bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Coordinate based meta-analysis of whole-brain voxel-based morphometry studies does not show evidence of grey matter loss specific to PTSD

Christopher R. Tench

Radu Tanasescu

Ketan D. Jethwa

Cris S. Constantinescu

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Neuroimaging studies have detected structural alteration in post-traumatic stress disorder (PTSD), but findings are inconsistent. This might be explained by heterogeneity between subjects with PTSD in terms of common comorbidities such as depressive and anxiety disorders and also in traumatic experience. Despite this, coordinate based meta-analysis (CBMA) has been used to try and identify localised grey matter changes, and does suggest some PTSD specific pathology. However, there are multiple technical issues that make the meta-analytic evidence questionable, warranting a re-evaluation.

A literature search for voxel-based morphometry studies was performed. Only whole-brain studies using subjects with a current diagnosis of PTSD, and having a comparison group of either healthy or trauma exposed controls, were included. Twenty one voxel-based morphometry studies met the inclusion criteria. CBMA was performed to identify altered grey matter (GM) structures.

Using a novel coordinate based random effect size meta-analysis, no grey matter structure was identified as being consistently altered in PTSD compared to controls. This was also verified using the activation likelihood estimate algorithm.

There is no evidence, from CBMA, of consistent localised grey matter changes specific to PTSD. Inconsistency may reflect true heterogeneity in PTSD pathology or methodological issues with imaging and/or analysis, limiting the detection of PTSD specific pathology.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Post-traumatic stress disorder (PTSD) is a mental health condition characterised by symptoms of excessive fear response that can be reliably provoked. Regardless of trigger, PTSD causes clinically significant distress, and impairs social interaction and capacity to work. Neuroimaging has been used to explore structural brain differences between PTSD and control groups using voxel based morphometry (VBM), which compares either GM concentration or absolute volume [1]. These are essential to the understanding of the condition, and indeed neurocircuitry models of PTSD have been suggested [2] based, at least in part, on evidence gathered from neuroimaging [3]. The amygdala is involved with fear processing and hyperactivity of the amygdala is associated with PTSD. This is thought to be the result of impaired inhibitory activity from the ventral/medial prefrontal cortex (vmPFC) and hippocampus [2]. Evidence of reduced volume of these structures in PTSD patients has come from region of interest (ROI) based neuroimaging studies and has been summarised in meta-analyses [4, 5]. The functional relationship of the amygdala and the medial prefrontal cortex has also been found to be inversely correlated to symptom severity [6].

Neuroimaging experiments involving a-priori regions of interest have been performed in PTSD, but are inevitably prone to publication bias. VBM is less biased in that regions hypothesised to be affected can be analysed in context with whole brain effects; publication bias is still a concern, but not directed specifically at hypothesised brain structures. Reports of VBM experiments include the coordinates (in standardised stereotactic space) of abnormalities detected anywhere in the brain. To be of any significance beyond the study cohort the results must be at least partially PTSD specific. However, studies often involve small subject numbers, which can limit generalisability [2]. Furthermore, while within-study the subjects may have experienced the same traumatic event, between-studies there is heterogeneity. There is also heterogeneity between subjects in terms of common comorbidities such as depressive and anxiety disorders [7]. Consequently, assessment of generalizable PTSD specific pathology using neuroimaging is not straightforward.

Whole-brain (WB) VBM analyses lend themselves to a coordinate based meta-analysis (CBMA) approach. The aim of CBMA is to use the coordinates from experiments testing

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

related hypotheses to find the anatomical regions where they cluster together representing agreement across studies; such agreement is an unlikely chance event, suggesting true pathology. Compared to VBM, where the resulting clusters represent voxel-wise significant grey matter differences, CBMA results generally relate only to the distribution of reported foci (coordinates) of the VBM clusters [8-10]. As with any meta-analysis (MA) the aim is to obtain better estimates of true effect, in this case localised changes in brain structure associated with PTSD, than provided by individual studies. Multiple analyses have been published and each indicates that, despite the inherent heterogeneity, there is observable PTSD specific pathology. Unfortunately, technical limitations of these studies render the results, and conclusions, questionable. Here a new coordinate based random effect size (CBRES) meta-analysis algorithm is used to identify which of the findings across VBM studies are consistently observed features of PTSD, while addressing the issues that limit the validity of previous analyses.

