Introduction

Posttraumatic stress disorder (PTSD) is often a chronic and debilitating disorder characterized by four symptom clusters: re-experiencing, avoidance, arousal, and negative changes in mood and cognition, as defined in the DSM-V (APA, 2013). Epidemiological studies have shown that the disorder has a high lifetime prevalence rate of 6.8% in the general population (Gradus, 2007; Kessler et al., 2005). The prevalence rate in Iraq and Afghanistan veterans is significantly higher, which is estimated to range from 8.5% to 24.5% in the USA (Hoge et al., 2006, 2007; Milliken et al., 2007), with lower rates in UK veterans ranging from 4% to 6% (Browne et al., 2007; Hotopf et al., 2006). The Department of Veteran Affairs (VA) estimates that only 9.5% of veterans diagnosed with PTSD are actually receiving treatment (Seal et al., 2010). There may be many reasons for such a small minority of veterans seeking treatment. Some research has shown that many fear the associated stigma related to seeking treatment, as well as institutional barriers such as lack of skill and sensitivity by VA staff (Quimette et al., 2011). Another possibility may be due to the marginal efficacy of current pharmacotherapy and psychotherapeutic options. Currently, only two pharmaceuticals are approved for treating PTSD: sertraline and paroxetine (Jeffreys, 2009; Pollack et al., 2001). Many psychotherapeutic options are available but have high dropout rates for a variety of reasons. First, trauma often affects the victim’s ability to form trusting interpersonal relationships, which can affect the “working alliance” between the patient and therapist (Doukas et al., 2014). Additionally, many people with PTSD have a small window of “optimal arousal” or “therapeutic threshold,” which limits therapeutic effectiveness and contributes to a high dropout rate (Foia and Kozak, 1986). A key symptom of PTSD is avoidance, so it is no surprise that re-emerging thoughts brought up in therapy can overwhelm the patient and cause them to dropout. Eftekhari et al. (2013) explored some possible explanations for the high dropout rates in service members using the VA but found that the majority (40.8%) reported non-specific reasons, while 35.6% reported that the therapy increased distress.

Prolonged exposure (PE) therapy is one of the most widely accepted treatments for PTSD and was specifically designed for treating the disorder (Foia, 2011). It requires the patient to re-live their traumatic experiences repeatedly within a safe context in a
process referred to as “flooding.” Constant exposure to the traumatic thoughts decoupled from actual threat can induce extinction of the trauma response (McLean et al., 2015). However, only a small minority of veteran patients (6.3%) are treated with PE (Shiner et al., 2013) because it is emotionally demanding and often aggravates the patient’s symptoms before they improve (Steenkamp and Litz, 2014). Considering that only 20–30% of patients with PTSD respond to pharmacotherapy (Stein et al., 2009) and dropout rates from psychotherapy are estimated to be 30% (Cloitre, 2009), it is apparent that new treatment options must be developed.

An emerging treatment for PTSD uses ±3,4-methylenedioxymethamphetamine (MDMA) as a therapeutic catalyst during psychotherapy sessions. MDMA is a ring-substituted amphetamine, with structural similarities to the hallucinogenic drug mescaline (Green et al., 2003). The drug was originally used as a psychotherapeutic adjunct by psychiatrists and psychologists until it started to be used recreationally by the public in the 1980s. In 1985, it was categorized as a Schedule I drug and banned from medical use. The subjective psychoactive effects of MDMA include reduced anxiety, acute antidepression, increased insight (largely dependent on 5-HT transporter modulation), accelerated thinking and euphoria (modulated in part by D2 receptors), enhanced visual and auditory perception (modulated in part by the 5-HT2 receptors), and increased prosocial behaviors such as a sense of trust and bonding (partly dependent on increased oxytocin release; Dumont et al., 2009; Kirkpatrick et al., 2015; Liechti and Vollenweider, 2001). It is hypothesized that the increased sense of bonding and trust allows for a better “working alliance” between the therapist and patient (a key issue in psychotherapy with PTSD patients), while the reduced anxiety and increased insight widen the window of “optimal arousal” and “therapeutic threshold” (Mittheofer et al., 2011).

Some have claimed that unique psychopharmacological properties of MDMA make the drug well suited for treating PTSD (Amoroso, 2015; Johansen and Krebs, 2009). Carhart-Harris et al. (2014) investigated the neural and psychological responses to positive and negative autobiographical memories after participants ingested MDMA. The participants who ingested MDMA reported their worst memories were less negative than those who had not ingested the drug did. The researchers found that MDMA attenuated activation in the left anterior cingulate cortex, left amygdala, and temporal cortex, while activating executive regions of the hippocampus, which other studies (e.g., Brenner et al., 2005) have shown to be important brain regions involved with PTSD.

