Coordinate based meta-analysis of networks in neuroimaging studies

Tench CR[1,2]*
Christopher.Tench@Nottingham.ac.uk

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

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

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

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

Auer DP[2-5]
dorothee.auer@nottingham.ac.uk

Cottam WJ[2-5]
William.cottam@nottingham.ac.uk

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

[2]NIHR Nottingham Biomedical Research Centre, Queen's Medical Centre, University of Nottingham, Nottingham, UK.

[3]Arthritis Research UK Pain Centre, University of Nottingham, Nottingham, UK.

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

[5]Sir Peter Mansfield Imaging Centre, School of Medicine, University of Nottingham, Nottingham, UK.

*Corresponding author

Keywords
Meta-analysis, neuroimaging, voxel-based morphometry, functional MRI, networks, graphs

Highlights
- Meta-analyses coordinates from published studies as coactivation networks
- Uses the same data as popular coordinate based meta-analysis
- Freely available software

Abstract

Meta-analysis of summary results from published neuroimaging studies independently testing a common hypothesis is performed using coordinate based meta-analysis (CBMA), which tests for consistent activation (in the case of functional MRI studies) of the same anatomical regions. Using just the reported coordinates it is also possible to meta-analyse coactivated regions to reveal a network-like structure of coordinate clusters (network nodes) distributed at the coactivated locations and a measure of the coactivation strength (network edges), which is determined by the presence/absence of reported activation.

Here a new coordinate-based method to estimate a network of coactivations is detailed, which utilises the Z score accompanying each reported. Coordinate based meta-analysis of networks (CBMAN) assumes that if the activation pattern reported by independent studies is truly consistent, then the relative magnitude of these Z scores might also be consistent. It is hypothesised that this is detectable as Z score covariance between coactivated regions provided the within study variances are small. Advantages of using the Z scores instead of coordinates to measure coactivation strength are that censoring by the significance thresholds can be considered, and that using a continuous measure rather than a dichotomous one can increase statistical power.

CBMAN uses maximum likelihood estimation to fit multivariate normal distributions to the standardised Z scores, and the covariances are considered as edges of a network of coactivated clusters (nodes). Here it is validated by numerical simulation and demonstrated on real data used previously to demonstrate CBMA. Software to perform CBMAN is freely available.

Introduction

Coordinate based meta-analysis (CBMA) is an approach commonly used to estimate the consistently observable effects from multiple independent, but related by a shared hypothesis, neuroimaging studies [1-10]. It is employed to meta-analyse (amongst others) voxel-based morphometry (VBM), functional magnetic resonance imaging (fMRI), and functional positron emission tomography (PET) studies, and uses only the reported summary statistics; coordinates and/or Z scores. In the case of fMRI, for example, the results reveal estimates of the distribution of activation peaks (clusters of activation foci) [1, 4, 11]. CBMA's generally report this distribution as multiple spatially isolated clusters of coordinates and it is their anatomical locations on which the interpretation and conclusion are based.

With coordinate based meta-analysis (MA) results are largely determined by the consistency with which clusters are reported as activated (in the case of fMRI studies) by the independent studies. There is, however, considerable interest in approaches that relate the spatially isolated clusters. Meta analytic connectivity modelling (MACM) [12] uses coordinates from large databases such as brainmap (http://www.brainmap.org/) to identify regions frequently coactivated (in the case of fMRI, co-altered in the case of VBM) across multiple domains. The ability to identify network-like features using a coordinate based meta-analysis has also been proposed [13-19]. These methods use voxel-wise analysis, predefined regions of interest, or significant clusters from the popular activation likelihood estimate (ALE) CBMA algorithm [1, 3, 5, 6] to the define network nodes, and coactivation strength (network edges) is based only on reported coordinates. Results indicate anatomical regions frequently coactivated (activated together within the same study), which is distinct from CBMA where only the frequency is important.

It might be argued that if the hypothesis generates consistently observable spatial effects then the reported Z scores accompanying each coordinate might also be consistent except for study heterogeneity. For example, the relative magnitude of activation between the anatomical regions should be such that studies reporting the smallest (largest) activation strength in one region also report the smallest (largest) in a coactivated region; for positively correlated activation strength. Consequently, if the within-study sample variance is small the Z scores reported by multiple independent studies may be correlated between coactivated clusters. Indeed, this has previously been shown using clusters derived from a CBMA [10], motivating development of the new analysis method proposed here.

Coordinate based meta-analysis of networks (CBMAN) analyses coactivation of clusters by estimating the covariance of standardised reported Z scores. A multivariate normal (MVN) distribution model is fitted and is parameterised such that there are means and variances for each cluster, plus correlations relating the standardised Z scores in coactivated clusters. It is these correlations that form the edges in the network, while the coactivated clusters are the nodes. To test for significant correlation between pairwise clusters a permutation

method is used, which after correction for multiple statistical tests reveals a network of significant coactivation. By comparison with approaches using just coordinates, the use of standardised reported Z scores provides the opportunity to consider censoring by the study thresholds and may be more sensitive in some cases because dichotomising (coordinate present/absent within a cluster) can reduce statistical power [20]. However, interpretation is different in that coordinates do not explicitly imply correlated Z scores. It should be noted that despite the network-like nature of coactivation methods, they are indicative only of consistent coactivation and cannot be interpreted in the same way as the possibly more commonly considered brain networks defined by temporal dependency of neuronal activation patterns of anatomically separated brain regions [21]. Nevertheless, coordinate based coactivation methods provide an alternative to CBMA, which can be subjected to network analysis [14, 16].

This article describes coordinate based meta-analysis of networks. The methods involved in fitting MVN distributions are detailed, including the subtleties involved with censored data. Ability to estimate correlation with censored data is demonstrated by numerical experiments. Simulated networks are used to demonstrate the algorithm and real coordinate data from painful stimulus fMRI studies, and VBM studies of multiple sclerosis, employed to show feasibility. Type 1 error rate control is by false discovery rate (FDR) [22]. The software to perform CBMAN is provided to use freely (search NeuRoi), and test experiments provided for validation.

Methods

The cluster forming algorithm

There is no strict definition of a cluster in coordinate-based analysis, with each algorithm having different cluster forming approaches. CBMAN finds peaks in estimated study density and assigns the reported coordinates to them before further processing to obtain valid clusters. A minimum number of coordinates defines cluster validity, and the choice of minimum acts as a constraint on the assignment of coordinates to clusters. For example, a collection of three coordinates might be assigned to a bigger cluster/clusters if the minimum was greater than 3, whereas they might form their own cluster with a minimum of 3 or less. In common with the popular density-based spatial clustering of applications with noise (DBSCAN) algorithm CBMAN requires clusters have at least four coordinates because that is the minimum number needed to define a three-dimensional cluster volume. Sensitivity to this choice is considered in the results.

Cluster forming requires the spatial density of reporting studies, which is estimated on a voxel-by-voxel basis by summation of Gaussian kernels over independent studies

$$D(r) = \sum_{studies\ i} exp\left(-\frac{(r-r_i)^2}{2\sigma^2}\right) \qquad (1)$$

where r_i is the coordinate from study i that is nearest to the voxel centred on r. The parameter σ must be large enough to smooth away false density peaks while being small enough to preserve true density peaks; a density peak is a voxel where the study density estimate is higher than that in neighbour voxels. Just as with fMRI processing the choice of smoothing kernel can influence the results, for example in network analysis [23], and a suitable choice is not always apparent a-priori. Fortunately, the spatial properties of clusters formed by activation coordinates reported by independent studies have been investigated and found to be consistent [9]. Specifically the spatial spread of coordinates has a standard deviation between 3mm and 5mm with a large peak at around 4mm, which is modelled here with a truncated normal $N(4,0.4^2)$ distribution.

Using this empirical spatial spread, choice of σ can be made in a principled way by simulating clusters with coordinates having standard spatial deviation drawn randomly (per simulation) from the truncated $N(4,0.4^2)$. These coordinates are smoothed to produce a density image and the number of detectible peaks counted. The value chosen for σ is the minimum producing no more than one peak in the large majority of simulations. Avoiding false density peaks is important since they can cause clusters to split into multiple clusters. It is anticipated that larger numbers of coordinates require less smoothing, so σ is estimated for different numbers of coordinates to allow an empirical relationship to be established for extrapolation. The analysis of the parameter σ is presented in the results.