Four meta-analyses of VBM in PTSD have been published [11-14], each with technical limitations. The assumption of independence of the studies is important in MA [15] but is often violated by multiple experiments using the same subjects [16]. In CBMA correlated studies can bias the results by imposing apparent concordance; all four published meta-analyses included multiple studies as independent, but which shared subjects. One analysis [11] included a study of a-priori region of interest [17], rather than whole brain analysis, in violation of the CBMA methodological assumptions. Known software implementation issues [18] affected two of the analyses [11, 12]. Finally two studies [13, 14] employed the effect size signed differential mapping (ES-SDM) algorithm [19]. This method has a slightly different aim to others, attempting to estimate the voxel-wise grey matter difference from just the reported foci summaries. To achieve this, however, it employs no principled control of the type 1 error rate [20], making it difficult to interpret as a meta-analysis, which by definition requires robust statistics. It has been shown that if the included studies report false positive results, these will be echoed in the results [8]. Indeed the authors of the method recommend analysing only high quality studies [21], however multiple of the included VBM studies used liberal uncorrected p-value thresholding that likely resulted in false positive findings [20]. Consequently the results of these analyses are difficult to interpret.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A statistically robust re-evaluation of the evidence for generalisable structural alterations from WB studies of PTSD is warranted. The aim of this study is to use CBRES meta-analysis to explore consistent abnormal localised GM volumes or concentrations. Particular emphasis is placed on the analysis of independent WB data; where studies use common subjects, coordinates are pooled as suggested in [10, 22] to form independent sets. Differences due to experimental factors, such as using corrected/uncorrected statistical threshold or comparing PTSD subjects to healthy controls (HC) or trauma exposed controls (TEC), that may affect the MA are considered using a sensitive omnibus test of differences [23]. Furthermore, an algorithm (ClusterZ) employing a principled [20], and interpretable, threshold approach to control false positive results is used. Finally, CBMA studies rarely provide the data used for analysis making them difficult to verify or reproduce. Therefore, in line with recommendations for scientific data analysis reporting [24], the coordinates used in this study are provided to allow further evaluation/validation.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Ethical approval

No ethical approval was required for this study.

Study search and inclusion criteria

A literature search was performed using the search terms: [("post-traumatic stress disorder" OR "PTSD") AND ("voxel based morphometry" OR "grey matter volume" OR "gray matter volume")]. Studies included in previously published CBMAs of PTSD were also considered for inclusion.

Inclusion criteria were: studies involving adult subjects with a current PTSD diagnosis, the comparison of grey matter volume or concentration in PTSD to either HC or TEC, reported coordinates in either Talairach [25] or MNI [26] space. Studies that only performed ROI analyses violate the assumptions of the CBMA and were excluded; for studies performing both whole-brain and ROI analyses, only the whole-brain results were included.

Effect of experimental factors

Meta-analysis combines results from experiments that, in principle, measure the same effect in order to improve estimates. In doing so it is necessary to check for heterogeneity among the included studies that might arise due to experimental factors that do not relate to the effect. In conventional MA this is possible by computation of an inconsistency measure [27]. Such methods are not available in CBMA, but a test of whether variance in the spatial distribution of reported coordinates is partly explained by experimental factors is possible (see for example [28]) using the omnibus test. The algorithm is detailed in [23], but briefly described here.

The algorithm tests the combined p-values (sum, over coordinates, of the log p-values) using a method analogous to Fisher's combined probability test; but without the assumption of independent, uniformly distributed, p-values. Dividing the experiments into two groups by an experimental factor allows a p-value to be computed for each coordinate using a permutation test of difference between groups [23]. The resulting test statistic is the combined p-values. Under the null hypothesis there is no difference in the spatial distribution of coordinates

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between the groups, so a similarly distributed combined p-value would be obtained for a random permutation of the grouping variable. To perform the omnibus test the null distribution of combined p-values is generated using many random permutations of the grouping variable. An extreme value of the test statistic, compared to this null distribution, provides evidence of a significant difference in coordinate pattern due to the experimental factor. Such differences indicate sources of heterogeneity mean the CBMA would need to be interpreted with caution.

Coordinate based meta-analysis

The aim of CBMA is to identify statistically where the reported coordinates (foci of GM loss) from independent studies testing a common hypothesis are in spatial concordance. The results are clusters representing the spatial distribution of reported foci located where there is consensus across studies that there is a significant effect (for example grey matter loss or bold response). The anatomical locations of the clusters are then assumed to indicate pathology relevant to the common hypothesis. Statistical significance is determined using a null hypothesis that models lack of spatial concordance between studies by replacing the reported coordinates by coordinates drawn at random from grey matter image mask. It is then required that the observed coordinates agree spatially significantly better than random coordinates.

This study employs a new algorithm [8], ClusterZ, which also considers the Z scores or t-statistics commonly reported with each coordinate. Each reported Z score is first transformed into an effect size and variance that is independent of the number of subjects in the study. Then in each cluster formed by the spatial concordance of coordinates a conventional random effect size meta-analysis, or meta-regression, is performed. Interpretation of each cluster is as for conventional MA, and the use of standard visualisation and data-checking tools such as forest plots is possible. ClusterZ has multiple advantages over other algorithms. In particular by utilising the reported Z scores the p-values are based on an effect size rather than a density of reported coordinates, which might depend on the size of the structure affected. Furthermore, ClusterZ has an easy to interpret approach to principled [20] type 1 error rate control through the use of the false cluster discovery rate (FCDR). This is the proportion of the clusters declared significant that would be expected just by chance. Being based on the false discovery rate (FDR) method [29] this is less conservative than the cluster-wise family wise error (FWE)

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

method recommended with the ALE algorithm, and avoids the issue of employing voxel-wise FDR that can lead to false positive results [30].

