Coordinate based meta-analysis does not show grey matter atrophy in narcolepsy

Tench CR[1] Christopher.Tench@Nottingham.ac.uk

Radu Tanasescu[1,2] Radu.Tanasescu@nottingham.ac.uk

Cottam WJ[3] William.cottam@nottingham.ac.uk

Constantinescu CS [1] Cris.Constantinescu@Nottingham.ac.uk

Auer DP[3] dorothee.auer@nottingham.ac.uk

[1]Division of Clinical Neurosciences, Clinical Neurology, University of Nottingham, Queen's Medical Centre, Nottingham, UK

[2]Department of Neurology, University of Medicine and Pharmacy Carol Davila Bucharest, Colentina Hospital, Bucharest, Romania

[3]Division of Clinical Neuroscience, Radiological Sciences, University of Nottingham, Queen's Medical Centre, Nottingham, UK

Background

Narcolepsy is a lifelong debilitating condition caused by hypocretin deficiency [1] with symptoms including excessive sleep disorder. In idiopathic narcolepsy there is a highly selective loss of 85–95% of hypocretin producing hypothalamic neurons [2], and this has motivated multiple whole-brain voxel based morphometry (VBM) [3] studies exploring MRI detectable structural changes to the grey matter (GM) in narcolepsy. Findings are inconsistent, with some reporting multiple GM changes compared to healthy control subjects, for example [4], and others reporting none [5]. This inconsistency is not atypical of neuroimaging studies, where small numbers of subjects and experimental variability can produce results that are not generalizable [6]. This has driven the development of coordinate based meta-analysis (CBMA), which statistically combines the significant coordinates from multiple studies to distinguish study specific effects from grey matter alteration that is consistent across studies. The aim of meta-analysis is objective appraisal of evidence from multiple independent studies [7]. There are several CBMA algorithms available, but all have a similar approach. Using the coordinates of grey matter loss reported by VBM studies, a measure of spatial consistency of reporting across studies is computed on a voxel-wise basis. To assess statistical significance these measures are compared to those obtained when the coordinates are replaced by random coordinates. Therefore, after controlling for multiple voxel-wise statistical tests, significant results are indicative of stronger agreement across studies than expected by random chance.

Weng et al [8] recently performed a CBMA of eight VBM studies of narcolepsy utilising the activation likelihood estimation (ALE) algorithm [9, 10]. This showed multiple regions of consistent GM loss including the hypothalamus. These results were subsequently demonstrated to be due to an implementation issue with the software [11], and re-analysis using a corrected version (GingerALE 2.3.3) detected no consistent significant regional GM loss [12]. This reanalyses employed the false discovery rate (FDR) method of controlling the type 1 error [13], which is no longer the recommended option in GingerALE [14] and has been superseded [15]. This has prompted a second re-evaluation of the narcolepsy data by Zhong and colleagues [16] employing the signed differential mapping (SDM) CBMA algorithm [17, 18]. The results apparently reaffirm the initial assessment that there is a consistent pattern of grey matter loss in narcolepsy. However, while the criticism of the use of FDR by Zhong and colleagues may have some validity, the SDM algorithm employed instead offers no principled control of the type 1 error rate [19]. The method therefore provides no quantitative evidence that the results are truly significant given the very large number (2×10^{5}) of voxel-wise statistical tests, which is in contradiction to the aim of meta-analysis [7]. The aim of this commentary is to clarify this situation using the latest version of the ALE algorithm [15], which avoids the criticism of FDR and offers principled control of the family wise error (FWE) rate.

Coordinate based meta-analysis of voxel based morphometry studies of narcolepsy

Details of the VBM studies included in the original meta-analysis [8] are given in table 1; the coordinates are also available to download from the previous commentary [12]. There are 8 studies involving 149 subjects with narcolepsy and 162 healthy control subjects. A total of 72 coordinates of

grey matter loss were reported by these studies. Only half of the studies used a principled method to control the type 1 error rate, while half used uncorrected p-values.