Clustering begins by computing an image of study density from which peaks (proposed cluster centres) are detected (see figure 1(a) for example). Coordinates (cluster members)

are then assigned to the cluster centres such that they are only assigned to the nearest, and only the closest from each study is assigned. This produces clusters with at most one contributed coordinate per study. However, there is nothing to prevent assignment of coordinates that are unrealistically far from the bulk of the cluster (outlier coordinates), as indicated by the study density dropping with distance from the peak before increasing again at the outliers (see figure 1b). To eliminate outliers, the cluster members are used to compute a density image by application of the smoothing kernel (equation (1)) with σ conditional on the number of member coordinates. Then from the image peak density an isolated image object, spatially bound by a minimum density threshold of 1 (the density of a single coordinate), is obtained by region growing. Member coordinates falling outside this object are considered outliers and removed from the cluster and the process repeated until there are no further eliminations.
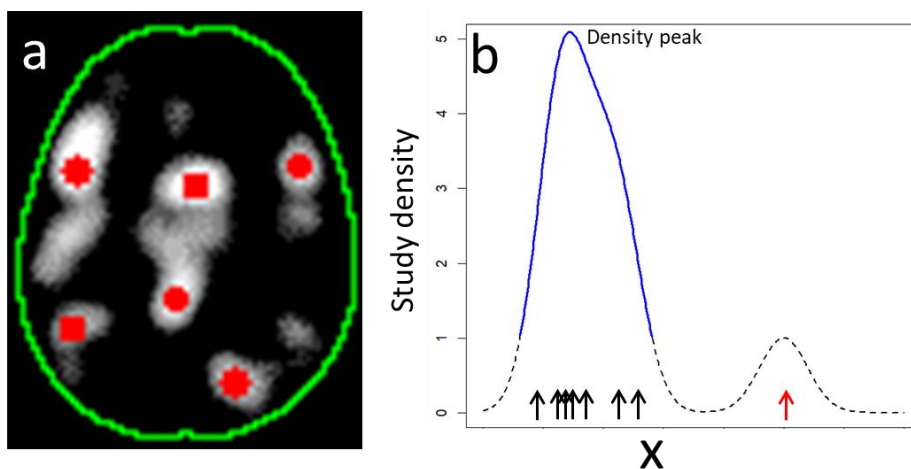

Figure 1. Panel (a) shows the study density image with detected density peaks (red) overlayed. These peaks form the proposed cluster centres to which coordinates are assigned. This can assign outlier coordinates outside the main cluster as shown in panel (b) where the coordinates within the main cluster are indicated by black arrows and an outlier, which causes a second peak, is indicated by a red arrow. Any coordinates falling outside of the main cluster bounded by a minimum density of 1 (the density of a single coordinate and shown as a blue curve in the figure) is considered an outlier and removed.

An iterative process of removing any clusters having fewer than four member coordinates is then performed, involving removing the cluster having the fewest member coordinates and the lowest density at its peak first, before repeating the clustering process until all clusters have four or more coordinates.

Standardised Z scores

The summary effect sizes reported by most studies are the peak Z score; t statistics and uncorrected p-values are converted to Z scores. This score depends on the number of

subjects in the study, so may deviate considerably from being normally distributed, which is undesirable when fitting MVN distributions. CBMAN uses standardised Z scores, as used in coordinate based random effect size (CBRES) meta-analysis [11] and where standardised is by the number of subjects

$$\varepsilon_i^a = \frac{z_i^a}{\sqrt{n_i^*}}. \qquad (2)$$

This specifies the effect size ε from study i in cluster a as a Z score standardised using the number of subjects n^* in study i. The number of subjects depends in whether the study is of a single group or a comparison between two groups. For single study groups the value of n^* is the number of subjects. For two group comparison studies the standardiser is

$$n_i^* = \frac{n_{i1} \times n_{i2}}{n_{i1} + n_{i2}}, \qquad (3)$$

where n_{i1} is the number of subjects in group 1 and n_{i2} is the number of subjects in group 2.

Fitting multivariate normal distributions to standardised Z scores

The general form of the k dimensional multivariate normal distribution is

$$f(\boldsymbol{\varepsilon}) = \frac{exp\left(-\frac{1}{2}(\boldsymbol{\varepsilon}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\varepsilon}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}}. \qquad (4)$$

For k clusters discovered by the clustering algorithm the network has k nodes and the assumption of CBMAN is that the standardised reported Z scores, ε, are distributed as specified in equation (4). The parameters, to be estimated by maximum likelihood, are then: $\boldsymbol{\mu}$ a column vector of means with k dimensions, and $\boldsymbol{\Sigma}$ a symmetric covariance matrix of size $k \times k$. Estimating the parameters of the MVN distribution can be a high dimensional problem if k is large; and as will be shown, a nonlinear one. Fortunately, it is a property of the MVN that marginal distributions over subsets of the dimensions are themselves MVN distributions with the same subset of parameters. In CBMAN this fact is utilised to estimate parameters by fitting bivariate normal (BVN) distributions to pairwise clusters, reducing the problem to estimating multiple sets of 5 parameters.

To fit BVN distributions to the standardised effect sizes in pairwise clusters at least two clusters must form. The five parameters to estimate by using maximum likelihood estimation (MLE) are: two means (μ) and two standard deviations (σ) plus a correlation (ρ). Standardised Z score pairs for study i in clusters a & b are distributed as

$$f\left(\varepsilon_i^a, \varepsilon_i^b\right) = \frac{1}{2\pi\sigma_a\sigma_b\sqrt{1-\rho^2}} exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{\left(\varepsilon_i^a-\mu_a\right)^2}{\sigma_a^2} + \frac{\left(\varepsilon_i^b-\mu_b\right)^2}{\sigma_b^2} - \frac{2\rho\left(\varepsilon_i^a-\mu_a\right)\left(\varepsilon_i^b-\mu_b\right)}{\sigma_a\sigma_b}\right]\right). \qquad (5)$$

To estimate the parameters the log likelihood (LL) is maximised. There is one additive term to LL for each study in the analysis. A subtlety in the calculation is that some studies may not

contribute a coordinate, and associated Z score, to either or both of the clusters. In this instance the effect size is interval censored; it is known only that the value does not exceed a threshold level, and it is assumed that the censored value is drawn from the BVN distribution. Statistically the contribution to the LL of interval censoring is computed by integrating over regions/lines of the BVN distribution, which is why the problem of fitting the MVN distribution in CBMAN is non-linear.

When the Z score from a study is not censored (the study contributes a coordinate to both clusters) the additive contribution to the LL is just the log of equation (5).

If the study contributes to only one of the clusters, say cluster *b*, one Z score is censored and known only to fall between ±α, where α is derived from the study threshold for significance using equation (2); the threshold is often reported, but is estimated by the minimum reported *Z* score otherwise. Then the contribution to LL is computed by integration of equation (5),

$$log\left(\int_{-\alpha}^{\alpha} f(x, \varepsilon^b) dx\right); \qquad (6)$$

and similar if censored in cluster *b*. The integral in equation (6) can be computed analytically using the error function and the conditional distribution for the standardised Z score in cluster *a* given the standardised Z score in cluster *b*, however for this report it is computed numerically [24] using Simpson's rule.

Another scenario considered by CBMAN is when a study contributes a coordinate to neither cluster. In this case both Z scores are interval censored and known only to fall between thresholds ±α. The contribution to LL is then

$$log\left(\int_{-\alpha}^{\alpha} \int_{-\alpha}^{\alpha} f(x, y) dx dy\right), \qquad (7)$$

which is also computed numerically.

There is one other type of censoring common in neuroimaging studies whereby coordinates are listed but no Z score is reported; such studies report only if the effect is positive or negative. The Z scores are then left or right censored, known only to be less than -α (left censored) or greater than +α (right censored) and assumed to be distributed according to equation (5). This type of censoring can be considered by integration over regions of the BVN distribution just as for interval censoring. The general LL term is

$$log\left(\int_{-m}^{n} \int_{-p}^{q} f(x, y) dx dy\right), \qquad (8)$$

where the integral limits depend on whether censoring is left or right, or interval if the study reports no coordinate within the cluster. Table 1 indicates the limits for each scenario. Integrals with these limits are computed numerically, with ∞ limits approximated by α +6;

this multiple standard deviations beyond the typical values and increasing this has been found not to alter the results in any significant way.

Censor type	Lower integral limit	Upper integral limit
Left	$-\infty$	$-\alpha$
Right	α	∞
Interval	$-\alpha$	$+\alpha$

Table 1. Integral limits for the calculation, using equation (8), of the LL contribution for studies that do not report Z scores, t statistics, or uncorrected p-values.

Thresholding the correlation between coactivated clusters

The correlation of standardised Z scores (edges) between coactivated clusters (nodes), estimated by maximising the likelihood, determines the coactivation network in CBMAN. Since estimates are unlikely to be exactly zero this produces a fully connected network (every node connected to every other node), which is not very revealing. A method of thresholding the edges is needed to reduce the network to meaningful connections between nodes by some criteria. In network analysis of the brain this thresholding is still an open question but requiring correlation to be above a minimum magnitude is a common strategy [25]. Optionally the user may set a correlation threshold in CBMAN, however this is only for the purpose of visualising simplified networks rather than producing meaningful results because the number of coordinates per cluster is not fixed; correlation between two small clusters would require a larger threshold than between two large clusters for the same evidence of correlation. In CBMAN the primary method of thresholding is statistical and uses a permutation test. The aim is to identify correlation of reported standardised Z scores between clusters, and the test used addresses the problem that censored studies can impose nonzero correlation even in the absence of correlated Z scores. For example, if half of studies contribute uncensored and uncorrelated Z scores to coactivated clusters, while the other half are censored, the dichotomy of studies can produce apparent significant correlation; such dichotomy might be of interest but is not the aim of CBMAN. To ensure that the test is sensitive to correlated standardised Z scores and not the result of censoring, only the uncensored data are permuted while censored data acts to constrain the bivariate normal model fit. The constrained correlation under permutation then acts as the null hypothesis.

Testing for significant correlation of standardised Z scores between two coactivated clusters requires that at least five studies report uncensored results in both clusters. While with just four studies there are 24 permutations yielding a minimum possible p-value of 0.042, which is slightly less than the typical 0.05, this is ultimately unlikely to be declared significant after correction for multiple tests and would increase the number of tests making the correction

more conservative. Significance testing proceeds by fitting of the BVN distribution to estimate the correlation of effects before repeatedly re-estimating having permuted the uncensored effects in one cluster; this is the normal approach to estimating significance of correlation using a permutation test but the null hypothesis is not zero correlation because of censoring. The p-value is the proportion of correlation estimates that are as, or more, extreme under permutation than the unpermuted estimate; positive correlation must be more positive, while negative correlation must be more negative. Permutation only affects the correlation parameter of the BVN so the problem reduces to a computationally undemanding one dimension. If there are 7 or fewer uncensored effects all possible permutations are computed (for 7 the number of permutations is just 7!=5040) using Heap's algorithm [26], otherwise the p-value is estimated using 5000 random permutations.

