
Bioimage informatics

Three-dimensional cardiovascular imaging-genetics: a mass univariate framework

Carlo Biffi^{1,2}, Antonio de Marvao², Mark I. Attard²,
Timothy J. W. Dawes^{2,3}, Nicola Whiffin^{2,3,4}, Wenjia Bai¹, Wenzhe Shi¹,
Catherine Francis^{3,4}, Hannah Meyer⁵, Rachel Buchan^{3,4},
Stuart A. Cook^{2,3,4,6,7}, Daniel Rueckert¹ and Declan P. O'Regan^{2,*}

¹Department of Computing, Imperial College London, South Kensington Campus, London, UK, ²Cardiovascular Magnetic Resonance Imaging and Genetics, MRC London Institute of Medical Sciences, Imperial College London, Hammersmith Hospital Campus, London, UK, ³Quantitative Physiology and Genetics, National Heart and Lung Institute, Imperial College London, London, UK, ⁴NIHR Cardiovascular Biomedical Research Unit, Royal Brompton and Harefield NHS Trust, London, UK, ⁵European Molecular Biology Laboratory (EMBL), European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK, ⁶Department of Cardiology, National Heart Centre Singapore, Singapore and ⁷Programme in Cardiovascular and Metabolic Disorders, Duke National University Singapore, Singapore

*To whom correspondence should be addressed.

Associate Editor: Robert Murphy

Received on May 5, 2017; revised on August 10, 2017; editorial decision on August 31, 2017; accepted on September 1, 2017

Abstract

Motivation: Left ventricular (LV) hypertrophy is a strong predictor of cardiovascular outcomes, but its genetic regulation remains largely unexplained. Conventional phenotyping relies on manual calculation of LV mass and wall thickness, but advanced cardiac image analysis presents an opportunity for high-throughput mapping of genotype-phenotype associations in three dimensions (3D).

Results: High-resolution cardiac magnetic resonance images were automatically segmented in 1124 healthy volunteers to create a 3D shape model of the heart. Mass univariate regression was used to plot a 3D effect-size map for the association between wall thickness and a set of predictors at each vertex in the mesh. The vertices where a significant effect exists were determined by applying threshold-free cluster enhancement to boost areas of signal with spatial contiguity. Experiments on simulated phenotypic signals and SNP replication show that this approach offers a substantial gain in statistical power for cardiac genotype-phenotype associations while providing good control of the false discovery rate. This framework models the effects of genetic variation throughout the heart and can be automatically applied to large population cohorts.

Availability and implementation: The proposed approach has been coded in an R package freely available at <https://doi.org/10.5281/zenodo.834610> together with the clinical data used in this work.

Contact: declan.oregan@imperial.ac.uk

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

One of the most complex unanswered questions in cardiovascular biology is how genetic and environmental factors influence the structure and function of the heart as a three-dimensional (3D)

structure (Li *et al.*, 2016). This is relevant for understanding the penetrance and expressivity of variants associated with inherited cardiac conditions as well as exploring the biology of heart development and within-population variation. Cardiac magnetic resonance

(CMR) is the gold-standard for quantitative imaging (Hundley et al., 2010), providing a rich source of anatomic and motion-based data, however conventional phenotyping relies on manual analysis reducing the variables of interest to global volumes and mass. Computational image analysis, by which machine learning is used to annotate and segment the images, is gaining traction as a means of representing detailed 3D phenotypic variation at thousands of vertices in a standardized coordinate space (Fig. 1) (Fonseca et al., 2011; Young and Frangi, 2009).

One approach to inference is to transform the spatially correlated data into a smaller number of uncorrelated principal components (Medrano-Gracia et al., 2015), however these modes would not provide an explicit model relating genotype to phenotype. A more powerful approach may be to derive a statistic expressing evidence of a given effect at each vertex of the 3D model, hence creating a so-called statistical parametric map—a concept widely used in functional neuroimaging (Friston, 2007). In this paper we extend techniques developed in the neuroscience domain to cardiovascular imaging-genetics by implementing a mass univariate framework to map associations between genetic variation and a 3D phenotype. Such an approach would provide overly conservative inferences without considering spatial dependencies in the underlying data and so we validated the translation of threshold-free cluster-enhancement (TFCE) to cardiovascular phenotypes for the sensitive detection of coherent signal (Smith and Nichols, 2009) as well as implementing robust control for multiple testing. The feasibility of the proposed methodology to derive computationally efficient inferences on imaging-genetics datasets has been tested through experiments on clinical and synthetic data using an R package developed for this purpose.