Study	Subjects with narcolepsy	Healthy controls	Reported coordinates	Corrected statistics used
[20]	29	29	4 Courtesy of authors	Yes
[4]	12	32	19	No
[5]	15	15	0	Yes
[21]	12	12	2	Yes
[22]	19	16	2	No
[23]	17	17	29	No
[24]	29	29	11	No
[25]	16	12	5	Yes

Table 1. Eight VBM studies of narcolepsy included in the CBMA.

The data were subjected to a fourth analysis using the most recent type 1 error control scheme [15] implemented into GingerALE (version 2.3.6). An uncorrected p-value threshold of $p<0.001$ was combined with a cluster-level FWE rate threshold of 0.05, and 1000 iterations were performed; these are the recommended settings.

Results

A single cluster (109 voxels) was detected indicating consistent regional GM atrophy of the hypothalamus, centred on Talairach coordinate (4·8, -0·9,-12·2) mm. Detection of this cluster invites closer inspection of the contributing studies for the purpose of validation and interpretation. The hypothalamic coordinate included in the Joo et al [24] study was not significant by whole-brain VBM but was included as a significant a-priori region of interest (ROI) result, it should therefore not have been included in the meta-analysis; ROI studies violate the assumptions of CBMA. Furthermore, the study by Buskova et al [22] is not a whole-brain VBM study, but rather an ROI study specifically of the hypothalamus in narcolepsy with cataplexy and should have also been excluded from the CBMA. Removing these coordinates eliminated the significant result detected by GingerALE.

Summary and conclusion

Coordinate based meta-analyses using the GingerALE software may not be valid if the version used pre-dates the fixes to the known implementation issues [11]. The impact can be considerable, as demonstrated by this reanalysis of the narcolepsy data using a fixed version of the software [12]. One limitation of the analysis was the use of FDR, which is no longer the recommended ALE method and has motivated a further reanalysis using SDM, which seemed to reaffirm the original findings. However, SDM has never been validated, instead an uncorrected p-value threshold was determined empirically from a single study on known positive data then assumed to be generally applicable [17]. This is not a safe assumption [19], and is the reason FDR and FWE methods are used. Consequently SDM provides no quantitative evidence that the results are critical of the null hypothesis. The presented third reanalysis overcomes all of the criticisms raised by using the cluster-level FWE control scheme with working implementation (version 2.3.6) of GingerALE [11], but found no evidence for spatially consistent effects reported by the 8 narcolepsy studies.

These analyses all suffer from the limitations of the data. Eight (seven without the ROI study) studies are intuitively too few for reasonable statistical power, but the concept of power in CBMA is complicated by its explicit dependency on the specifics of the contributing studies. Nevertheless, a recent quantitative analysis of statistical power of the ALE algorithm [14] concluded that 20 studies are a minimum for typical effects (around 25% of studies reporting the effect), and that with just 7 valid studies the estimated power is only around 20%. Indeed, with 7 studies the statistical power is such that only strong effects (most studies reporting coordinates in the same anatomical location) would be detectible by ALE, and these would be apparent by direct inspection of the studies. Furthermore, half of the studies used uncorrected p-value thresholds and some involved few subjects, which is known to greatly increase the risk of false positive results [6] that might have reduced the power to detect narcolepsy specific effects even further. Other limitations relate to the reporting of VBM studies, which can make it difficult to identify exactly what data to include. This resulted in the inclusion of a non-significant hypothalamic coordinate, which was reported [24] in a table that combined both whole-brain VBM and a-priori ROI results; ROI analyses are often reported alongside whole-brain analyses when there are a-priori regions of interest. Another study was presented as a VBM study throughout [22], but on careful reading turns out to be a region of interest study specifically of the hypothalamus. It is particularly important to avoid including ROI studies in CBMA as it biases the results and can erroneously reaffirm a-priori models of pathology. This emphasizes the hurdles to performing a robust CBMA which necessitates: sufficient numbers of studies for reasonable statistical power to detect typical effects, close scrutiny of any studies contributing to significant clusters in CBMA to verify validity, and principled control of the type 1 error rate to provide quantitative evidence that the results are critical of the null hypothesis.