The number of edges is quadratic in the number of nodes in the network so correction for multiple statistical tests is necessary and is handled by FDR. The false discovery rate is known to cause issues when used in voxel-wise analyses because of the possibility of very many positive voxels, which causes a high expected number of false positive voxels potentially producing multiple false positive clusters. However, since CBMAN tests network-edge-wise rather than voxel-wise it produces a relatively small number of positive results and consequently a small expected number of false positives; less than 1 if fewer than 20 significant effects for FDR of 0.05. This property was demonstrated previously by comparing the voxel-wise FDR option in the ALE algorithm to the cluster-wise FDR used by the CBRES algorithm [11].

Experiments

In this report the concept and computation are validated using simulated data. To demonstrate applied utility, coordinates extracted from published studies are used. In the absence of available software for other coactivation based methods [13-19], the popular ALE algorithm is used for comparison. While this does not produce results based on coactivation it is useful to contrast the results with CBMA, and where there is agreement on the clusters the two methods are expected to be similar given the same data, which would in part validate the clustering algorithm used in CBMAN.

Coordinate systems and images

CBMAN uses the grey matter tissue class image from the ICBM 452 atlas [27]. Transformation between MNI (Montreal Neurological Institute) and Talairach [28] space is performed as detailed in [29].

Establishing clustering settings

An empirical function relating the smoothing kernel width (σ) to the number of member coordinates (n) in a cluster is derived by simulation. This is done in a principled way given that coordinates within clusters have a spatial standard deviation of between 3mm and

5mm with a mode at around 4mm [9]. For each value of σ, 5000 simulations are performed for clusters of 2, 3, 4, 5, 10, 15, 20, and 30 coordinates, and in each the number of density peaks counted. For a given n the aim is to choose σ such that on average false peaks (more than one peak per simulated cluster) are rare while also avoiding over smoothing; false peaks can prevent formation of clusters by incorrectly splitting the coordinates from one cluster into multiple clusters. This is achieved in practice by the minimum σ to produce only a small percentage of false peaks per experiment while also avoiding false negatives whereby coordinates are excluded from the simulated clusters.

Simulated meta-experiments

Numerical simulation is used to validate the code and to test: the ability to estimate parameters of the BVN distribution when data are censored, and the ability to identify known networks.

To test parameter estimation for censored data BVN data (50 experiments in total) are simulated with random parameters μ_a, μ_b, σ_a, σ_b, ρ; mean parameters are uniform random $0.6 \leq \mu \leq 1.2$, standard deviations are uniform random $0.2 \leq \sigma \leq 0.6$, and correlation is uniform between ±1. These means and standard deviations represent the range of results presented in [10]. Data are interval censored if they are $<3.09/\sqrt{20}=0.7$ (representing the minimum for commonly used Z score threshold of 3.09, and 20 subjects), and 10% of studies are left/right censored; reflecting the observation that the majority of studies do report Z scores. The known simulation parameters and the corresponding estimates are qualitatively compared by scatter plot for examples with 20 studies and 50 studies to test the ability of MLE to estimate the BVN parameters. To show the impact of censoring on estimation, least squares estimation is used to estimate the parameters on the uncensored data and the correlation parameter compared, graphically, to the MLE estimate after censoring; a histogram of the proportion of censored data is also depicted.

A simulated network experiment involves 6 coactivated clusters and four different covariance matrices generating four different networks. Z scores are simulated with a mean of 5, and a standard deviation of 1. Censoring is at a Z score of 3.09, the number of subjects is 20, and the number of studies 30. The correlation parameter is set at a level of 0.707 ($R^2=0.5$), and correction for multiple statistical tests is using FDR=0.05. The first network simulated involves a diagonal covariance matrix, so no network is expected since the correlations should be statistically zero; this is an important example as no network should be produced providing the correction for multiple tests is successful. A second network is simulated with covariance matrix having only one off-diagonal covariance for which the expected network has two nodes and one edge. A third network is simulated with a block diagonal covariance matrix generating two independent networks. Finally, a covariance matrix with all off-diagonal elements involving the correlation is expected to produce a fully connected network.

Real data example: fMRI of painful stimulation

An example of a functional MRI meta-analysis, painful stimulation, was provided in the CBRES MA report. Here that data are used for CBMAN.

Real data example: Voxel based morphometry of Multiple Sclerosis

It is well known that multiple sclerosis causes atrophy of the grey matter, and repeatable patterns of atrophy have been demonstrated by multiple VBM studies. The data used in this example are taken directly from the CBRES report. Both CBRES meta-analysis and CBMAN are performed.

Sensitivity of CBMAN to the clustering algorithm

The clustering algorithm employed by CBMAN has two parameters: the false peak rate and the minimum number of coordinates needed to form a valid cluster. In CBMAN the chosen FPR is 1% and minimum number of coordinates is 4. To test the sensitivity to these parameters both the MS and pain coordinate data are analysed for FPRs of 0.1%, 0.5%, 1%, and 5% and minimum number of coordinates of 2, 3, 4, and 5.

Results

Establishing clustering settings

By simulating coordinates with standard deviation drawn at random from $N(4,0.4^2)$ truncated to between 3mm and 5mm an empirical function relating the smoothing kernel σ to the number of coordinates was established. Simulations were performed for clusters of 2, 3, 4, 5, 10, 15, 20, and 30 coordinates each for a range of σ, and for each the presence of false peaks (more than one peak per simulated cluster) noted. Figure (2) shows examples of average false peak rates (FPR) as a function of σ for 5, 10, 15, and 20 coordinates, and also shows that exponential functions describe this data (dotted lines). By estimating the exponential trend the value of σ producing average false peak rates of 0.1%, 0.5%, 1%, 5%, and 10% per cluster was computed for each number of coordinates and plotted in figure (3).
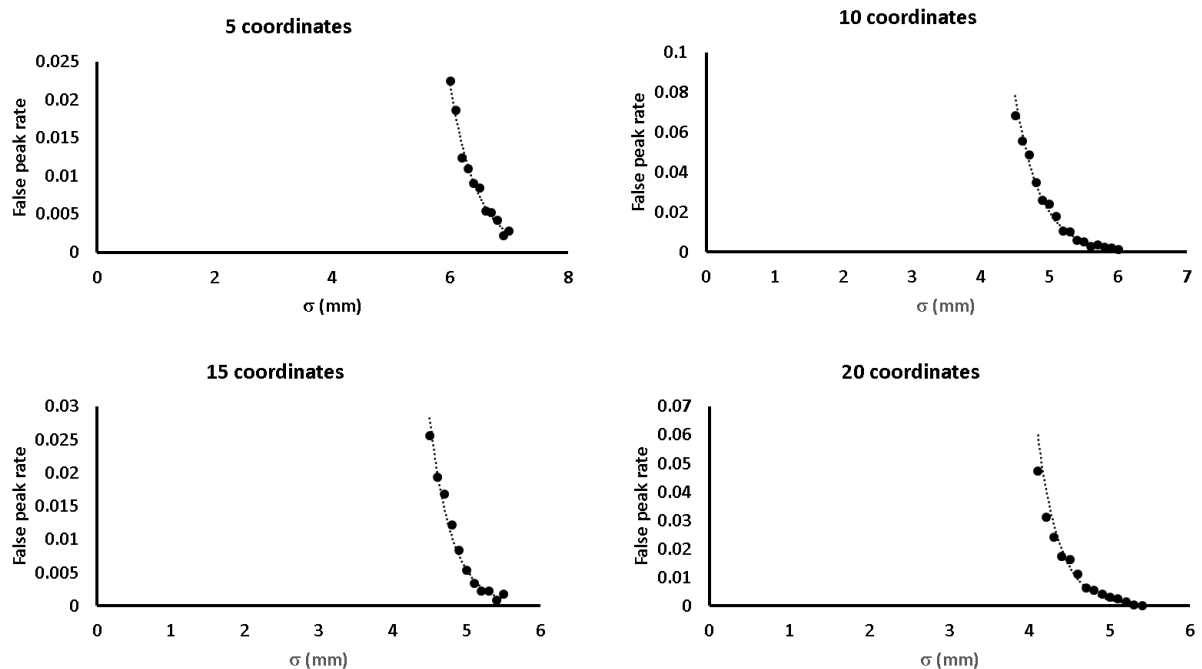