2 Materials and methods

2.1 Study population

The healthy volunteers dataset used in this study is part of the UK Digital Heart Project at Imperial College London (Bai et al., 2015) (see Supplementary Data S1 for full cohort characteristic and acquisition details). To capture the whole-heart phenotype, a high-spatial resolution 3D balanced steady-state free precession cine sequence was performed on a 1.5-T Philips Achieva system (Best, the Netherlands). Images were stored on an open-source database (MRIdb, Imperial College London, UK) (Woodbridge et al., 2013). Conventional volumetric analysis of the cine images was performed using CMRtools (Cardiovascular Imaging Solutions, London, UK) following a standard protocol (Schulz-Menger et al., 2013).

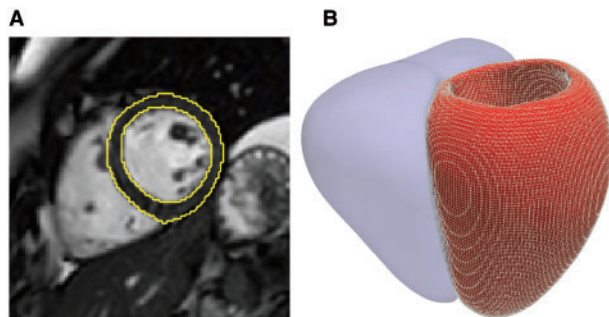


Fig. 1. Computational image analysis. **(A)** Short axis cardiac magnetic resonance image demonstrating automated segmentation of the endocardial and epicardial boundaries of the left ventricle. **(B)** The segmentation is used to construct a three dimensional mesh of the cardiac surfaces (left ventricle shown as a mesh, right ventricle shown as a solid) that is co-registered to a standard coordinate space. Phenotypic parameters, such as wall thickness, are then derived for each vertex in the model

Genotyping of common variants was carried out using an Illumina HumanOmniExpress-12v1-1 single nucleotide polymorphism (SNP) array (Sanger Institute, Cambridge). Clustering, calling and scoring of SNPs was performed using Illumina GenCall software. Samples were pre-phased with SHAPEIT (Delaneau et al., 2013) and imputation was performed using IMPUTE2 (Howie et al., 2009) with the UK10K dataset as a reference (www.uk10k.org). Quality of the genotypes was evaluated both on a per-individual and per-marker level using in-house Perl scripts. SNPs were removed if they had a Impute Information (INFO) score < 0.4 , missing call rate in more than 1% of samples, minor allele frequency of less than 1% or deviated significantly from Hardy-Weinberg equilibrium ($P > 0.001$). Only non-related individuals with ‘CEU’ ethnicity were retained. The total genotyping rate in these individuals was 0.997 and the total number of variants available was 9.4 million.

2.2 Atlas-based segmentation and co-registration

All image processing was performed with Matlab (MathWorks, Natick, Mass). A validated cardiac segmentation and co-registration framework was used which has previously been described in detail (Bai et al., 2015; de Marvao et al., 2015). A 3D shape model (at end-diastole and end-systole) was created encoding phenotypic variation in our study population at 49 876 epicardial vertices and visualized in a standard coordinate space (Fig. 1) (Bai et al., 2015). Wall thickness (WT) was measured by computing the distance between respective vertices on the endocardial and epicardial surfaces at end-diastole.

2.3 Overview of the approach

In the following sections we introduce a framework for deriving associations between clinical/genetic parameters and a 3D cardiac phenotype which is outlined in Figure 2. Briefly, a general linear model is fitted at each vertex to extract the regression coefficient associated with the variable of interest (mass univariate regression). Threshold-free cluster enhancement (TFCE) is then applied to boost belief in extended areas of coherent signal in the derived vertex-wise statistical map. The points where a significant effect exists are determined by applying TFCE on the obtained t-statistic map and on N t-statistic maps obtained through permutation testing, derived under the null hypothesis of no effect. Then, at each vertex the frequentist probability of having obtained a higher TFCE score by chance is regarded as the P -value related to the regression coefficient β . Finally, the derived P -values are adjusted for multiple testing. The permutation testing procedure employed by this approach is the Freedman-Lane procedure (Freedman and Lane, 1983), whilst a false discovery rate (FDR) correction using the Benjamini-Hochberg procedure (Benjamin and Hochberg, 1995) is applied to correct for multiple testing.