The typical treatment course for MDMA-assisted psychotherapy (MDMA-AP) is one to three drug sessions lasting eight hours and several follow-up non-drug sessions. Two clinical trials have been completed and published with promising results (i.e., Mittheofer et al., 2011; Oehen et al., 2013). In contrast, PE therapy sessions typically last about an hour, and the number of sessions ranges from about 6 to 19 (Powers et al., 2010).

One of the best ways to determine if an emerging treatment is worth pursuing is to compare it to existing treatments. Some suggest that it is an ethical imperative to compare the results of new treatments with those of “best-available” treatments (Hill, 1994). There have been several meta-analyses published comparing the efficacy of new PTSD treatments to PE (Benish et al., 2008; Sherman, 1998; Van Etten and Taylor, 1998), but MDMA-AP has yet to be compared to an existing “best-available” treatment.

Often, statistically significant results using p-values are overemphasized, while the magnitude of the results is overlooked (Cohen, 1995; Dar et al., 1994). This is problematic, as a large sample size can produce statistically significant p-values while the treatment effect may be negligible. Inversely, a small sample size can produce insignificant p-values but have a large effect size (Fritz et al., 2012). Therefore, effect size is often a more useful metric when comparing treatment studies with large differences in sample size (Kraemer and Kupfer, 2006).

In this analysis, the effect sizes of PE in treating PTSD are compared to those of the published MDMA-AP trials. Mittheofer et al. (2011) conducted a randomized, placebo-controlled, double blind, crossover design study consisting of 23 participants. Oehen et al. (2013) conducted the second published MDMA-AP study to date, with 12 treatment-resistant PTSD patients. The present study compares these studies to the only meta-analysis published on the effectiveness of PE by Powers et al. (2010). Both the primary outcome measures (clinician-observed PTSD symptoms) and secondary outcome measures (self-reported symptoms) of the three studies are discussed.

Methods

Procedure

The primary goals of this study are to report a preliminary meta-analysis of MDMA-AP and to compare the results to PE, the most widely accepted treatment of PTSD. This was done by first conducting a literature search for published clinical trials of MDMA-AP and a reliable meta-analysis on PE. Effect sizes are compared, as well as dropout rates. The dropout rates were reported in the MDMA-AP clinical trials but not the PE meta-analysis, so those were calculated and reported as well.

Study selection. Three databases were used (PsycINFO, PubMed, and GoogleScholar) to find the only meta-analysis on PE and the two published articles on MDMA-AP. The search for the meta-analysis of PE included the terms “meta,” “metaanalytic,” “exposure,” “PTSD,” and “prolonged,” while the search for MDMA-AP clinical trials included the terms “MDMA,” “double blind,” “placebo controlled,” “posttraumatic stress,” and “therapy.” Inclusion criteria for the MDMA-AP studies were as follows: (1) participants had to meet DSM-III-R, DSM-IV, or DSM-IV-R criteria for PTSD; (2) the study had to be randomized, double blind, and placebo controlled (unless a crossover design was used); (3) and there had to be enough participants to provide inferential statistics. Three MDMA studies were found, but one was excluded (i.e., Bouso et al. 2008) because it was terminated prematurely, did not provide adequate experimental controls, and did not report inferential statistics.

Effect size calculation. Once the MDMA studies were selected, the effect sizes for both primary and secondary outcome measures were compared to PE. The meta-analysis on PE by Powers et al. (2010) provided Hedges’ g, while the MDMA studies reported Cohen’s d. The effect sizes were corrected according to Hedges and Olkin (1985) using the formula:
\[ g = d \left(1 - \frac{3}{4(n_1 + n_2)} - 9 \right) \]

Mitchoefer et al. (2011) only reported the effect size for primary outcomes, so the effect size for secondary outcomes were converted from the reported \( F \)-test to Cohen’s \( d \) using the formula:

\[ d = \sqrt{\left(\frac{n_1 + n_2}{n_1 n_2}ight) \left(\frac{n_1 + n_2}{n_1 + n_2 - 2}\right)} \]

(Thalheimer and Cook, 2002). The average effect size for the MDMA studies was then calculated using the same formula:

\[ \bar{g} = \frac{\sum w_j g_j}{\sum w_j} \]

used by Powers et al. (2010), where \( w_i \) is the weight for each study and \( g_i \) is the effect size for each study.