Figure 2. Average false peak rates per cluster shown as a function of the smoothing kernel σ for between 5 and 20 coordinates. The fitted trend lines (dotted lines) have the form r(σ)=A×exp(-b×σ), where A and b are parameters.
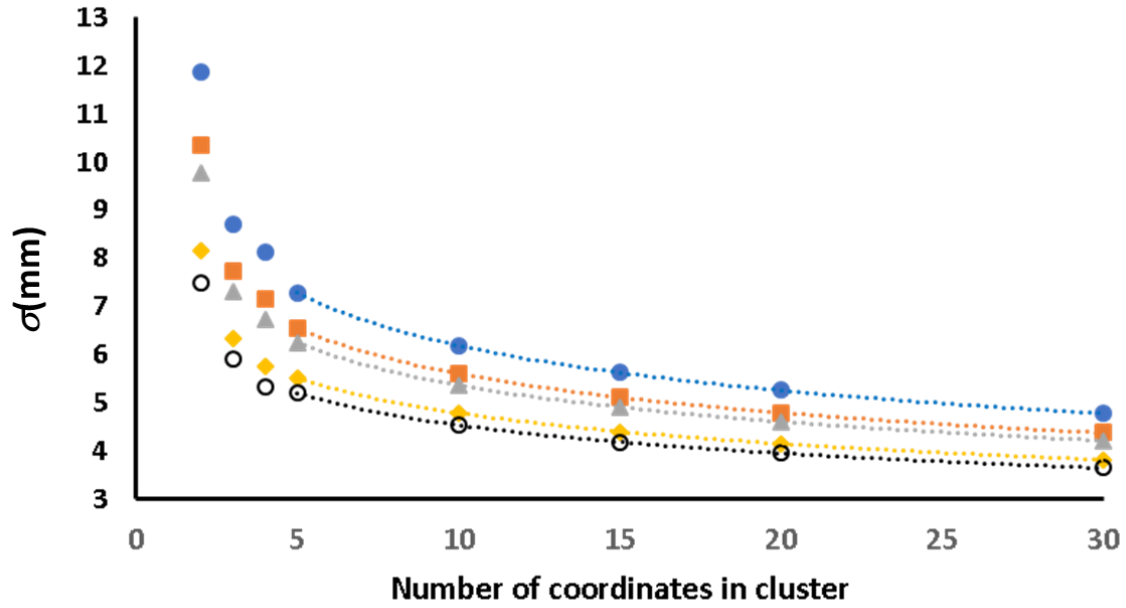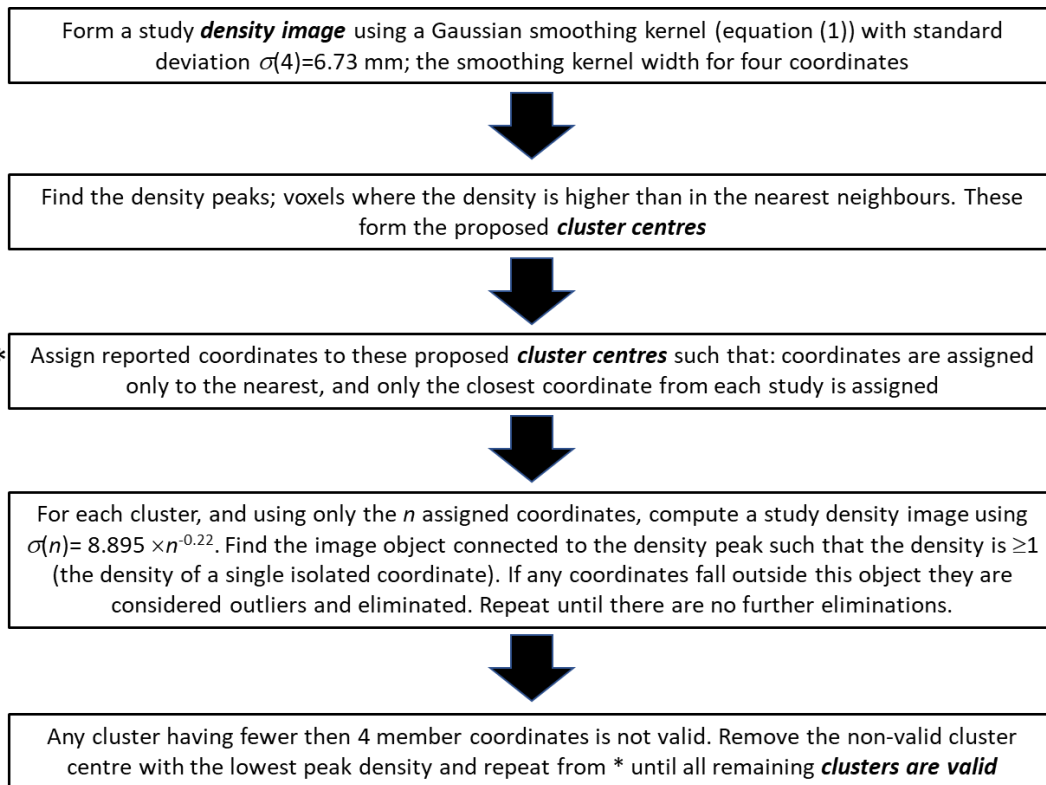
Figure 3. The estimated value of the smoothing kernel width σ as a function of the number of coordinates in a cluster for average false peak rates of 0.1% (blue circle), 0.5% (orange square), 1% (grey triangle), 5% (yellow diamond), and 10% (unfilled circle). For five or more coordinates there is an empirical relationship between σ and the number of coordinates (n) with the form $\sigma(n)=A \times n^{-x}$ having parameters A and x. For $n \leq 4$ σ is not well described by this power function.

Choice of model for σ from this analysis is a trade-off between excessive smoothing with σ too large and false density peaks if σ is too small. Given the false peak rate the number of study density peaks per study such that there is good chance (80%) no false peaks is binomial: n=220 for FPR=0.1%, n=44 for FPR=0.5%, n=23 for FPR=1%, n=4 for FPR=5%, and n=2 for FPR=10%. Clearly for a false peak rate of 0.1% the number of peaks per study is in the order of the number of coordinates typically, and so represents smoothing by an unnecessarily large σ. On the other hand, for FPR of 5% or 10% even a few study density peaks result in a high chance of false peaks. The pain and MS data used in this report produce around 30 peaks, so choice of 1% false peak rate used here would seem reasonable, with about 74% chance of no false peaks in these examples. Consequently, the empirical relationship between σ and the number of coordinates n in the cluster used in CBMAN is

$$\sigma(n) = 8.895 \times n^{-0.22} \qquad\qquad (9)$$

for $n \geq 5$ coordinates, while $n<5$ is treated as special cases. Since this choice is not the only one possible, others will be considered in the results on the MS and pain coordinate data. The final clustering procedure used in CBMAN is depicted in figure (4).

Form a study **density image** using a Gaussian smoothing kernel (equation (1)) with standard deviation $\sigma(4)$=6.73 mm; the smoothing kernel width for four coordinates

⬇

Find the density peaks; voxels where the density is higher than in the nearest neighbours. These form the proposed **cluster centres**

⬇

* Assign reported coordinates to these proposed **cluster centres** such that: coordinates are assigned only to the nearest, and only the closest coordinate from each study is assigned

⬇

For each cluster, and using only the n assigned coordinates, compute a study density image using $\sigma(n)$= 8.895 $\times n^{-0.22}$. Find the image object connected to the density peak such that the density is $\geq$1 (the density of a single isolated coordinate). If any coordinates fall outside this object they are considered outliers and eliminated. Repeat until there are no further eliminations.

⬇

Any cluster having fewer then 4 member coordinates is not valid. Remove the non-valid cluster centre with the lowest peak density and repeat from * until all remaining **clusters are valid**

Figure 4. The procedure for discovering valid clusters for CBMAN.

Figure (5) demonstrates the average simulated cluster volumes produced using equation (9) and the cluster forming algorithm used by CBMAN. An important feature is that σ is reducing with the number of coordinates making up the cluster resulting in volumes that tend to converge for higher numbers of coordinates, which is not true for fixed smoothing [8, 9, 11]. Indeed, fixed smoothing kernels can paradoxically result in escalating false positives as the number of studies analysed increases [8] because cluster volumes also increase to inevitably include coordinates that do not belong. Importantly, using equation (9) to define the smoothing kernel the rate of false negative coordinates (simulated coordinates not included in the cluster) was always below 0.36% for these simulations.
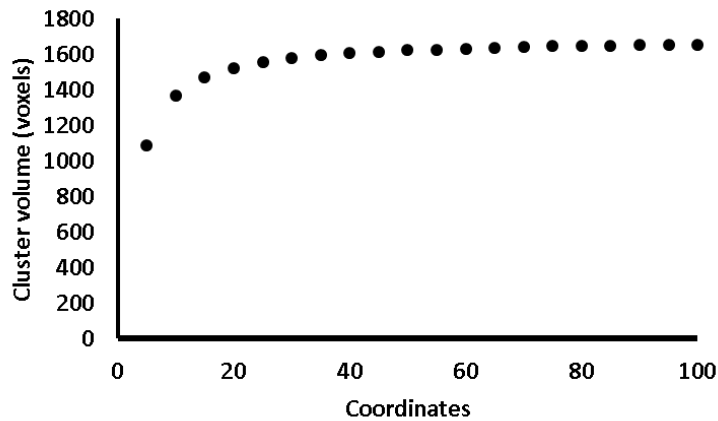
Figure 5. The simulated cluster volumes produced using σ giving an estimated false peak rate of 1%. The cluster volume tends to a fixed volume for large numbers of coordinates because of the monotonically reducing σ.

Simulated data experiment

It is a requirement that accurate estimates of correlation of standardised Z scores between coactivated clusters is possible for a reasonable number of studies; typically, CBMA involves 10's of studies. For simulated BVN distributed data, using realistic standardised Z scores and censoring thresholds, figure (6) shows error due to maximum likelihood estimation of parameters. With 20 studies (top) the parameters can be estimated, but as expected the sampling error is evident. Nevertheless, simulations of strong correlation are quite reliable. Increasing to 50 studies considerably improves estimates and therefore reliability of the analysis.