Experimental procedure

Data extraction

Studies identified were scrutinised for: coordinate space used (MNI or Talairach), number of subjects, subject characteristics, and whether corrected statistical analysis was employed. Coordinates and Z scores from each study were extracted, and those in MNI space converted to Talairach [31]. Talairach labels were determined as detailed in [32]. Both GM volume increase and decrease were extracted; the convention used is that PTSD patients having increased GM volume relative to controls have $Z>0$, while relative GM loss is indicated by $Z<0$. Extracted data are checked by two authors (CRT and RT). In addition demographic and relevant clinical variables were also extracted and tabulated.

Analysis

Studies entered into the CBMA must be independent to avoid bias. Coordinates from multiple experiments in single studies using the same subjects were therefore pooled into single independent studies. To look for shared subjects across different studies the papers were scrutinised for similar cohort demographics. Furthermore, ClusterZ automatically indicates where different studies report occurrences of identical coordinates, which is unlikely for independent experiments, as a pointer to potentially correlated reports. Different studies sharing subjects were pooled to form a single independent study. Where there was doubt, the authors were contacted.

Difference in spatial distribution of coordinates due to experimental factors was tested for using the omnibus test. Factors considered were: use of corrected or uncorrected p-values, control group (TEC or HC), and treated or untreated PTSD subjects. A p-value of <0.05 was considered significant, and indicative of a source of heterogeneity.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Coordinate based random effect size meta-analysis was performed, using ClusterZ, on studies comparing PTSD subjects to trauma exposed controls or healthy controls both collectively and by control subject type. An FCDR of 0·05 was used, meaning only 5% of clusters declared significant were expected by chance. Post hoc analysis was performed to detect the next most significant clusters (FCDR>0.05) to make sure none were just missed; FCDR estimates are provided for these clusters.

Results

Search results

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) [33] flow chart is given for both VBM studies in figure 1.

Fig 1. PRISMA flowchart depicting the study selection process for systematic reviews.

VBM studies

Twenty one whole-brain VBM studies, reporting 24 relevant experiments involving a total of 168 coordinates, were suitable for inclusion; see table 1. These VBM studies overlapped substantially with those included in previous CBMAs, with some recent studies added [34, 35] and others omitted: Bryant et al [17], Eckhart et al [36], and Cortese et al [37] because they are ROI studies, Felmingham et al [38] because coordinates were only given for a-priori hypothesised regions, Hakamata et al [39] because the coordinates given fell outside of both Talairach and MNI space and no other coordinate space was specified, and Jatzko et al [40] because whole brain VBM results were not reported.

As noted by Li and colleagues [12], some of the studies used the same subjects. The survivors of a fire in the Hunan province of China were used in three of the studies [41-43], and two studies [44, 45] shared the same victim of rape subjects; the coordinates from these studies were merged such as to create two independent studies. The effect of pooling on significant clusters was also considered when subsets of subjects were potentially shared; the clusters should be robust against pooling of the data to ensure they are not driven by replicated results from the same subjects. The two studies by Nardo et al [46, 47] had some subjects in common

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(confirmed by author), and two studies potentially a subset of coal mine flood survivors [48, 49]. Taking into account these duplications: a total of (approximately, without complete knowledge of duplicated subject use) 316 PTSD subjects and 382 matched controls were studied.

Effects of experimental factors

A comparison between VBM studies employing corrected and uncorrected p-values using the omnibus test detected no difference in the spatial distribution of coordinates (p=0·7). Similarly comparing VBM studies using healthy controls to VBM studies using trauma exposed controls also detected no difference in the spatial distribution of coordinates (p=0·6). Of those VBM studies that reported treatment status of the patients only four reported using a treated cohort, therefore no comparison was performed using the omnibus test.

From these comparisons there was no detectable evidence that the spatial distribution of coordinates was modified by the main experimental factors. The experiments were therefore considered homogeneous for the purpose of analysis.

Meta-Analysis of VBM studies

On performing CBRES meta-analysis of all studies listed in table 1, no significant clusters were detected. The most significant cluster was not detectable below an FCDR of 0·34.

Results from the present analysis are contrary to previous CBMAs, which all reported multiple significant clusters. Therefore, independent validation was sought by performing the CBMA with GingerALE (version 2.3.6) using the recommended cluster based method method [50]. No significant clusters were detected in this analysis.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

PTSD is a condition that is inhomogeneous in terms of comorbidities, and traumatic experience is certainly individual. It is perhaps of no surprise that the results of VBM studies are inconsistent. Despite this multiple coordinate based meta-analyses have suggested generalizable pathology specific to PTSD. Such findings are important as they can inform models of PTSD. Given the technical limitations of previously published CBMAs of PTSD, and that none provided data for validation, a re-evaluation of the neuroimaging evidence for consistently detectable neuropathology was warranted. Meta-analysis of the whole-brain VBM studies revealed no significant clusters, indicating a lack of detectable agreement across VBM studies of PTSD.