2.4 Mass univariate analysis

The association between a ventricular phenotype mapped onto a 3D mesh and one or more clinical variables can be described using a general linear model of the form $Y = \beta X + \epsilon$, where Y is a vector of, for example, WT values at each vertex and X is a design matrix that can be used to model the effect of interest and contains in each column the subject’s values of clinical co-variables as well as the intercept term. These variables can be numerical (such as age or weight), categorical or expressing interaction between them. In particular, categorical variables can be exploited to express either different categories (such as gender or ethnicity) or the presence/absence in a binary form of a clinical condition (such as the presence of genetic mutation or a specific disease). β is the regression coefficient vector to be estimated and ϵ represents the noise or error term, which is

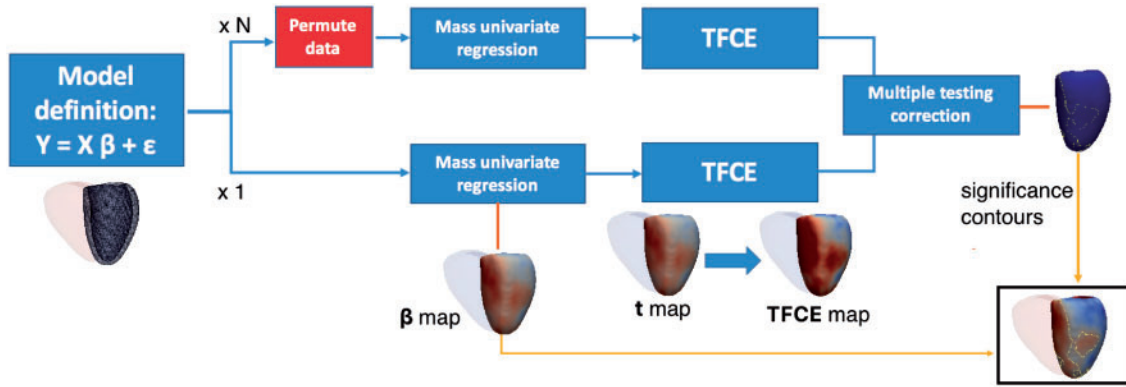


Fig. 2. Outline of three-dimensional mass univariate framework. A statistical atlas provides point-wise measures of ventricular geometry and function which can be linked to a given predictor through a general linear model. Using mass univariate regression, three-dimensional maps of a test statistic and the degree of association (β) can be derived. Threshold free cluster enhancement (TFCE) coupled with permutation testing produces vertex-wise P -values weighted to the degree of coherent spatial support. Finally, P -values are corrected for multiple testing. Regression coefficients enclosed by significance contours are represented on a model of the left ventricle

assumed to be a zero-mean Gaussian process and represents the variability of Y not explained by the model. The regression coefficient can be standardized by normalizing to mean 0 and unit-variance the columns of X and Y . As a result, β will represent the amount of variation of Y in units of its standard deviation when X is increased by one standard deviation, allowing comparisons between variables.

The same model can be fitted at each ventricular vertex independently (mass univariate regression) and statistics can be extracted and corrected for multiple testing in order to test one or more statistical hypotheses. In a parametric framework, the t -statistic computed as $t = \frac{\beta}{s.e.(\beta)}$ is typically used in the neuroimaging literature to contrast the null hypothesis $H_0 : \beta = 0$ (no association between the predictor and the phenotype under study), where $s.e.(\beta)$ is the standard error of the estimator of β (Friston, 2007). The regression coefficients β and their related P -value thus obtained can be plotted to display, at high resolution on the whole 3D ventricle, the magnitude and spatial distribution of a given association. However, this approach underestimates associations where the signal is more spatially correlated than noise coherence. For this reason non-parametric statistics such as TFCE are valuable to increase the statistical power of the approach.

2.5 Threshold-free cluster enhancement on a cardiac atlas

The value of a statistic h obtained through mass univariate regression at a vertex p —a t -statistic in our context—is transformed by TFCE using the following integral:

$$TFCE(p) = \int_{h=0}^{h_p} e(h)^E h^H \delta h \simeq \sum_{h=0}^{h_p} e(h)^E h^H \Delta h \quad (1)$$

In the equation h_p is the value of the vertex statistic, $e(h)$ is the extent of the cluster with cluster-forming threshold h that contains p , and E and H are two parameters usually set to 0.5 and 2 for empirical and analytical motivations (Smith and Nichols, 2009). In computational algorithms the integral is estimated using a discrete sum. The computational model of the heart is defined as a 3D mesh composed of non-congruent triangles where at each vertex pointwise phenotypic variables are stored for each subject. The translation of TFCE to a cardiac model is not straightforward as the model is not composed of a regular grid of voxels (as in brain imaging applications) but instead forms a mesh of vertices. We addressed this

problem by associating an area to each vertex i equivalent to the mesh area closest to that vertex. In computing Eq. 1 at each vertex, the most time consuming part is deriving $e(h)$ —the area of all the elements connected to p that have a statistic value greater or equal to h —as a different $e(h)$ needs to be computed for each vertex of the mesh and for each term of the sum. However, the TFCE score associated to a vertex p of a specific h in the summation consists of the same score that should be associated with all the vertices which contribute to $e(h)$. Therefore, the computational time of the TFCE method can be significantly reduced by sampling the interval between the maximum and minimum statistic h in the statistic map so as to use each sampled value as a threshold \hat{h} for the selection of vertices with a greater statistic value in the case of positive threshold, or less than \hat{h} in the case of a negative threshold. The edges of the graph are defined from a list containing the nearest neighbours of each vertex, and which is filtered at each iteration to contain only the vertices selected by the thresholding operation, resulting in one or more graphs of connected vertices. In this way, all the possible patterns of signal on the ventricle can be discovered without relying on assumptions about the geometry of cluster shapes. For all the obtained graphs including more than two vertices the TFCE score can be computed and associated to all the vertices that belong to them. The final TFCE score is the sum of all the TFCE scores thus obtained.