While figure (6a) shows that estimation of parameters from a small sample is possible when the data are censored, it does not show how much variance might be due to the censoring. By estimating the parameters before censoring using analytic least squares estimation, the component of variance due to censoring can be explored. Figure (6b) shows the correlation parameter estimated by MLE on censored data plotted against the least squares estimate on uncensored data. It also shows the distribution of the proportion of data censored. Clearly the major source of variance in the parameter estimates is the sampling, rather than the censoring; despite an average of 38% of data being censored in this experiment.
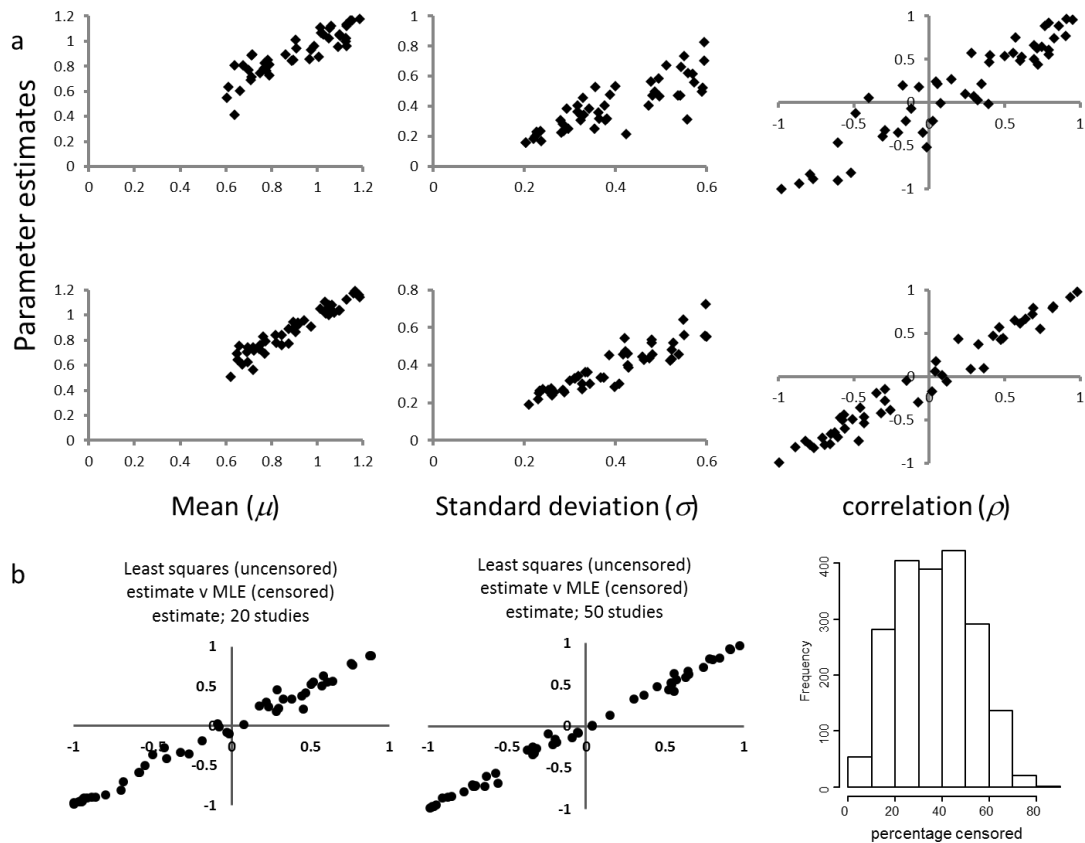
Figure 6. Demonstration of parameter estimation using simulated censored data. In figure (6a) the top row is an example with 20 simulated studies, representing a small meta-analysis. The bottom row of figure (6a) uses 50 simulated studies, representing a moderate sized meta-analysis. Estimates are better for the larger number of studies, as expected. In figure (6b) the impact of censoring is considered by comparison of analytic least squares estimates of the parameters using uncensored data and the MLE estimates using censored data. These show that censoring is not a major source of variance on the parameter estimates, despite 38% data censoring on average in this experiment.

By simulating known coactivation networks CBMAN can be shown to detect MVN censored data and threshold the results by FDR. Networks with 6 clusters (nodes) were generated with various MVN covariance matrices. The 6 clusters were all reliably detected by CBRES MA for each network variant, and the resulting clusters shown in figure (7).
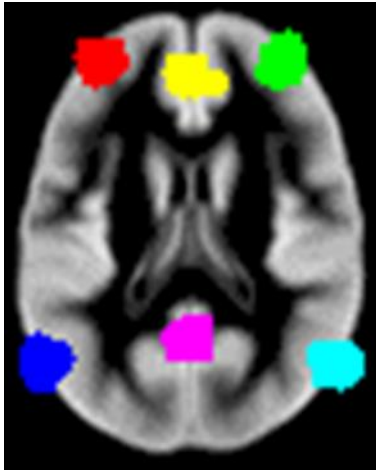
Figure 7. The clusters detected by CBRES MA of the simulated network test data; these results are independent of the network structure.

Figure (8) shows the results of CBMAN on 4 different network configurations that use the 6 clusters shown in figure (7); graphs are produced using automatically generated R [30] code and requires the igraph package [31]. The first (top left) shows the network with a diagonal covariance matrix where all correlations are statistically zero, hence no network is detected. This example also highlights a difference between CBMAN and other coactivation methods using only coordinates, which might detect a significant coactivation network despite the zero covariance of the standardised Z scores. The second network (top right) includes only one off-diagonal element (covariance) and so the network consists of only two nodes connected by one edge. The third example (bottom left) shows a network with a block diagonal covariance matrix, which forms two independent networks as depicted. Finally (bottom right) the covariance matrix specifies a network where effect sizes in each cluster are correlated with every other, as successfully found by CBMAN. Contrasting figure (7) with results shown in figure (8) demonstrates the difference between CBMA and CBMAN; CBMAN considers how the clusters coactivate as measured by correlated reported Z scores, while CBMA has no mechanism to consider such correlation.
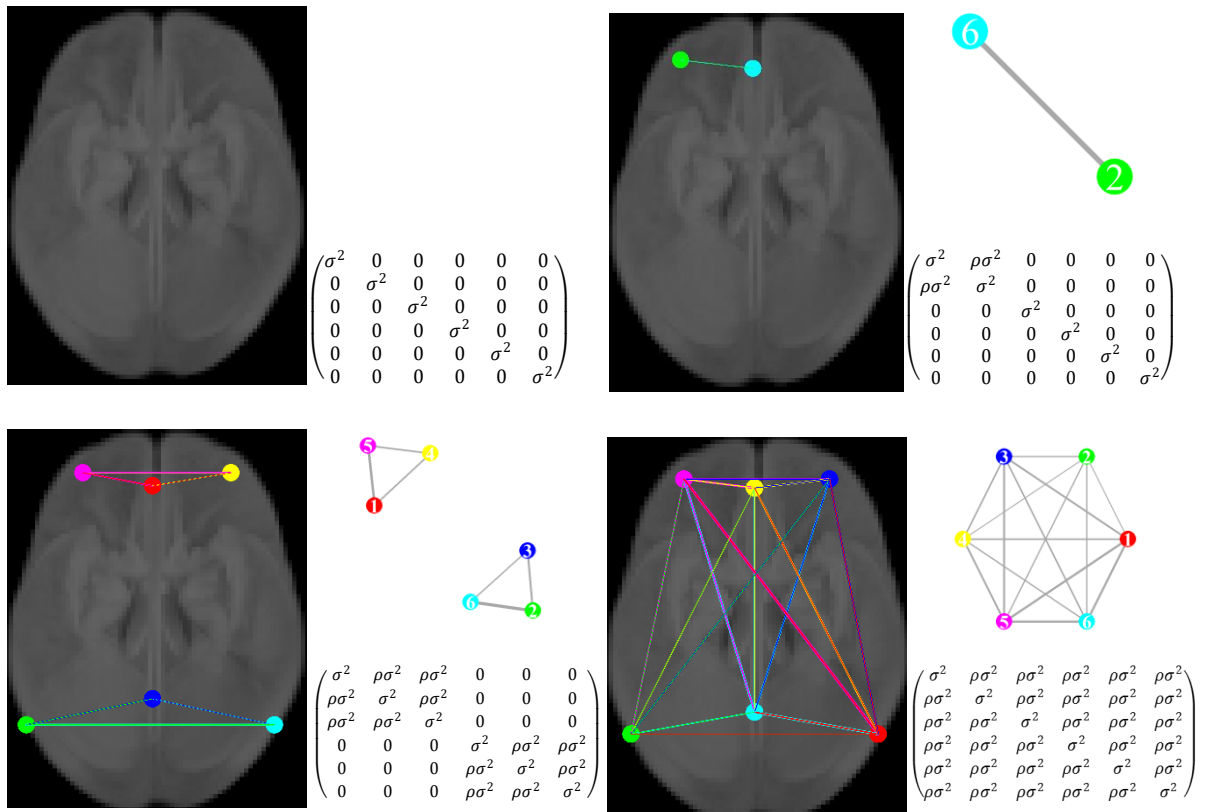
Figure 8. Results of coordinate based meta-analysis of networks. Four networks are simulated using the clusters shown in figure (7). In each case the statistically significant nodes (clusters) and edges are shown, along with the associated graph and the form of the covariance matrix of the multivariate normal distribution of effects that define the network.