The results presented here contradict those from previously published CBMAs, despite the similarity of included studies. Explanations include software implementation issues, and methods employed to correct for multiple comparison. Other issues with previous CBMAs include: use of ROI studies that can bias the results as they are generally designed around neurocircuitry models of PTSD, and treating studies as independent when in fact they use the same subjects. However, it is not possible to know exactly what has caused this apparent contradiction without access to the coordinate database used; the coordinate databases used here are provided as supplementary material. To independently check that the results were reproducible, the same coordinates were analysed using the GingerALE software (version 2.3.6) and no clusters were detected.

Ideally the results of neuroimaging studies on PTSD would be generalizable. However, it is the nature of traumatic events that they often involve few people, so many studies use small (same traumatic event) sets of subjects. Lack of statistical power might then mean that real effects are simply missed; false negatives. To counter this liberal, uncorrected, statistical thresholds are used by many (see table 1), but this risks false positive results; indeed, one author noted that application of type 1 error control by correction for multiple comparisons completely eliminated all significant results [52]. The heterogeneity (different traumatic experience) of subjects between studies is another candidate explanation, since the specifics of the trauma are relevant to PTSD [53]. Heterogeneity in terms of experimental factors, such as comparison to a healthy or trauma exposed control group, might contribute to the overall

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lack of detected consistency, but this was not detectable using the omnibus test. Elapsed time since trauma (duration of disease) is another factor that is heterogeneous between studies. Administration of rating scales to assess affective symptoms in the included studies revealed that anxiety and depressive symptoms are common, and some studies included patients fulfilling the diagnostic criteria for a co-morbid depressive disorder. This reduces the diagnostic purity of the cohort and introduces a confounding factor. Past or present pharmacological treatment may also have an effect on imaging findings. Wherever possible, a sensitive omnibus test was employed to explore whether differences in experimental design had resulted in heterogeneity, but none were found to explain a significant amount of the variance in reported coordinates.