2.6 Permutation testing

The P -value associated with the regression coefficient computed at each atlas vertex can be derived via permutation testing. In particular, by permuting N times the input data, N TFCE scores maps can be obtained. It is important that the adopted permutation strategy guarantees the exchangeability assumption, *i.e.* permutations of Y given X do not alter the joint distribution of the dependent variables under the null hypothesis. In the proposed context, the Y values themselves are not exchangeable under the null hypothesis, as the predictors included in the model together with the variable under study are nuisance variables that could explain some portion of variability of Y . In order to address this problem, among a number of available techniques, the Freedman-Lane procedure (Freedman and Lane, 1983) has proved to provide the best control of statistical power and false positives (type 1 error) (Winkler *et al.*, 2014). This procedure proceeds as follows. If Z contains all the nuisance variables previously contained in X , the general linear model can be

rewritten as $Y = \beta X + \gamma Z + \epsilon$. Then, instead of permuting Y and extracting β , the procedure computes the residual-forming matrix $R_Z = 1 - XX^T$ and performs N different permutations by computing the model $\widehat{P}_N R_Z Y = \beta X + \gamma Z + \epsilon$ at each point, where \widehat{P}_N is the permutation operator. For a full derivation of the Freedman-Lane strategy see [Winkler et al. \(Winkler et al., 2014\)](#).

2.7 False discovery rate correction for multiple comparisons

A multiple testing problem arises by testing tens of thousands of statistical hypotheses simultaneously. Control of the family wise error rate at 5% could be derived by extracting the maximum score from each map derived via permutation testing and by using the 95th percentile as a threshold for significance. However, in this context such a correction could be overly conservative as we are rarely interested in the exact number of vertices that reach significance. The main goal is to detect extended areas of coherent signal and therefore we can accept a maximum fixed percentage of false discoveries as provided by false discovery rate (FDR) procedures. In particular, these procedures can be applied to adjust the voxelwise P -values obtained at each vertex by computing the ratio between the number of times in which a TFCE score greater than the measured one has been obtained and the number of permutations N . We have found adaptive procedures such as the two-stage Benjamini-Hochberg ([Benjamini et al., 2006](#)) not suitable for our dataset, since it led to lower P -values and increased areas of significance, as also reported in the neuroimaging literature ([Reiss et al., 2012](#)). For this reason, the original Benjamini-Hochberg (BH) ([Benjamin and Hochberg, 1995](#)) procedure has been employed for this work. It is important to underline that both FDR correction procedures are valid when the tested hypotheses are independent or satisfy a technical definition of dependence called positive regression dependency on a subset ([Benjamini and Yekutieli, 2001](#)). This condition for Gaussian data is translated into the requirement that the correlation between null voxels or between null and signal voxels is zero or positive, and for smoothed image data as those that compose a cardiac atlas, this assumption is generally considered satisfied ([Friston, 2007](#)).

2.8 Software

The proposed mass univariate framework has been coded as an R package (`mutools3D`) which benefits from the use of vectorized operations. Matrices containing the phenotypic data and templates to visualise the 3D models are also available with the software. Linear regression assumptions must be met in order to obtain reliable inferences (for a review of them and their importance in a mass univariate setting see [Supplementary Data S2](#)). Particularly important in this context are multicollinearity and heteroscedasticity problems which should be checked and solved for each model definition. For the latter, the R package implements mass univariate functions exploiting HC4m heteroscedasticity consistent estimators ([Cribari-Neto and da Silva, 2011](#)).

3 Results

3.1 GWAS replication study

As an exemplar application, six out of nine exonic SNPs which have previously shown an association with LV mass in a case-control genome wide association study (GWAS), using echocardiography for phenotyping ([Arnett et al., 2009](#)), were also identified in the UK Digital Heart Project genotypes and were assessed for replication.