Painful stimulus in healthy subjects

Functional MRI studies using painful stimulation of healthy subjects were collected and used to demonstrate CBRES MA previously [11]. This data includes 22 independent studies of 315 healthy volunteers and consists of 361 coordinates and the associated Z scores. Figure (9) shows clusters found significant by CBMAN (using FDR=0.05) and by the ALE algorithm (using only the coordinates and the recommended cluster level FWE (Family Wise Error) at 0.05 and cluster forming threshold of 0.001). There are many significant correlations as indicated by the graph plot, and often these are between symmetric clusters. An example of a significant correlation is plotted in figure (10) along with an example of where correlation is not significant. Comparing the ALE result to that of CBMAN shows more significant results from CBMAN for this data. This might be due to the use of FDR instead of the more conservative FWE, or because of the different hypothesis being tested. Nevertheless, where the two algorithms agree the clusters are quite similar, which helps to validate the CBMAN clustering method.
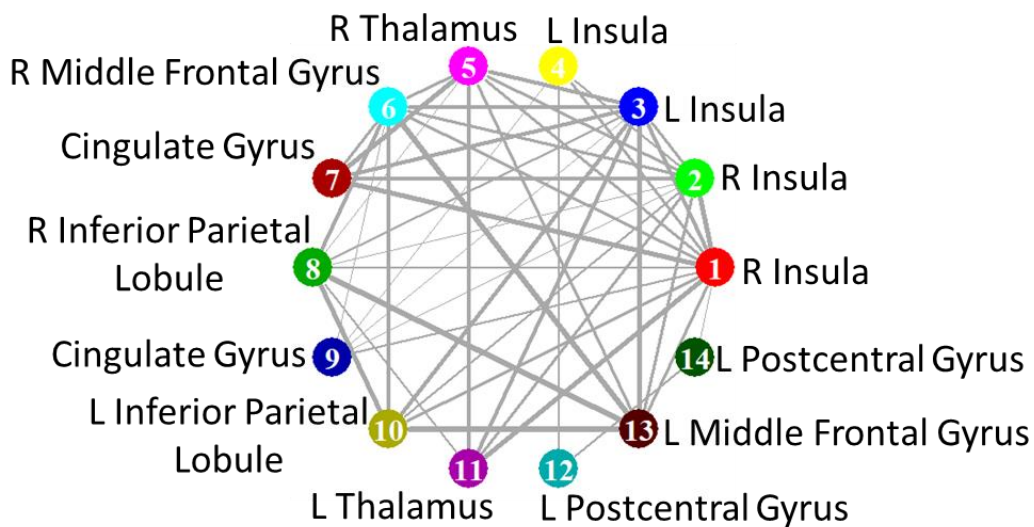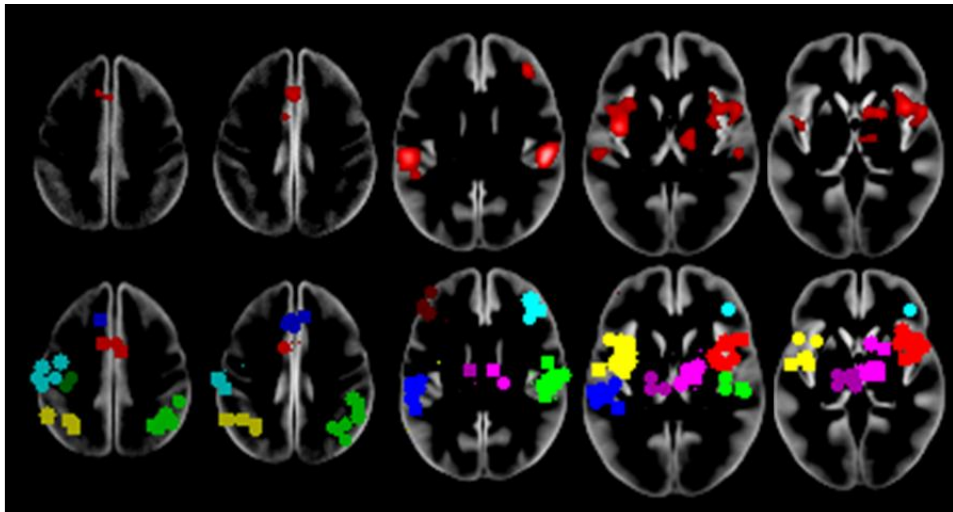
Figure 9. The ALE MA (top cluster images) and CBMAN (bottom cluster images) analysis on functional MRI studies of painful stimulus. The graph indicating correlation of standardised Z scores between many coactivated clusters is also shown (bottom row).
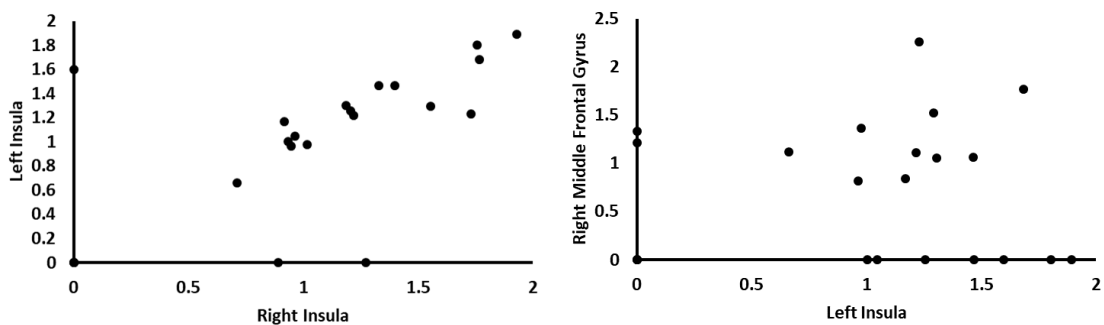

Figure 10. Scatter plot demonstrating the positive correlation of standardised Z scores in the left/right insula (clusters 1 & 4 in graph in figure (9)) and a non-significant correlation for left insula v right middle frontal gyrus (cluster 4 & 6 in graph in figure (9)) for the pain coordinate data. Markers falling on the axes indicate censored effects.

Voxel based morphometry in Multiple Sclerosis

VBM studies of MS patients were collected to demonstrate CBRES MA previously [11], and are used unaltered here to demonstrate CBMAN. This data consists of 27 independent studies of grey matter atrophy in MS and clinically isolated syndrome patients comparing to healthy controls; in total the study includes 871 patients and 671 healthy controls and includes 333 coordinates and associated Z scores. Figure (11) shows the clusters declared significant by both the ALE method and CBMAN; ALE analysis uses only the coordinates and the recommended cluster level FWE at 0.05 and cluster forming threshold of 0.001, while CBMAN used FDR=0.05. The correlation between symmetric clusters is generally significant. An example of significant correlation is plotted in figure (12) along with an example where the correlation is not significant. Just as with the pain coordinate analysis, there are differences between the ALE results and the CBMAN results but there is also good agreement for some clusters.
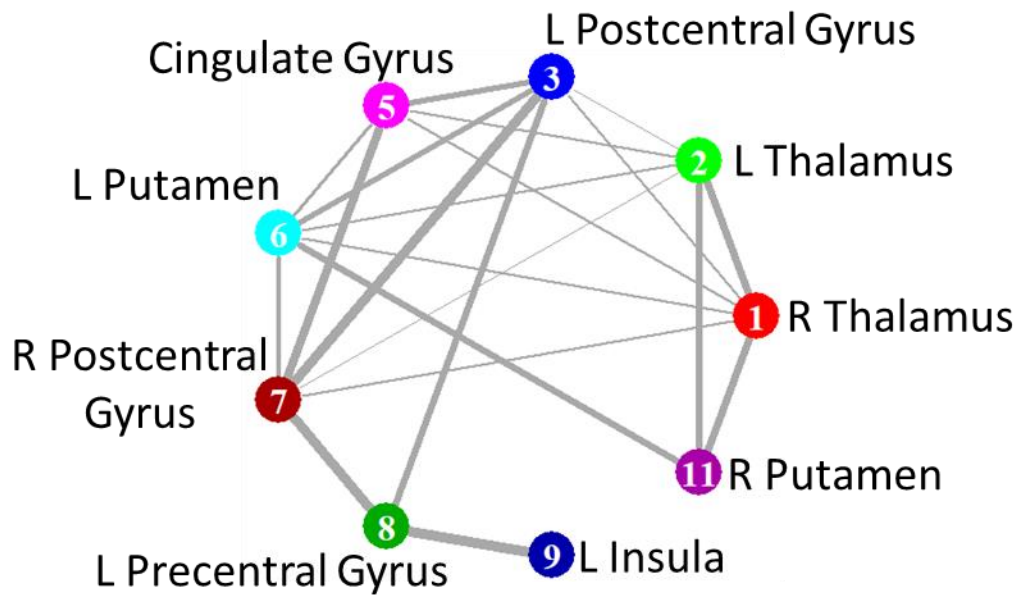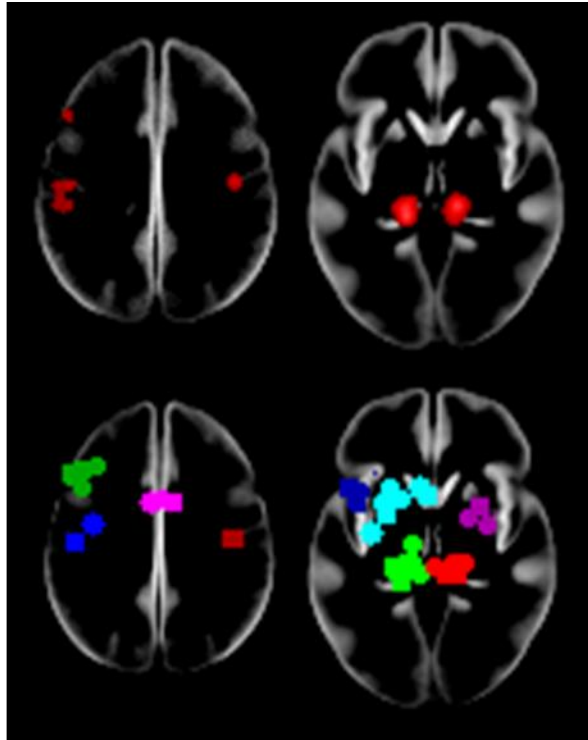
Figure 11. Clusters found by ALE meta-analysis (top row of cluster images) and CBMAN for VBM studies of MS. The graph indicating correlation of standardised Z scores between coactivated clusters is also shown (bottom row).
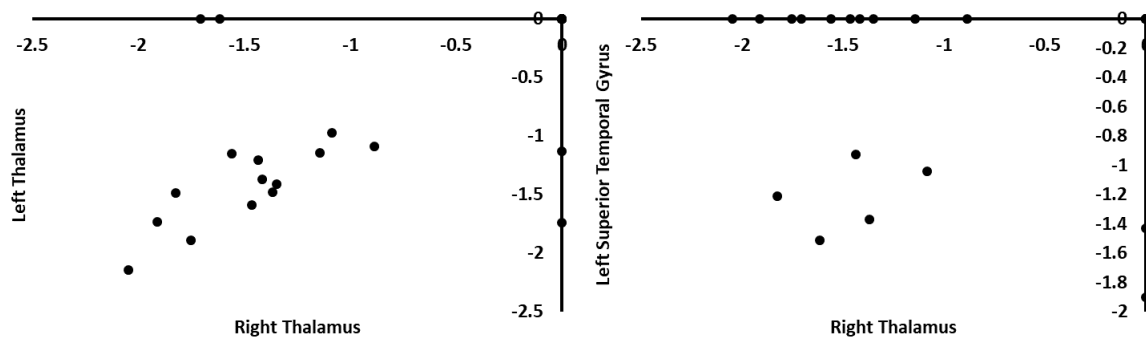
Figure 12. Scatter plot demonstrating the positive correlation of standardised Z scores in the left/right thalamus (clusters 1 & 2 in graph in figure (11)) and non-significant correlation between the right thalamus and left superior temporal gyrus (Thalamus is cluster 1 in figure (11) while the temporal gyrus cluster is missing because of lack of significance) for the MS coordinate data. Markers falling on the axes indicate censored effects.