**Calculation of dropout rates.** Powers et al. (2010) did not calculate an average dropout rate. Therefore, dropout percentages of each study within the PE meta-analysis were calculated and are reported here. Only participants assigned to the PE condition in each study were included in the calculation. The MDMA studies employed crossover and active placebo designs, so the total number of participants dropping out of the study was divided by the total number of participants in the study. The percentages per treatment type and standard deviations are reported below.

**Results**

**Hypothesis 1: MDMA-AP will have a larger cumulative effect size than PE will for primary outcome measures**

The overall effect size reported in the meta-analysis of PE was large for primary outcome measures (Hedges’ \( g = 1.08 \); SE = 0.20; 95% CI 0.69–1.46; \( p < 0.001 \)). The cumulative effect size for primary outcome measures calculated for MDMA-AP in this analysis was also large (Hedges’ \( g = 1.17 \); SE = 0.09; 95% CI 0.38–1.90; \( p = 0.033 \); Table 1).

**Hypothesis 2: MDMA-AP will have a larger cumulative effect size than PE will for secondary outcome measures**

The overall effect size reported in the meta-analysis of PE was also large for secondary outcome measures (Hedges’ \( g = 0.77 \); SE = 0.12; 95% CI 0.53–1.01; \( p < 0.001 \)). The effect size for secondary outcome measures was not reported by Mitchoefer et al. (2011), but significant improvements were found in the MDMA-AP group (time\( \times \)group interaction \( F[1, 17] = 3.290; p = 0.027 \)). The \( F \)-test was used to calculate the effect sizes for the secondary outcome measures, which was large (Hedges’ \( g = 0.83 \)). Oehen et al. (2013) also did not report the effect size for secondary outcome measures, but this was found to be large (Hedges’ \( g = 0.97 \)) by another researcher (Chabrol and Oehen, 2013). The cumulative effect size for secondary outcome measures calculated for MDMA-AP in this analysis was large (Hedges’ \( g = 0.87 \); 95% CI 0.01–1.79; \( p = 0.049 \); Table 1).

**Hypothesis 3: MDMA-AP will have lower cumulative dropout rates than PE will**

The average percentage of participants that dropped out of the studies included in the PE meta-analysis was 27.0% (SD = 10.8%). An average of 12.7% (SD = 5.6%) of participants dropped out of the MDMA-AP studies (Table 1).

**Validity of meta-analysis**

**Heterogeneity.** Heterogeneity was calculated using the \( F \)-statistic instead of Cochran’s Q (as done in the PE meta-analysis) because it is known to be a better reporter of heterogeneity (Hardy and Thompson, 1998; Paul and Donner, 1992). The \( F \)-statistic describes the percentage of heterogeneity across studies rather than the variation due to chance (Higgins et al., 2003). An \( F \) of 0% indicates no observed heterogeneity, with increasing values meaning more heterogeneity (Higgins et al., 2003). For the MDMA-AP studies, \( F = 0.0% \). Powers et al. (2010) reported Cochran’s \( Q(12) = 59.90 (F = 79.9%) \).

**Publication bias.** Research has shown that significant results in clinical trials are more than three times as likely to be published than those with insignificant results (Dickersin et al., 1987). Because of this publication bias, also referred to as the “File Drawer Problem,” effect sizes reported in meta-analyses may be overestimated. A fail-safe \( N \) should be reported in meta-analyses to account for the possible null effects of unpublished work (Rosenthal, 1979). The fail-safe \( N \) determines the number of null-effect studies required to reverse statistical significance of the findings in a meta-analysis. The formula:

\[ N: X = \frac{K(KZ^2 - 2.706)}{2.706} \]

was used to compute the fail-safe \( N \) using \( z \)-scores, where \( K \) is the number of studies and \( Z \) is the mean \( Z \) from each study (Rosenthal, 1991). The number of studies needed to reduce the overall effect size to a non-significant level must exceed 5 \( K+10 \).
in a robust meta-analyses, or 75 and 20 for the PE and MDMA-AP studies, respectively (Rosenthal, 1991). Powers et al. (2010) reported that 446 current or future unpublished studies with an effect size of 0 would be required to bring the overall effect size of the meta-analysis within the non-significant range. The number of studies with an effect size of 0 required to bring the MDMA-AP meta-analysis within non-significant range was calculated to be 135, meaning both the PE and MDMA-AP meta-analyses are robust.