For each SNP, WT at each vertex in the 3D model in 1124 healthy Caucasian subjects was tested for association with the posterior estimate of the allele frequency by a regression model adjusted for age, gender, body surface area (BSA) and systolic blood pressure (SBP). The tested SNPs are rs409045, rs6450415, rs1833534, rs6961069, rs10499859 and rs10483186 and cohort characteristics are reported in [Supplementary Data S5](#). Regression diagnosis through Breusch-Pagan and White's test showed how the homoscedasticity assumption was violated at a large number of vertices, therefore mass univariate regression was corrected using HC4m heteroscedasticity consistent estimators ([Cribari-Neto and da Silva, 2011](#)). Regarding the assumption of multicollinearity, the condition number of the model matrix was 2.19 while the variance inflation factor was equal to 1.06, suggesting a very low degree of multicollinearity. All the simulations were executed on a high performance computer (Intel Xeon Quad-Core Processor (30 M Cache, 2.40 GHz), 36 Gb RAM), using the analysis pipeline and R package proposed in this paper ([Fig. 2](#)). A multiple comparisons procedure correcting for the number of vertices and the number of SNPs tested was applied by simultaneously testing in a BH FDR-controlling procedure all the TFCE-derived P -values from all the models as suggested in ([Benjamini and Yekutieli, 2005](#)). The number of permutations was fixed to 10 000 and simulations required less than 3 h each. Finally, as a result of a preliminary study we conducted (full details in [Supplementary Data S3](#)), TFCE parameters E and H were set to 0.5 and 2, as suggested in the original TFCE paper ([Smith and Nichols, 2009](#)), since this choice provides good sensitivity and specificity on a range of synthetic signals.

Four SNPs showed a significant association with WT as reported in [Figure 3](#). These are rs409045 (maximum regression coefficient $\bar{\beta} = -0.1$, percentage of the LV area significant $S = 13\%$), rs6450415 ($\bar{\beta} = -0.11$, $S = 11\%$), rs6961069 ($\bar{\beta} = -0.09$, $S = 44\%$) and rs10499859 ($\bar{\beta} = 0.1$, $S = 41\%$). Conventional linear regression analysis using LV mass and the same model for all the SNPs did not reach significance ([Supplementary Data S5](#)).

3.2 Assessment of sensitivity, specificity and false discovery rate on synthetic data

Sensitivity, specificity and the rate of false discoveries of the proposed pipeline were estimated using synthetic data. A 3D model showing no correlation between WT and the posterior estimate of the allele frequency X_{snp} of a non-associated SNP (rs4288653) adjusted for age, gender, BSA and SBP was used to generate background noise. A synthetic data signal was generated by summing to the WT values of each subject a term $I \beta X_{snp}$ at each vertex, where I is the signal intensity and β is a map of regression coefficients. Two contrasting β maps (signal A and B) obtained from real clinical data were chosen and are available in [Supplementary Data S6](#). Signal A was characterized by non-null β coefficients covering the 10% of total area of the LV and scaled to the $(0, 1]$ range, while signal B presented non-null regression coefficients scaled to the $[-1, 0)$ range in a more extended area covering the 60% of the LV surface. By subsampling the number of subjects N and the signal intensity I , different signals to be detected by the proposed standard mass univariate pipeline were obtained. The number of permutations for each simulation was fixed to 5000 and results were linearly interpolated and plotted on the contour plots shown in [Figure 4](#).

Sensitivity increased at larger sample sizes N and signal intensities I , reaching the greatest values with the most extended signal (signal B) as expected. Given the sample size of our GWAS replication study and the intensity of the associations found, these results