Sensitivity of CBMAN to the clustering algorithm

The pain and MS coordinate data were reanalysed with false peak rates of 0.1% to 5% and minimum number of coordinates for valid clusters between 2 and 5. Resulting networks of coactivated clusters are plotted such as to preserve the Talairach coordinates of the cluster peaks; for a peak at Talairach location x, y, z the location on the plot is $(z+z_0)\times x/(x^2+y^2)$, $(z+z_0)\times y/(x^2+y^2)$, where z_0 is the minimum Talairach Z coordinate. The plot axes indicate the x=0 and y=0 lines as well as the z=0 circle, where inside the circle z is negative. This makes comparison of the clusters between analyses easier since the same cluster peak location appears on the plot in the same location.

Figure (13) shows sensitivity to the choice of false peak rate: left is the MS VBM coordinates example and right is the pain fMRI coordinates example. In each figure the FPR is shown top left. Comparing to the FPR used in CBMAN (1%) the most different is at FPR 5%, where the smoothing kernel width is smallest and consequently some of the clusters have not formed as anticipated. For FPR of 0.1% and 0.5% there are differences, but many of the clusters have formed and many of the significant edges are the same as those at FPR of 1%. FPR of 0.1% arguably over smoothes to achieve low risk of false peaks. The choice between FPR of 0.5% or 1% is less clear, but the strongest edges (thickest lines) are similar, as are the clusters, and either choice would lead to largely similar interpretation of the analysis.
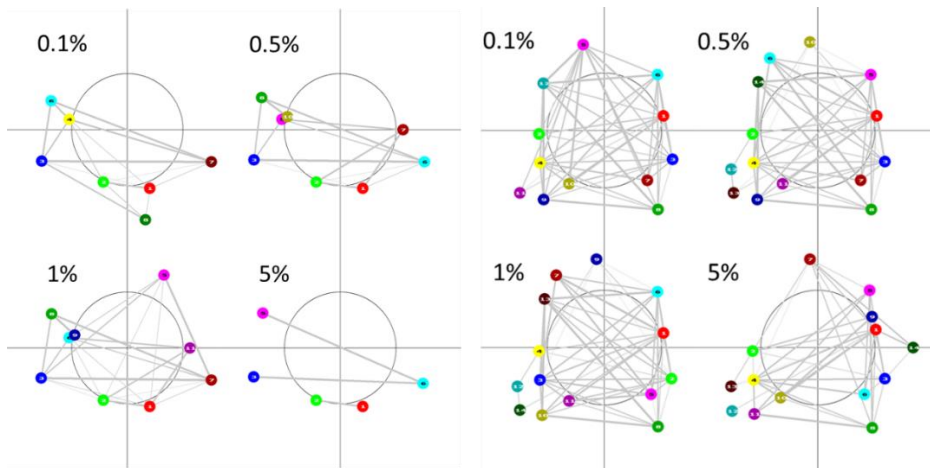
Figure 13. Clustering algorithm sensitivity to false peak rate. For 1% or lower many of the clusters are similar, as are some of the edges. At 5% the reduced smoothing kernel width has eliminated some clusters by causing false peaks.

Figure (14) shows sensitivity to the choice of minimum coordinates for valid clusters: left is the MS VBM coordinates example and right is the pain fMRI coordinates example. In each figure the minimum is shown top left. Comparing to the minimum used in CBMAN (4) the most different is with a minimum of 2, where the smoothing kernel width is large (see figure (3)) so over smoothing has supressed the density peaks. For minimum of 3 or 5 there are differences, but many of the clusters have formed and many of the significant edges are the same as those at a minimum of 4. Choice of 3, 4, or 5 produces relatively similar results that would not greatly modify the interpretation of the analysis. Increasing the minimum further eliminates clusters formed from too few coordinates.
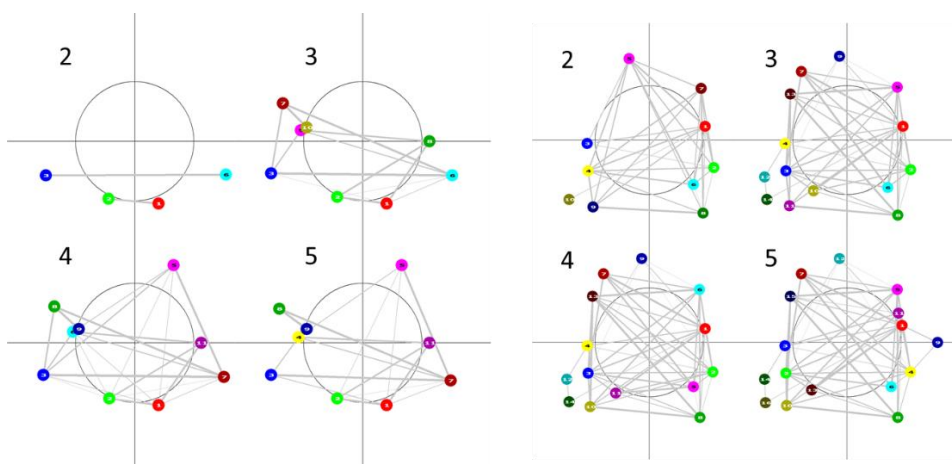

Figure 14. Clustering algorithm sensitivity to the choice of minimum number of coordinates for a valid cluster. For a minimum of 4 or 5, there are few differences in the results. For a minimum of 2 the smoothing kernel width is quite large (see figure (3)) and has smoothed away some peaks. Increasing the minimum beyond 5 eliminates clusters that are too small.

Discussion

Here a method of performing a meta-analysis of summary results reported by functional MRI or voxel-based morphometry studies has been presented. The fundamental assertion is that the multiple summaries (coordinates and Z scores) reported by each study represent a network of coactivated (co-altered in the case of VBM) effects. CBMAN uses only the data needed to perform coordinate based meta-analysis but tests a different hypothesis, and software to perform the analysis (NeuRoi) is freely available.

It is an implicit assumption of CBMA that the results tabulated in fMRI and VBM articles are independent, or at least any dependency is not considered. However, multiple activations/grey matter changes might also be considered as a pattern of coactivated or co-altered effects. In CBMAN the pattern is formed by multiple spatially separated but coactivated clusters with standardised Z scores that correlate. Fitting a high dimensional MVN distribution to the standardised Z scores is a non-linear problem, however the mathematical property that the marginal bivariate normal distributions involving pairwise clusters have the same parameters as the respective MVN parameters makes fitting feasible in the presence of censored data. Maximum likelihood estimation for censored data is by integration over portions of the BVN distribution, which is accurately handled numerically.

Coordinate based meta-analysis of networks and coordinate based meta-analysis should be considered complementary. If CBMA reveals multiple significant clusters while CBMAN reveals no clusters, this might indicate lack of coactivation despite the consistent involvement of regions highlighted by the clusters or uncorrelated reported Z scores; a simulated example of this was demonstrated using a network generated by MVN effects with a diagonal covariance matrix (see figure (8)). On the other hand, if standardised Z scores are detectibly correlated between spatially separate but coactivated regions CBMAN might reveal a network whether such effects are detected by CBMA or not.

The CBMAN algorithm relies on coordinate clustering using a density-based method. There is no strict definition of cluster but using the observed features of coordinates [9] the smoothing kernel width was chosen in a principled way through simulation. It also requires a minimum number of coordinates making a valid cluster, and here 4 is chosen to be in line with the popular DBSCAN algorithm [32] that requires a minimum D+1, where D is the dimensionality of the space. It was shown in figures (9) and (11) that the clusters detected by CBMAN are similar to those reported by the popular ALE algorithm, which has been utilised in many meta-analyses; while this is not a validation, it at least indicates a consistency with an established method despite the different approach to clustering. Nevertheless, these parameters might have been chosen differently so the impact of alternative reasonable choices has been explored. Figures (13) and (14) depict analyses using different false peak rates and minimum numbers of coordinates for valid clusters, and the results are quite predictable from the affect these parameters have on the smoothing

kernel width (figure (3)) used to define the clusters. For 5% FPR, which has a relatively small smoothing kernel width, it is expected that some clusters fail to form and indeed multiple clusters edges are no longer found significant compared to FPR 1%. For FPR less than 1% the clusters are modified as the smoothing kernel width increases, but many of the findings remain consistent. Figure (3) shows how σ increases rapidly below a minimum of 4 or 5 coordinates per valid cluster. Therefore, requiring only 2 coordinates to define a valid cluster demands a large smoothing kernel width to achieve 1% FPR and consequently many of the study density peaks used to define clusters are not detected. For a minimum of 3 coordinates most of the coactivation network structure has been detected compared to the chosen minimum of 4. Increasing the minimum further eventually leads to elimination of clusters that are too small.