Summary and Conclusions

The present coordinate based random effect size meta-analysis of voxel-based morphometry studies produced no evidence of consistent neuroimaging detectable structural abnormality in PTSD. The lack of results compared to previous coordinate based meta-analyses of PTSD is attributed to: software issues in previous studies, careful consideration for the independence of the studies included in the analysis, and explicit exclusion of results from a-priori region of interest analyses. These results suggest that either neuroimaging experiments of PTSD have a level of reproducibility that is not detectable by CBMA, or that pathology in PTSD is heterogeneous.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Ashburner, J. and K.J. Friston, *Voxel-based morphometry--the methods.* Neuroimage, 2000. **11**(6 Pt 1): p. 805-21.
2. Rauch, S.L., L.M. Shin, and E.A. Phelps, *Neurocircuitry models of posttraumatic stress disorder and extinction: human neuroimaging research--past, present, and future.* Biol Psychiatry, 2006. **60**(4): p. 376-82.
3. Hull, A.M., *Neuroimaging findings in post-traumatic stress disorder.* The British Journal of Psychiatry, 2002. **181**(2): p. 102-110.
4. Karl, A., et al., *A meta-analysis of structural brain abnormalities in PTSD.* Neurosci Biobehav Rev, 2006. **30**(7): p. 1004-31.
5. Woon, F.L. and D.W. Hedges, *Hippocampal and amygdala volumes in children and adults with childhood maltreatment-related posttraumatic stress disorder: a meta-analysis.* Hippocampus, 2008. **18**(8): p. 729-36.
6. Shin, L.M., et al., *Regional cerebral blood flow in the amygdala and medial prefrontal cortex during traumatic imagery in male and female Vietnam veterans with PTSD.* Arch Gen Psychiatry, 2004. **61**(2): p. 168-76.
7. Brady, K.T., et al., *Comorbidity of psychiatric disorders and posttraumatic stress disorder.* J Clin Psychiatry, 2000. **61 Suppl 7**: p. 22-32.
8. Tench, C.R., et al., *Coordinate based random effect size meta-analysis of neuroimaging studies.* Neuroimage, 2017. **153**: p. 293-306.
9. Turkeltaub, P.E., et al., *Meta-analysis of the functional neuroanatomy of single-word reading: method and validation.* Neuroimage, 2002. **16**(3 Pt 1): p. 765-80.
10. Wager, T.D., M. Lindquist, and L. Kaplan, *Meta-analysis of functional neuroimaging data: current and future directions.* Soc Cogn Affect Neurosci, 2007. **2**(2): p. 150-8.
11. Kuhn, S. and J. Gallinat, *Gray matter correlates of posttraumatic stress disorder: a quantitative meta-analysis.* Biol Psychiatry, 2013. **73**(1): p. 70-4.
12. Li, L., et al., *Grey matter reduction associated with posttraumatic stress disorder and traumatic stress.* Neurosci Biobehav Rev, 2014. **43**: p. 163-72.
13. Meng, L., et al., *Trauma-specific Grey Matter Alterations in PTSD.* Sci Rep, 2016. **6**: p. 33748.
14. Meng, Y., et al., *Anatomical deficits in adult posttraumatic stress disorder: a meta-analysis of voxel-based morphometry studies.* Behav Brain Res, 2014. **270**: p. 307-15.
15. Haidich, A.B., *Meta-analysis in medical research.* Hippokratia, 2010. **14**(Suppl 1): p. 29-37.
16. Tramer, M.R., et al., *Impact of covert duplicate publication on meta-analysis: a case study.* BMJ, 1997. **315**(7109): p. 635-40.
17. Bryant, R.A., et al., *Rostral anterior cingulate volume predicts treatment response to cognitive-behavioural therapy for posttraumatic stress disorder.* J Psychiatry Neurosci, 2008. **33**(2): p. 142-6.
18. Eickhoff, S.B., et al., *Implementation errors in the GingerALE Software: Description and recommendations.* Hum Brain Mapp, 2016.
19. Radua, J., et al., *A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps.* Eur Psychiatry, 2012. **27**(8): p. 605-11.
20. Bennett, C.M., G.L. Wolford, and M.B. Miller, *The principled control of false positives in neuroimaging.* Soc Cogn Affect Neurosci, 2009. **4**(4): p. 417-22.
21. Radua, J., et al., *Meta-analytical comparison of voxel-based morphometry studies in obsessive-compulsive disorder vs other anxiety disorders.* Arch Gen Psychiatry, 2010. **67**(7): p. 701-11.
22. Turkeltaub, P.E., et al., *Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses.* Hum Brain Mapp, 2012. **33**(1): p. 1-13.
23. Tench, C.R., et al., *Coordinate based meta-analysis of functional neuroimaging data using activation likelihood estimation; full width half max and group comparisons.* PLoS One, 2014. **9**(9): p. e106735.
24. Diggle, P.J., *Statistics: a data science for the 21st century.* Journal of the Royal Statistical Society: Series A (Statistics in Society), 2015. **178**(4): p. 793--813.
25. Talairach, J. and P. Tournoux, *Co-planar stereotaxic atlas of the human brain*. 1988, New York: Thieme.
26. Evans, A.C., et al. *3D statistical neuroanatomical models from 305 MRI volumes*. in *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.* 1993.
27. Higgins, J.P.T., et al., *Measuring inconsistency in meta-analyses.* BMJ, 2003. **327**(7414): p. 557-560.
28. Tanasescu, R., et al., *Functional reorganisation in chronic pain and neural correlates of pain sensitisation: a coordinate based meta-analysis of 266 cutaneous pain fMRI studies.* Neurosci Biobehav Rev, 2016.
29. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society. Series B (Methodological), 1995. **57**(1): p. 289-300.
30. Eickhoff, S.B., et al., *Behavior, Sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation.* NeuroImage, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

