benzodiazepines and cognitive deficit complaints: A risk factor study. *Italian Journal of Neurological Sciences, 14*, 429–435.

Chen, Y. (1990). Long-term benzodiazepine treatment: Is it ever justified? *Human Psychopharmacology: Clinical and Experimental, 5*, 301–312.

Cooper, A. J. (1982). Benzodiazepines: Toward more logical use. *Scottish Medical Journal, 27*, 297–304.

*,†Curran, H. V. (1992). Memory functions, alertness and mood of long-term benzodiazepine users: A preliminary investigation of the effects of a normal daily dose. *Journal of Psychopharmacology, 6*, 69–75.

Curran, H. V., Bond, A., O'Sullivan, G., Bruce, M., Marks, I., Lelliot, P., Shine, P., & Lader, M. (1994). Memory functions, alprazolam and exposure therapy: A controlled longitudinal study of agoraphobia with panic disorder. *Psychological Medicine, 24*, 969–976.

Curran, S., Hindmarch, I., Wattis, J. P., & Shillingford, C. (1990). Critical flicker fusion in normal elderly subjects: A cross-sectional community study. *Current Psychology: Research and Reviews, 9*, 25–34.

Golombok, S. (1989). Causes, effects and treatment of long-term benzodiazepine use: A review of psychological perspectives. *Human Psychopharmacology: Clinical and Experimental, 4*, 15–20.

Golombok, S., Moodley, P., & Lader, M. (1988). Cognitive impairment in long-term benzodiazepine users. *Psychological Medicine, 18*, 365–374.

*,†Gorenstein, C., Bernik, M. A., & Pompeia, S. (1994). Differential acute psychomotor and cognitive effects of diazepam on long-term benzodiazepine users. *International Clinical Psychopharmacology, 9*, 145–153.

*,†Gorenstein, C., Bernik, M. A., Pompeia, S., & Marcourakis, T. (1995). Impairment of performance associated with long-term use of benzodiazepines. *Journal of Psychopharmacology, 9*, 313–318.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

Kilic, C. (1999). Long-term effects of alprazolam on memory: A 3.5 year follow-up of agoraphobia/panic patients. *Psychological Medicine, 29*, 225–231.

Lader, M. H. (1982). Psychological effects of buspirone. *Journal of Clinical Psychiatry, 43*, 62–67.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lucki, I., & Rickels, K. (1986). The behavioral effects of benzodiazepines following long-term use. *Psychopharmacology Bulletin, 22*, 424–433.

Lucki, I., Rickels, K., & Geller, A. M. (1985). Psychomotor performance following long-term use of benzodiazepines. *Psychopharmacology Bulletin*, *21*, 93–96.

Lucki, I., Rickels, K., & Geller, A. M. (1986). Chronic use of benzodiazepines and psychomotor and cognitive test performance. *Psychopharmacology*, *88*, 426–433.

*,† Petursson, H., Gudjonsson, G. H., & Lader, M. H. (1983). Psychometric performance during withdrawal from long-term benzodiazepine treatment. *Psychopharmacology*, *81*, 345–349.

*,† Rickels, K., Lucki, I., Schweizer, E., Garcia-Espana, F., & Case, W. G. (1999). Psychomotor performance of long-term benzodiazepine users before, during, and after benzodiazepine discontinuation. *Journal of Clinical Psychopharmacology*, *19*, 107–113.

Rosenthal, R. (1991). *Meta-analytic procedures of social research* (Rev. ed.). London: Sage Publications.

*,† Sakol, M. S., & Power, K. G. (1988). The effects of long-term benzodiazepine treatment and graded withdrawal on psychometric performance. *Psychopharmacology*, *95*, 135–138.

* Salzman, C., Fisher, J., Nobel, K., & Glassman, R. (1992). Cognitive improvement following benzodiazepine discontinuation in elderly nursing home residents. *International Journal of Geriatric Psychiatry*, *7*, 89–93.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.

*,† Tata, P. R., Rollings, J., Collins, M., Pickering, A., & Jacobson, R. R. (1994). Lack of cognitive recovery following withdrawal from long-term benzodiazepine use. *Psychological Medicine*, *24*, 203–213.

*,† Toenne, U., Hiltunen, A. J., Vikander, B., Engelbrektsson, K., Bergman, H., Bergman, I., Leifman, H., & Borg, S. (1995). Neuropsychological changes during steady-state drug use, withdrawal and abstinence in primary benzodiazepine-dependent patients. *Acta Psychiatrica Scandinavica*, *91*, 299–304.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised manual*. San Antonio, TX: The Psychological Corporation.