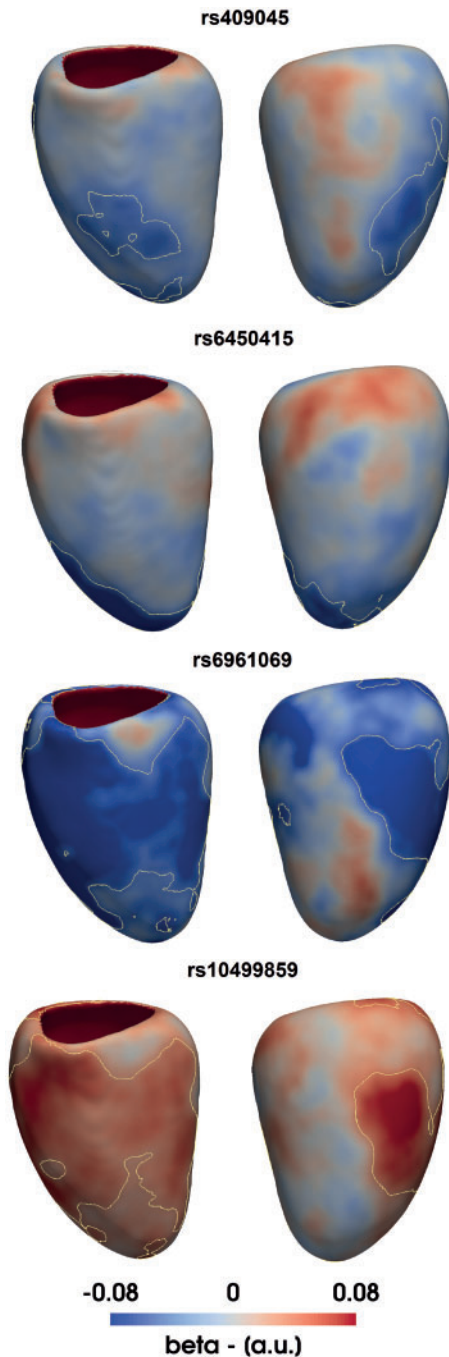


Fig. 3. Applying three-dimensional analysis to single nucleotide polymorphism (SNP) replication. β coefficients are plotted on the surface of the left ventricle for the effect of 4 distinct SNPs on wall thickness (WT) adjusted for age, gender, body surface area and systolic blood pressure. Yellow contours enclose standardized regression coefficients reaching significance after multiple testing

would assign a sensitivity of 70% for the first two discovered SNPs and more than 90% for the other two. Moreover, the rate of false discoveries was 0 for all the results of signal A and below 5% except for few simulations involving signal B and sample sizes greater than 1600 (results reported in [Supplementary Data S6](#)). This effect is due to the large synthetic signal extension, which causes TFCE to extend its support to vertices near the true signal which show the same direction of effect. This is not considered a major limitation as TFCE

will not enhance clusters that originate only from noise. Finally, sensitivity, specificity and the rate of false discoveries were also computed using our pipeline without TFCE—which showed how application of TFCE provides a relevant increase of up to 50% in sensitivity which only comes at the expense of a small decrease in specificity on large extended signals ([Supplementary Data S6](#)).

Finally, we have performed a comparative study between TFCE and a standard cluster-based thresholding method as reported in the original TFCE paper on brain imaging data ([Smith and Nichols, 2009](#)), which has been implemented in the proposed R package ([Supplementary Data S4](#)). Overall, in agreement with the neuroimaging literature, the sensitivity of the cluster-extent based thresholding method was lower than TFCE and proved to be very dependent on the cluster-forming threshold. Moreover, higher false discovery rates and lower specificities characterized cluster-extent based thresholding results in all cases when their sensitivity was comparable or greater than TFCE.

4 Discussion

The environmental and genetic determinants of cardiac physiology and function, especially in the earliest stages of disease, remain poorly characterized and morphological classification relies on one-dimensional metrics derived by manual image segmentation ([Khoury et al., 2010](#)). In contrast, computational cardiac analysis provides precise 3D quantification of shape and motion differences between disease groups and normal subjects ([Medrano-Gracia et al., 2013](#)). We have extended the application of these techniques by designing a general linear model framework that provides a powerful approach for modelling the relationship between phenotypic traits, genetic variation and environmental factors using high-fidelity 3D representations of the heart. By translating statistical parametric mapping techniques originally developed for brain mapping to the cardiovascular domain we exploit spatial dependencies in the data to identify coherent areas of biological effect in the myocardium. This framework also accounts for multiple testing correction at tens of thousands of vertices which is the main drawback of this class of techniques. In particular, the application of TFCE leads to a notable increase in power of the mass univariate approach at the expense of only a slight increase of the false discovery rate in large extended signals.

Genetic association studies using conventional 2D imaging leave much of the moderate heritability of LV mass unexplained ([Fox et al., 2013](#); [Semsarian et al., 2015](#); [Vasan et al., 2009](#)). One contribution may be the lack of phenotyping power of conventional imaging metrics, which require manual analysis and are insensitive to regional patterns of hypertrophy. Our simulations on synthetic data show that our approach has the power to detect anatomical regions associated with even small genetic effect sizes. In the reported exemplar application, we replicated the effect of four SNPs discovered in a GWAS for LV mass using a 3D WT phenotype with TFCE applied, while none of the SNPs replicated with conventional LV mass analysis. The genotype-phenotype associations that we report reflect that cardiac geometry is a complex phenotype with a highly polygenic architecture dependent on anatomical patterns of gene expression and spatially varying adaptations to haemodynamic conditions ([Srivastava and Olson, 2000](#); [Wild et al., 2017](#)).