Discussion

The results of this analysis suggest that MDMA-AP has comparable treatment outcomes to PE. The MDMA-AP studies showed a large cumulative effect size on primary outcome measures (Hedges’ g=1.17). The PE meta-analysis also reported a large effect size on primary outcome measures (Hedges’ g=1.08). The effect sizes for secondary outcome measures are large in both the MDMA-AP and PE studies.

One issue with PE is that the patient is put into a heightened state of arousal, with little time to process the experience before leaving the therapy session. The MDMA-AP had much longer therapy sessions typically lasting eight hours. PE and MDMA-AP offer two very different approaches to therapy. Some researchers and clinicians have claimed that PE is too “rigid” and “insensitive” to meet the needs of some patients (Feeney et al., 2003; Olatunji et al., 2009). In contrast, MDMA-AP offers a patient-centered approach, which allows the patient to explore aspects of the trauma that may be outside of the reaches of PE. This is not to say that PE is without proven benefits and efficacy. These results simply suggest that MDMA-AP may be a superior alternative for those who do not respond to PE, which is a much more available treatment.

PE has been shown to have high dropout rates (Schottenbauer et al., 2008), which may result from the avoidant nature of the disorder. Although the MDMA-AP studies had much smaller sample sizes, they had lower percentages of participants dropping out of treatment. One possible reason for this may be due to the long eight-hour therapy sessions, which may make the patient feel as though the therapist is more committed to their recovery compared with the 60-minute sessions typically offered by PE. Importantly, the MDMA-AP studies seemed to be very safe, as there were no psychiatric or physical emergencies. Mithoefer et al. (2011) reported that one of the two participants that dropped out did so because of difficulty with traveling to the study site, while the other dropped out because she was required to resume taking a medication for depression. Oehen et al. (2013) reported that two participants dropped out of treatment due to adverse effects, despite one of them being assigned to the active placebo group.

An important difference between the study designs is that the PE studies employed either psychological placebo conditions or waitlist controls (six employed psychological placebo, five waitlist controls, and two a combination of both). The MDMA-AP studies used active placebos, where participants in the control group were exposed to the same psychotherapy as those in the treatment group without the active dose of MDMA. This is important because the magnitude of effect sizes from the PE studies are based on those who received treatment and those who received nothing, whereas the magnitude of the effect sizes from the MDMA-AP studies are based on the effect of MDMA within a particular treatment.

Another potential confound of the PE meta-analysis is it included a considerable number of participants on various psychotropic medications, which may have inflated the effect sizes. The participants in the MDMA-psychotherapy studies were required to titrate off their medications five half-lives prior to treatment to avoid drug interactions and confounding data.

Some limitations should be mentioned. First, the meta-analysis on PE included 13 studies with much larger sample sizes (n=675), whereas the two MDMA-AP trials had much smaller sample sizes (n=37). However, part of the reason effect sizes are compared here is because they are largely unaffected by sample size. Another weakness is that there were differences in the participant demographics. The inclusion criteria for the PE meta-analysis were that participants simply had to meet full DSM criteria for PTSD. The two MDMA-AP studies required that the participants meet DSM criteria for PTSD, as well as have chronic and treatment resistant symptoms. For instance, the average duration of PTSD symptoms in Mithoefer et al. (2011) was 19 years. Powers et al. (2010) did not report the average duration of PTSD symptoms in the participants, so this comparison cannot be made. Finally, MDMA-AP and PE are very different therapies, which makes it difficult to make direct comparisons with conclusive results.

People with PTSD have a limited variety of treatment options. Many people suffering from PTSD cannot tolerate exposure therapies due to their emotionally taxing nature. Also, the response rate to pharmacotherapy is low in patients with PTSD. With an average of 22 veterans a day committing suicide, new innovations for treating mental illnesses such as PTSD is imperative (Kemp and Bossarte, 2013). At a cost of roughly US$43.2 billion annually, its economic impact is massive (Greenberg et al. 1999). With the VA spending millions of dollars on PTSD research (GAO, 2011) and billions in treatment (CBO, 2011), it is imperative that emerging treatments such as MDMA-AP become available, and are thoroughly considered, in order to help the many suffering from PTSD.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References


Bremner JD, Vernetten E, Schmahl C, et al. (2005) Positron emission tomographic imaging of neural correlates of a fear acquisition and...