31. Lancaster, J.L., et al., *Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template.* Hum Brain Mapp, 2007. **28**(11): p. 1194-205.
32. Lancaster, J.L., et al., *Automated Talairach atlas labels for functional brain mapping.* Hum Brain Mapp, 2000. **10**(3): p. 120-31.
33. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
34. Bossini, L., et al., *Morphovolumetric changes after EMDR treatment in drug-naive PTSD patients.* Riv Psichiatr, 2017. **52**(1): p. 24-31.
35. O'Doherty, D.C.M., et al., *Frontal and subcortical grey matter reductions in PTSD.* Psychiatry Res, 2017. **266**: p. 1-9.
36. Eckart, C., et al., *Structural alterations in lateral prefrontal, parietal and posterior midline regions of men with chronic posttraumatic stress disorder.* J Psychiatry Neurosci, 2011. **36**(3): p. 176-86.
37. Cortese, B.M., et al., *Burning odor-elicited anxiety in OEF/OIF combat veterans: Inverse relationship to gray matter volume in olfactory cortex.* J Psychiatr Res, 2015. **70**: p. 58-66.
38. Felmingham, K., et al., *Duration of posttraumatic stress disorder predicts hippocampal grey matter loss.* Neuroreport, 2009. **20**(16): p. 1402-6.
39. Hakamata, Y., et al., *Structure of orbitofrontal cortex and its longitudinal course in cancer-related post-traumatic stress disorder.* Neurosci Res, 2007. **59**(4): p. 383-9.
40. Jatzko, A., et al., *Hippocampal volume in chronic posttraumatic stress disorder (PTSD): MRI study using two different evaluation methods.* J Affect Disord, 2006. **94**(1-3): p. 121-6.
41. Chen, S., et al., *Insular cortex involvement in declarative memory deficits in patients with post-traumatic stress disorder.* BMC Psychiatry, 2009. **9**: p. 39.
42. Chen, S., et al., *Gray matter density reduction in the insula in fire survivors with posttraumatic stress disorder: a voxel-based morphometric study.* Psychiatry Res, 2006. **146**(1): p. 65-72.
43. Li, L., et al., *Magnetic resonance imaging and magnetic resonance spectroscopy study of deficits in hippocampal structure in fire victims with recent-onset posttraumatic stress disorder.* Can J Psychiatry, 2006. **51**(7): p. 431-7.
44. Sui, S.G., et al., *Abnormal grey matter in victims of rape with PTSD in Mainland China: a voxel-based morphometry study.* Acta Neuropsychiatr, 2010. **22**(3): p. 118-26.
45. Sui, S.-G., et al., *Abnormal cerebellum density in victims of rape with post-traumatic stress disorder: Voxel-based analysis of magnetic resonance imaging investigation.* Asia-Pacific Psychiatry, 2010. **2**(3): p. 129-135.
46. Nardo, D., et al., *Gray matter volume alterations related to trait dissociation in PTSD and traumatized controls.* Acta Psychiatr Scand, 2013. **128**(3): p. 222-33.
47. Nardo, D., et al., *Gray matter density in limbic and paralimbic cortices is associated with trauma load and EMDR outcome in PTSD patients.* Journal of Psychiatric Research, 2010. **44**(7): p. 477-485.
48. Chen, Y., et al., *Different regional gray matter loss in recent onset PTSD and non PTSD after a single prolonged trauma exposure.* PLoS One, 2012. **7**(11): p. e48298.
49. Zhang, J., et al., *Decreased gray matter volume in the left hippocampus and bilateral calcarine cortex in coal mine flood disaster survivors with recent onset PTSD.* Psychiatry Res, 2011. **192**(2): p. 84-90.
50. Eickhoff, S.B., et al., *Activation likelihood estimation meta-analysis revisited.* Neuroimage, 2012. **59**(3): p. 2349-61.
51. Gilbertson, M.W., et al., *Smaller hippocampal volume predicts pathologic vulnerability to psychological trauma.* Nat Neurosci, 2002. **5**(11): p. 1242-7.
52. Kasai, K., et al., *Evidence for acquired pregenual anterior cingulate gray matter loss from a twin study of combat-related posttraumatic stress disorder.* Biol Psychiatry, 2008. **63**(6): p. 550-6.
53. Breslau, N., et al., *Trauma and posttraumatic stress disorder in the community: the 1996 Detroit Area Survey of Trauma.* Arch Gen Psychiatry, 1998. **55**(7): p. 626-32.
54. Chao, L.L., M. Lenoci, and T.C. Neylan, *Effects of post-traumatic stress disorder on occipital lobe function and structure.* Neuroreport, 2012. **23**(7): p. 412-9.
55. Cheng, B., et al., *Gray Matter Alterations in Post-Traumatic Stress Disorder, Obsessive-Compulsive Disorder, and Social Anxiety Disorder.* Front Behav Neurosci, 2015. **9**: p. 219.
56. Corbo, V., et al., *Size versus shape differences: contrasting voxel-based and volumetric analyses of the anterior cingulate cortex in individuals with acute posttraumatic stress disorder.* Biol Psychiatry, 2005. **58**(2): p. 119-24.
57. Herringa, R., et al., *Post-traumatic stress symptoms correlate with smaller subgenual cingulate, caudate, and insula volumes in unmedicated combat veterans.* Psychiatry Res, 2012. **203**(2-3): p. 139-45.
58. Rocha-Rego, V., et al., *Decreased premotor cortex volume in victims of urban violence with posttraumatic stress disorder.* PLoS One, 2012. **7**(8): p. e42560.
59. Tan, L., et al., *Brain structure in post-traumatic stress disorder: A voxel-based morphometry analysis.* Neural Regen Res, 2013. **8**(26): p. 2405-14.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

60. Tavanti, M., et al., *Evidence of diffuse damage in frontal and occipital cortex in the brain of patients with post-traumatic stress disorder.* Neurol Sci, 2012. **33**(1): p. 59-68.
61. Thomaes, K., et al., *Reduced anterior cingulate and orbitofrontal volumes in child abuse-related complex PTSD.* J Clin Psychiatry, 2010. **71**(12): p. 1636-44.
62. Yamasue, H., et al., *Voxel-based analysis of MRI reveals anterior cingulate gray-matter volume reduction in posttraumatic stress disorder due to terrorism.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 9039-43.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