One of the main limitations of the presented framework is that high-spatial resolution CMR is not available in all cohorts, although conventional two-dimensional images may be super-resolved to provide similar shape models ([Oktay et al., 2017](#)). A second limitation

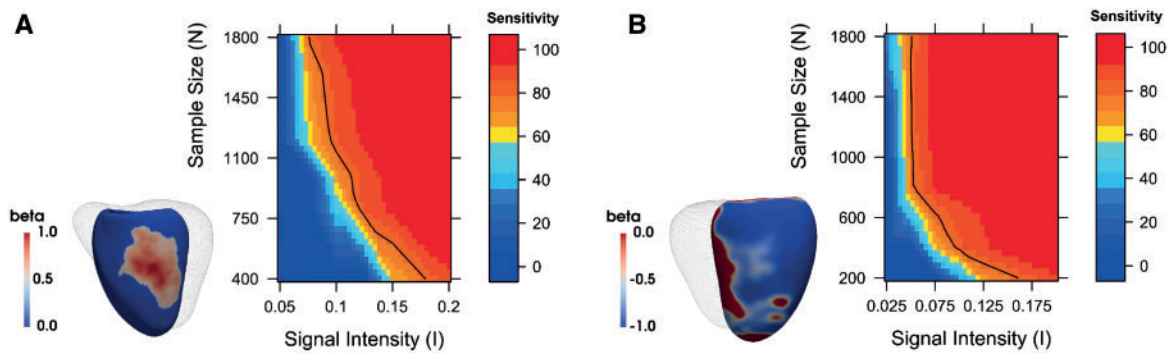


Fig. 4. Assessment of power using synthetic data. Plots of our framework's sensitivity at different sample sizes N and signal intensities I to detect a synthetic signal on (A) 10% and (B) 60% of the LV surface. A black line on the plots indicates a threshold of 80% sensitivity

is that the true association may not be linear in the model parameters and nonlinear models could better fit the data. However, the advantages favouring a general linear model are its simplicity, the ability to easily design and adjust the results for multiple factors and its wide use in biomedical statistics. A third limitation of this work is with regards to the experiments using synthetic data as we only assessed noise in our single centre population and did not generalize this to other cohorts. A general limitation of these approaches is that they do not establish causal relationships, such as the interaction between genetic variants, blood pressure and LV mass, although this may be addressed in future work by Mendelian randomization.

Mass univariate approaches do not directly consider the local covariance structure of the data, however this is accounted for when Random Field Theory or permutation tests define a threshold for significant activation (Bronzino and Peterson, 2015). In the neuroimaging literature, in the context of brain-wide candidate-SNP analyses, mass univariate approaches are used more extensively than multivariate approaches as the latter are less sensitive to regional effects and they require more observations than the dimension of the response variable (i.e. number of vertices) or the use of dimensionality reduction techniques (Friston, 2007).

As the methods are computationally efficient and require no human input for phenotypic analysis, it is feasible to scale up the pipeline to larger population cohorts such as UK Biobank, which aims to investigate up to 100 000 participants using MR imaging (Petersen et al., 2013). Applying these concepts to revealing the effect of rare variants on LV geometry in participants without overt cardiomyopathy (Schafer et al., 2017) and to vertex-wise genome-wide analyses also represent an interesting area of future work. In this latter context, multivariate approaches may show promise for modelling high-dimensional imaging and genetic data (Liu and Calhoun, 2014; Vounou et al., 2012). Finally, while we have focused on LV geometry and shape, the same approach can be applied to time-resolved vertex-wise data to create a functional phenotype for regression modelling.

5 Conclusion

We report a powerful and flexible framework for statistical parametric modelling of 3D cardiac atlases, encoding multiple phenotypic traits, which offers a substantial gain in power with robust inferences. We have implemented and validated the approach on both synthetic and clinical datasets, showing its suitability for detecting genotype-phenotype interactions on LV geometry. More generally, the proposed method can be applied to population-based

studies to increase our understanding of the physiological, genetic and environmental effects on cardiac structure and function.

Acknowledgements

The authors thank Dr. James Ware for advice on GWAS replication; our radiographers Ben Statton, Marina Quinlan and Alaine Berry; our research nurses Tamara Diamond and Laura Monje Garcia; and the study participants.

Funding

The study was supported by the Medical Research Council, UK; National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London, UK; British Heart Foundation grants PG/12/27/29489, NH/17/1/32725, SP/10/10/28431 and RE/13/4/30184; Academy of Medical Sciences Grant (SGL015/1006) and a Wellcome Trust-GSK Fellowship Grant (100211/Z/12/Z).

Conflict of Interest: none declared.