In conclusion these meta-analyses provide no evidence of consistent regional grey matter atrophy specific to narcolepsy. However, given the limitations of the studies analysed, absence of significance is not suggestive that regional grey matter atrophy in narcolepsy is heterogeneous.

Given the highlighted difficulties in selecting the valid coordinates for CBMA from each study, and in line with recommendations for scientific data analysis reporting [26], the coordinates used in this study are provided to allow further evaluation/validation.

1. Liblau, R.S., et al., *Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy.* Lancet Neurol, 2015. **14**(3): p. 318-28.
2. Thannickal, T.C., et al., *Reduced number of hypocretin neurons in human narcolepsy.* Neuron, 2000. **27**(3): p. 469-74.
3. Ashburner, J. and K.J. Friston, *Voxel-based morphometry--the methods.* Neuroimage, 2000. **11**(6 Pt 1): p. 805-21.
4. Kaufmann, C., et al., *Reduced cortical gray matter in narcolepsy: preliminary findings with voxel-based morphometry.* Neurology, 2002. **58**(12): p. 1852-5.
5. Overeem, S., et al., *Voxel-based morphometry in hypocretin-deficient narcolepsy.* Sleep, 2003. **26**(1): p. 44-6.
6. Button, K.S., et al., *Power failure: why small sample size undermines the reliability of neuroscience.* Nat Rev Neurosci, 2013. **14**(5): p. 365-76.

7.   Egger, M. and G.D. Smith, *Meta-Analysis. Potentials and promise.* BMJ, 1997. **315**(7119): p. 1371-4.

8.   Weng, H.H., et al., *Gray matter atrophy in narcolepsy: An activation likelihood estimation meta-analysis.* Neurosci Biobehav Rev, 2015. **59**: p. 53-63.

9.   Laird, A.R., et al., *ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts.* Hum Brain Mapp, 2005. **25**(1): p. 155-64.

10.   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.

11.   Eickhoff, S.B., et al., *Implementation errors in the GingerALE Software: Description and recommendations.* Hum Brain Mapp, 2016.

12.   Tanasescu, R., et al., *Coordinate based meta-analysis does not show grey matter atrophy in narcolepsy.* Neurosci Biobehav Rev, 2015. **57**: p. 297-8.

13.   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.

14.   Eickhoff, S.B., et al., *Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation.* Neuroimage, 2016. **137**: p. 70-85.

15.   Eickhoff, S.B., et al., *Activation likelihood estimation meta-analysis revisited.* Neuroimage, 2012. **59**(3): p. 2349-61.

16.   Zhong, J., et al., *Voxelwise meta-analysis of grey matter atrophy in narcolepsy.* Neurosci Biobehav Rev, 2016.

17.   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.

18.   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.

19.   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.

20.   Draganski, B., et al., *Hypothalamic gray matter changes in narcoleptic patients.* Nat Med, 2002. **8**(11): p. 1186-8.

21.   Brenneis, C., et al., *Voxel-based morphometry in narcolepsy.* Sleep Med, 2005. **6**(6): p. 531-6.

22.   Buskova, J., et al., *Reduced hypothalamic gray matter in narcolepsy with cataplexy.* Neuro Endocrinol Lett, 2006. **27**(6): p. 769-72.

23.   Kim, S.J., et al., *Gray matter deficits in young adults with narcolepsy.* Acta Neurol Scand, 2009. **119**(1): p. 61-7.

24.   Joo, E.Y., et al., *Gray matter concentration abnormality in brains of narcolepsy patients.* Korean J Radiol, 2009. **10**(6): p. 552-8.

25.   Scherfler, C., et al., *White and gray matter abnormalities in narcolepsy with cataplexy.* Sleep, 2012. **35**(3): p. 345-51.

26.   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.