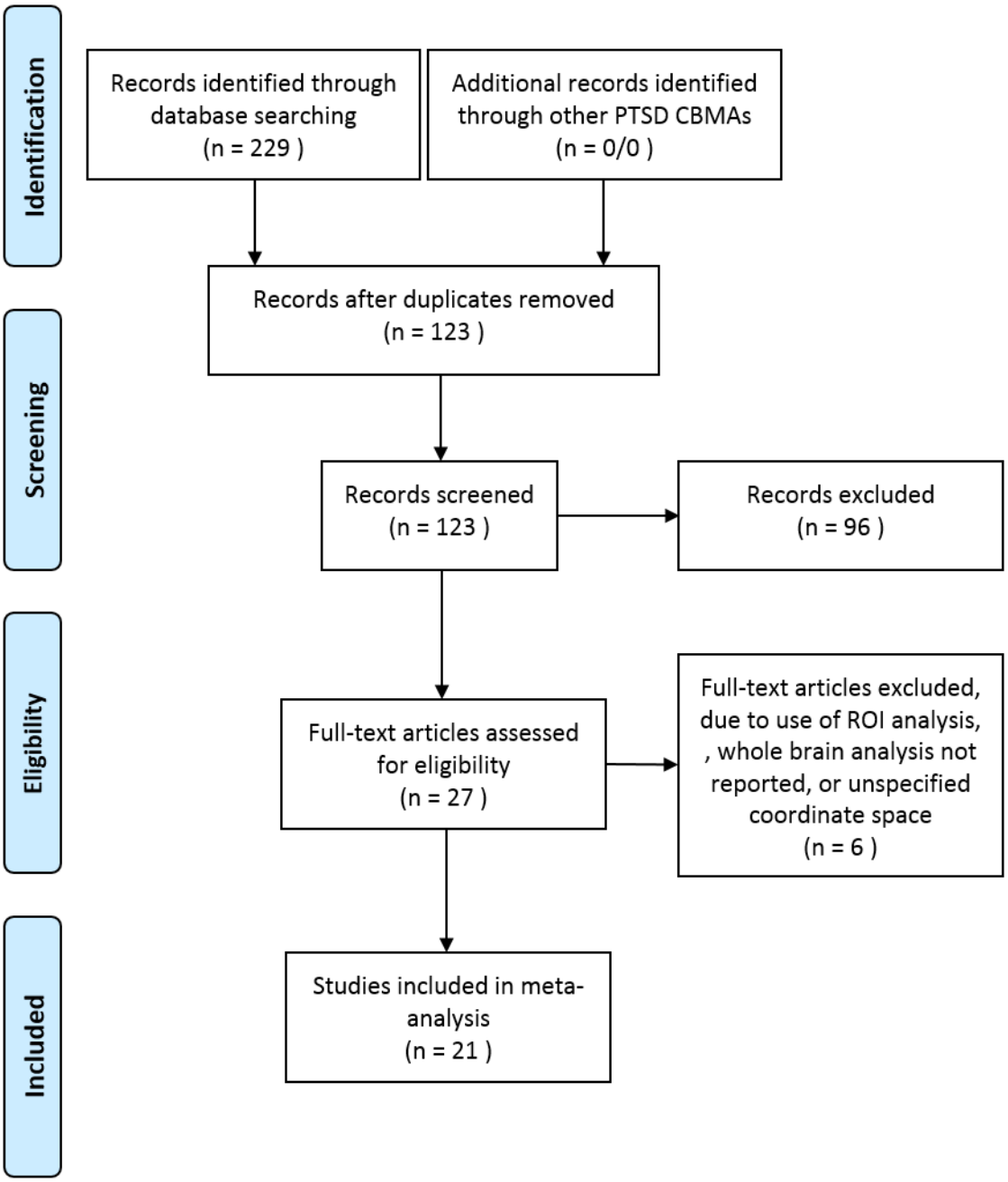

Fig 1. PRISMA flowchart depicting the study selection process for systematic reviews; numbers indicate VBM studies.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Study	Nature of trauma	Criteria used to diagnose PTSD	Time lapse between traumatic event and study	PTSD symptom severity assessment scale (score)	Psychiatric comorbidity diagnosed in addition to PTSD (n=)	Depression / anxiety symptom rating scale (score)	VBM volume/ concentration	Exposed to psychotropic medication	Mean PTSD Age (years) (sd/range)	Corrected p-values?	Coord-inates	PTSD n= (male)	Control n= Type
Bossini 2017 [34]	Mixed	DSM-IV	9 years	CAPS total (unknown)	No	-	Volume	No	40	Yes	7	19 (10)	19 HC
Chao 2012 [54]	Combat veterans	DSM-IV	Not reported	CAPS total (59.1)	No	BDI (18.2)	Volume	Yes, 6 patients taking a 'serotonergic antidepressant'	39 (11)	No	8	21 male	20 TEC
Chen 2006 [42]	Survivors of fire in Hunan province, China Nov 2003	DSM-IV	6.2 months	DEQ (43.12)	No	-	Concentration	No	35 (5)	Yes	4	12 (4)	12 TEC
Chen 2009 [41]	Survivors of fire in Hunan province, China Nov 2003	DSM-IV	Not reported	DEQ (43.12)	GAD (3)	-	Concentration	No	35 (5)	No	3	12 (4)	12 TEC
Chen 2012 [48]	Coalmine flood disaster	DSM-IV	6.2 months	CAPS total (78.7)	No	-	Volume	No	41 (7)	Yes	1 0	10 male	20 HC 10 TEC
Cheng 2015 [55]	Wenchuan Earthquake survivors	DSM-IV	12.5 months	CAPS total (56.4)	No	HAMA (10.9) HAMD (14.2)	Volume	No	26 (8)	Yes	3	30 (21)	30 HC
Corbo 2005 [56]	Motor vehicle accident victims	DSM-IV	Less than 4 weeks	CAPS PTSD score (64.71)	MDD (7) GAD (2) OCD (1)	BDI (7.21)	Concentration	No	33 (12)	Yes	5	14 (6)	14 HC
Herringa 2012 [57]	Combat veterans	DSM-IV	Not reported	CAPS total (47.5)	GAD (1)	BDI (10.5)	Volume	No	29 (4)	No	4	13 (11)	15 TEC
Kasai 2008 [52]	Combat veterans	DSM-IV	Not reported	CAPS total (73.3)	No	SCL-90 (2.5)	Concentration	Not reported	53 (3)	No	7	18 (unkno wn)	23 TEC
Li 2006 [43]	Survivors of fire in Hunan province, China Nov 2003	DSM-IV	6 – 8 months	DEQ (43.12)	Yes (2)	-	Concentration	No	35 (5)	Yes	4	12 (4)	12 TEC
Nardo 2010 [47]	Stockholm public transport accidents	DSM-IV	3 months – 6 years	-	No	-	Concentration	No	42 (9)	No	5	21 (15)	22 TEC
Nardo 2013 [46]	Stockholm public transport accidents	DSM-IV	3 months – 6 years	-	No	HAMD (30.0) HAMA (17.27)	Volume	No	42.41 (8.63)	No	5	15 (12)	17 TEC