References

- Arnett, D.K. et al. (2009) Genome-wide association study identifies single-nucleotide polymorphism in *knob1* associated with left ventricular mass in humans: the hypergen study. *BMC Med. Genet.*, **10**, 43.
- Bai, W. et al. (2015) A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion. *Med. Image Anal.*, **26**, 133–145.
- Benjamin, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B*, **57**, 289–300.
- Benjamini, Y. et al. (2006) Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, **93**, 491–507.
- Benjamini, Y. and Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.*, **29**, 1165–1188.
- Benjamini, Y. and Yekutieli, D. (2005) Quantitative trait loci analysis using the false discovery rate. *Genetics*, **171**, 783–790.
- Bronzino, J.D. and Peterson, D.R. (2015) *Biomedical Signals, Imaging, and Informatics*. CRC Press, Taylor & Francis, Boca Raton, FL, USA.
- Cribari-Neto, F. and da Silva, W.B. (2011) A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Adv. Stat. Anal.*, **95**, 129–146.
- de Marvao, A. et al. (2015) Precursors of hypertensive heart phenotype develop in healthy adults: a high-resolution 3D MRI study. *JACC Cardiovasc. Imaging*, **8**, 1260–1269.
- Delaneau, O. et al. (2013) Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods*, **10**, 5–6.
- Fonseca, C.G. et al. (2011) The Cardiac Atlas Project—an imaging database for computational modeling and statistical atlases of the heart. *Bioinformatics*, **27**, 2288–2295.

- Fox,E.R. *et al.* (2013) Genome-wide association study of cardiac structure and systolic function in African Americans: the Candidate Gene Association Resource (CARE) study. *Circ. Cardiovasc. Genet.*, **6**, 37–46.
- Freedman,D. and Lane,D. (1983) A nonstochastic interpretation of reported significance levels. *J. Bus. Econ. Stat.*, **1**, 292–298.
- Friston,K.J. (2007) *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier/Academic Press, Amsterdam Boston.
- Howie,B.N. *et al.* (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.*, **5**, e1000529.
- Hundley,W.G. *et al.* (2010) ACCF/ACR/AHA/NASCI/SCMR 2010 expert consensus document on cardiovascular magnetic resonance: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents. *J. Am. Coll. Cardiol.*, **55**, 2614–2662.
- Khourri,M.G. *et al.* (2010) A 4-tiered classification of left ventricular hypertrophy based on left ventricular geometry the Dallas heart study. *Circ. Cardiovasc. Imaging*, **3**, 164–171.
- Li,G. *et al.* (2016) Transcriptomic profiling maps anatomically patterned subpopulations among single embryonic cardiac cells. *Dev. Cell*, **39**, 491–507.
- Liu,J. and Calhoun,V. (2014) A review of multivariate analyses in imaging genetics. *Front. Neuroinf.*, **8**, 29.
- Medrano-Gracia,P. *et al.* (2013) Atlas-based analysis of cardiac shape and function: correction of regional shape bias due to imaging protocol for population studies. *J. Cardiovasc. Magn. Reson.*, **15**, 80.
- Medrano-Gracia,P. *et al.* (2015) Challenges of cardiac image analysis in large-scale population-based studies. *Curr. Cardiol. Rep.*, **17**, 1–7.
- Oktay,O. *et al.* (2017) Anatomically constrained neural networks (ACNN): Application to cardiac image enhancement and segmentation. *arXiv preprint arXiv:1705.08302*.
- Petersen,S.E. *et al.* (2013) Imaging in population science: cardiovascular magnetic resonance in 100, 000 participants of UK Biobank-rationale, challenges and approaches. *J. Cardiovasc. Magn. Reson.*, **15**, 1.
- Reiss,P.T. *et al.* (2012) Paradoxical results of adaptive false discovery rate procedures in neuroimaging studies. *NeuroImage*, **63**, 1833–1840.
- Schafer,S. *et al.* (2017) Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat. Genet.*, **49**, 46–53.
- Schulz-Menger,J. *et al.* (2013) Standardized image interpretation and post processing in cardiovascular magnetic resonance: Society for Cardiovascular Magnetic Resonance (SCMR) board of trustees task force on standardized post processing. *J. Cardiovasc. Magn. Reson.*, **15**, 35.
- Semsarian,C. *et al.* (2015) New perspectives on the prevalence of hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.*, **65**, 1249–1254.
- Smith,S.M. and Nichols,T.E. (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, **44**, 83–98.
- Srivastava,D. and Olson,E.N. (2000) A genetic blueprint for cardiac development. *Nature*, **407**, 221.
- Vasan,R.S. *et al.* (2009) Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. *JAMA*, **302**, 168–178.
- Vounou,M. *et al.* (2012) Sparse reduced-rank regression detects genetic associations with voxel-wise longitudinal phenotypes in alzheimer’s disease. *Neuroimage*, **60**, 700–716.
- Wild,P.S. *et al.* (2017) Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. *J. Clin. Invest.*, **127**, 1798–1812.
- Winkler,A.M. *et al.* (2014) Permutation inference for the general linear model. *NeuroImage*, **92**, 381–397.
- Woodbridge,M. *et al.* (2013) MRIdb: medical image management for biobank research. *J. Digit. Imaging*, **26**, 886–890.
- Young,A.A. and Frangi,A.F. (2009) Computational cardiac atlases: from patient to population and back. *Exp. Physiol.*, **94**, 578–596.