The CBMAN algorithm relies on a permutation test to threshold the edges with sufficient evidence for correlated Z scores. Permutation is restricted to the uncensored data and constrained by censored data. This is a modification of the normal permutation test for correlation, which is necessary because censored data can impose non-zero correlation even when the effects are permuted; testing against a null hypothesis of zero correlation can produce significant results even in the absence of correlated effects because of censoring. Significant results from this test have standardised Z scores that correlate statistically more than the non-zero null, which is more demanding than the usual null hypothesis of zero correlation.

In this report CBMAN has been validated using simulated data with known properties. Figure (6) indicates the ability to estimate a realistic range of parameters for 20 and 50 studies with simulated censoring and as expected estimation is better with more studies. The figure also shows that parameter estimates are not substantially affected even by large proportions of data being censored. Beyond validation coordinates from published studies have also been used to demonstrate example output from analysis of real data. The two examples are analyses of functional MRI studies (painful stimulation) and of VBM of grey matter (in multiple sclerosis). This data was originally collected to demonstrate CBRES meta-analysis and have not been altered for CBMAN, emphasizing that no additional data are required than for CBMA; only coordinates and Z scores are needed. The strongest correlations tended to relate to spatial symmetry in the results, and this was true for both functional and VBM analyses.

Like all coordinate-based algorithms results should be judged critically, as they have elements that are empirical rather than based on established fact. For example, some algorithms use a null hypothesis generated by randomising coordinates into a 3D space, yet there is no strong consensus on the method of randomisation or the specification of the space and each algorithm opts for different definitions. CBMAN does not perform randomisation, but empirical clustering algorithm choices are still required. Furthermore,

methods based on just coordinates tend not to take account of studies reporting no coordinates, while at least some of [11, 33, 34] those that also use the Z scores as an effect size make corrections for censoring. However, the Z scores themselves represent only a statistical effect and are not biologically meaningful. Other limitations include the choice of correction for multiple statistical tests, with some authors suggesting combining empirical p-value and z score thresholds [35] while others use either FDR [11] or FWE [4, 6]. Therefore, the results obtained from coordinate-based analyses are dependent not only on the data being meta-analysed, but also the choice of analysis method; a similar problem is true of the software packages used in the studies themselves [36]. Coordinate based methods should therefore be considered indicative of potentially interesting effects and used to generate hypotheses. These could then be tested in well-designed prospective studies.

The requirements of performing and reporting CBMAN analysis are similar to those of CBMA. Firstly, the method assumes that studies are independent. It is important that multiple experiments on the same subjects are not considered independent [4, 37] as this will produce a known form of bias common to meta-analysis [38], and consequently reduce the quality of evidence. It is also important to provide the data analysed along with any publication; typically, multiple experiments are reported per study and it can be difficult to know which experiments have been included, and therefore to reproduce the analysis, without the data. Provision of data in any meta-analysis is a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [39] requirement, and only involves inclusion of a small text file. The use of principled control of the type 1 error [40] is also necessary in meta-analysis to provide evidence of effect. In CBMAN, principled control of the type 1 error is by FDR [22], which should be set appropriately small (FDR≤0.1) to prevent possible excessive false positive results. For reports the location of the clusters (nodes) are relevant, just as for CBMA.

Summary and conclusions

Coordinate based meta-analysis is a very popular approach for discovering regions consistently activated (fMRI) or with altered grey matter (VBM) over multiple independent neuroimaging studies. Another approach is to analyse for regions consistently coactivated or co-altered, and this has previously been performed using just the reported coordinates. Here this approach is extended to require that the relative magnitude of activation, as measured by the Z score accompanying each coordinate, is also consistent as measured by correlation of standardised Z scores between coactivated regions. The results from CBMAN are a network of coactivated (or co-altered) spatial coordinate clusters (nodes) indicating spatial consistency of reported activation, connected by edges where the standardised reported Z scores have consistent relative magnitude across study. This may generate interesting hypotheses that can be tested in well-designed prospective studies.

References

1. 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.
2. Wager, T.D., et al., *Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging.* Neuroimage, 2003. **19**(3): p. 513-31.
3. 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.
4. 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.
5. Eickhoff, S.B., et al., *Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty.* Hum Brain Mapp, 2009. **30**(9): p. 2907-26.
6. Eickhoff, S.B., et al., *Activation likelihood estimation meta-analysis revisited.* Neuroimage, 2012. **59**(3): p. 2349-61.
7. Tench, C.R., et al., *Coordinate based meta-analysis of functional neuroimaging data; false discovery control and diagnostics.* PLoS One, 2013. **8**(7): p. e70143.
8. 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.
9. Eickhoff, S.B., et al., *Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation.* Neuroimage, 2016. **137**: p. 70-85.
10. Tench, C., et al. *Coordinate based random effect size meta-analysis shows regions of GM atrophy do not develop independently in MS and CIS*. in *MULTIPLE SCLEROSIS JOURNAL*. 2017. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND.
11. Tench, C.R., et al., *Coordinate based random effect size meta-analysis of neuroimaging studies.* Neuroimage, 2017. **153**: p. 293-306.
12. Robinson, J.L., et al., *Metaanalytic connectivity modeling: delineating the functional connectivity of the human amygdala.* Hum Brain Mapp, 2010. **31**(2): p. 173-84.
13. Xue, W., et al., *Identifying functional co-activation patterns in neuroimaging studies via poisson graphical models.* Biometrics, 2014. **70**(4): p. 812-22.
14. Lancaster, J.L., et al., *Automated analysis of meta-analysis networks.* Hum Brain Mapp, 2005. **25**(1): p. 174-84.
15. Neumann, J., et al., *Meta-analysis of functional imaging data using replicator dynamics.* Hum Brain Mapp, 2005. **25**(1): p. 165-73.
16. Neumann, J., et al., *Learning partially directed functional networks from meta-analysis imaging data.* Neuroimage, 2010. **49**(2): p. 1372-84.
17. Cauda, F., et al., *The morphometric co-atrophy networking of schizophrenia, autistic and obsessive spectrum disorders.* Hum Brain Mapp, 2018. **39**(5): p. 1898-1928.
18. Tatu, K., et al., *How do morphological alterations caused by chronic pain distribute across the brain? A meta-analytic co-alteration study.* Neuroimage Clin, 2018. **18**: p. 15-30.
19. Chu, C., et al., *Co-activation Probability Estimation (CoPE): An approach for modeling functional co-activation architecture based on neuroimaging coordinates.* Neuroimage, 2015. **117**: p. 397-407.

20. Altman, D.G. and P. Royston, *The cost of dichotomising continuous variables.* BMJ : British Medical Journal, 2006. **332**(7549): p. 1080-1080.

21. van den Heuvel, M.P. and H.E. Hulshoff Pol, *Exploring the brain network: a review on resting-state fMRI functional connectivity.* Eur Neuropsychopharmacol, 2010. **20**(8): p. 519-34.

22. 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.

23. Alakorkko, T., et al., *Effects of spatial smoothing on functional brain networks.* Eur J Neurosci, 2017. **46**(9): p. 2471-2480.

24. Press, W.H., et al., *Numerical recipes in C (2nd ed.): the art of scientific computing.* 1992: Cambridge University Press. 994.

25. Zalesky, A., A. Fornito, and E. Bullmore, *On the use of correlation as a measure of network connectivity.* NeuroImage, 2012. **60**(4): p. 2096-2106.

26. Heap, B.R., *Permutations by Interchanges.* The Computer Journal, 1963. **6**(3): p. 293-298.

27. Mazziotta, J., et al., *A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM).* Philos Trans R Soc Lond B Biol Sci, 2001. **356**(1412): p. 1293-322.

28. Talairach, J. and P. Tournoux, *Co-planar stereotaxic atlas of the human brain.* 1988, New York: Thieme.

29. 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.

30. Team, R.D.C., *R: A Language and Environment for Statistical Computing.* 2008, R Foundation for Statistical Computing.

31. Csardi, G. and T. Nepusz, *The igraph software package for complex network research.* InterJournal, Complex Systems, 2006. **1695**(5): p. 1-9.

32. Ester, M., et al., *A density-based algorithm for discovering clusters in large spatial databases with noise.* Kdd, 1996. **96**(34): p. 226-231.

33. Albajes-Eizagirre, A., et al., *Voxel-based meta-analysis via permutation of subject images (PSI): Theory and implementation for SDM.* Neuroimage, 2018. **186**: p. 174-184.

34. Costafreda, S.G., *Parametric coordinate-based meta-analysis: valid effect size meta-analysis of studies with differing statistical thresholds.* J Neurosci Methods, 2012. **210**(2): p. 291-300.

35. 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.

36. Bowring, A., C. Maumet, and T.E. Nichols, *Exploring the impact of analysis software on task fMRI results.* Hum Brain Mapp, 2019.

37. 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.

38. Tramèr, M.R., et al., *Impact of covert duplicate publication on meta-analysis: a case study.* BMJ, 1997. **315**(7109): p. 635-640.

39. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.

40.   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.