O'Doherty 2017 [35]	Not specified	DSM-IV	3 months – 10 years	CAPS total (95.6)	No	-	Volume	No	34.0	Yes	25 26	25 (12)	25 TEC 25 HC
Rocha-Rego 2012[58]	Victims of urban violence	DSM-IV	3.0 years	PCLC (not reported)	Yes (16)	-	Volume	Yes (treatment not specified)	43 (6)	No	4	16 (7)	16 TEC
Sui 2010[45]	Victims of rape	DSM-IV	3.79 years	CAPS total (74.45)	No	-	Concentration	Yes (treatment not specified)	25.55 (4.01)	No	6 11	11 female; 13 recruite dbut 2 removed due to scan artifact	8 TEC 12 HC
Sui 2010a [44]	Victims of rape	DSM-IV	4.5 years; for 13 subjects recruited	CAPS (not reported)	No	-	Concentration	Not specified	24 (5.77); for 13 subjects recruited	No	6	11 female; 2 removed due to scan artifact	12 HC
Tan 2013[59]	Survivors of fire in Hunan province, China June 2005	DSM-IV-R	2 – 3 years	CAPS total (58.1)	No	-	Volume	No	37.6 (3.7)	Yes	2	12 male	14 TEC
Tavanti 2011[60]	Unspecifie d trauma in adults	DSM-IV	7.36 years	CAPS total (81.5)	No	HAMD (15.32) HAMA (25.8)	Volume	No	38 (11)	Yes	26	25 (8)	25 HC
Thomaes 2010[61]	Victims of historical child abuse (physical and/or sexual abuse before the age of 16 years)	DSM-IV-R	Not specified	CAPS total (88.5)	MDD 21 Anxiety disorder 23	BDI (29.3)	Concentration	Yes; SSRI, mean 35.7mg FLX-eqv	35 (10)	No	6	31 female	28 HC
Yamasue 2003[62]	Tokyo subway attack in 1995	Not stated	5 – 6 years	CAPS total (62.2)	Yes (1)	-	Concentration	No	45 (16)	Yes	1	9 (5)	16 TEC
Zhang 2011[49]	Coalmine flood disaster in 2007	DSM-IV	6 months	CAPS total (78.7)	No	-	Volume	No	40.8 (6.83)	Yes	3	10 male	10 TEC

Table 1. VBM studies included in analysis. Of these 21 studies, 24 experiments were extracted: 15 comparing PTSD to a TEC control group, and 9 comparing to a HC group. Principled methods of controlling the type 1 error rate were employed by 11 of the studies. Eleven of the studies used VBM to analyse GM volume, while ten analysed GM concentration.
DSM: Diagnostic and Statistical Manual of Mental Disorders. CAPS: Clinician-Administered PTSD Scale. DEQ: Distressing Events Questionnaire. SCL-90: Symptom Check List-90.

bioRxiv preprint doi: https://doi.org/10.1101/265496; this version posted February 15, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

HAMA: Hamilton Depression Scale. HAMD: Hamilton Anxiety Scale. BDI: Beck Depression Inventory. GAD: Generalised anxiety disorder. OCD: obsessive compulsive disorder.