NeuroImage 94 (2014) 287–302

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Fast and accurate modelling of longitudinal and repeated measures neuroimaging data

Bryan Guillaume [a,b,c], Xue Hua [d], Paul M. Thompson [d], Lourens Waldorp [e], Thomas E. Nichols [b,f,g,*], for the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium
[b] Department of Statistics, University of Warwick, Coventry, UK
[c] Global Imaging Unit, GlaxoSmithKline, Stevenage, UK
[d] Imaging Genetics Center, Laboratory of Neuro Imaging, Dept. of Neurology & Psychiatry, UCLA School of Medicine, Los Angeles, CA 90095, USA
[e] Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands
[f] Warwick Manufacturing Group, University of Warwick, Coventry, UK
[g] Oxford Centre for Functional MRI of the Brain, University of Oxford, Oxford, UK

ARTICLE INFO

Article history:
Accepted 10 March 2014
Available online 18 March 2014

Keywords:
Longitudinal Modelling
Sandwich Estimator
Marginal Modelling
ADNI

ABSTRACT

Despite the growing importance of longitudinal data in neuroimaging, the standard analysis methods make restrictive or unrealistic assumptions (e.g., assumption of Compound Symmetry—the state of all equal variances and equal correlations—or spatially homogeneous longitudinal correlations). While some new methods have been proposed to more accurately account for such data, these methods are based on iterative algorithms that are slow and failure-prone. In this article, we propose the use of the Sandwich Estimator method which first estimates the parameters of interest with a simple Ordinary Least Square model and second estimates variances/covariances with the "so-called" Sandwich Estimator (SwE) which accounts for the within-subject correlation existing in longitudinal data. Here, we introduce the SwE method in its classic form, and we review and propose several adjustments to improve its behaviour, specifically in small samples. We use intensive Monte Carlo simulations to compare all considered adjustments and isolate the best combination for neuroimaging data. We also compare the SwE method to other popular methods and demonstrate its strengths and weaknesses. Finally, we analyse a highly unbalanced longitudinal dataset from the Alzheimer's Disease Neuroimaging Initiative and demonstrate the flexibility of the SwE method to fit within- and between-subject effects in a single model. Software implementing this SwE method has been made freely available at http://warwick.ac.uk/tenichols/SwE.

Introduction

Longitudinal data analysis is of increasing importance in neuroimaging, particularly in structural and functional MRI studies. There were over 1000 publications in 2012 to mention "longitudinal fMRI", which is 3.9% of all "fMRI" 2012 publications and up from 1.5% in 2000.[2] Unfortunately, while the current versions of the two most widely used packages (i.e. SPM and FSL) are computationally efficient, when they model more than two time points per subject they must make quite restrictive assumptions. In particular, FSL v5.0 must assume Compound Symmetry, a simple covariance structure where the variances and correlations of the repeated measures are constant over time, and a fully balanced design. SPM12 unrealistically assumes a common longitudinal covariance structure for the whole brain. This motivates recent publications proposing methods to better model neuroimaging longitudinal data (Bernal-Rusiel et al., 2013a, 2013b; Chen et al., 2013; Li et al., 2013; Skup et al., 2012), however, all of these methods entail iterative optimisation at each voxel.

In neuroimaging, the two most widely longitudinal approaches currently used are the Naïve Ordinary Least Squares (N-OLS) modelling and the Summary Statistics Ordinary Least Squares (SS-OLS) modelling. The N-OLS method tries to account for the intra-visit correlations existing in the data by including subject indicator variables (i.e. an intercept per

* Corresponding author at: Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom.

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

[2] Based on Pubmed searches of "longitudinal AND fMRI" in all fields, versus just "fMRI". This is a crude measure, but does reflect the growing role of this type of study.

http://dx.doi.org/10.1016/j.neuroimage.2014.03.029

subject) in an OLS model. This approach is fast, but does not allow one to make valid inferences on pure between-subject covariates (e.g., group intercept or gender) and is valid only under a balanced design and Compound Symmetry (CS). The SS-OLS method proceeds by first extracting a summary statistic of interest for each subject (e.g., slope with time) and then uses a group OLS model to infer on the summary measures. This method is also fast and has the advantage of reducing the analysis of correlated data to an analysis of independent data, but this summary data may be highly variable as it is based on single-subject fits. In the context of one-sample t-tests, Mumford and Nichols (2009) showed that this approach is robust under heterogeneity, but warned that it is probably not the case for more general regression models.

In biostatistics, the analysis of longitudinal data is a long-standing problem and is generally performed by using either Linear Mixed Effects (LME) models or marginal models. The LME models include random effects which account for the intra-visit correlations existing in the data. Nevertheless, they require iterative algorithms which are generally slow and may fail to converge to a correct solution. Another issue with LME models is the complexity of specifying and fitting the model. For example, the random effects and the covariance structure of the error terms need to be specified (e.g., only random intercepts? Also random slopes?) and, unfortunately, a misspecification of those may lead to invalid results. These are particularly serious problems in neuroimaging as model assessment is difficult and a single model must be used for the whole brain. As a consequence, the use of LME models in neuroimaging may be prohibitively slow, and may lead to statistical images with missing or invalid results for some voxels in the brain. To limit the convergence issues, one may be tempted to use a LME model with only a random intercept per subject. Unfortunately, like the N-OLS model, this model assumes CS which is probably not realistic, especially for long studies carried out over years and with many visits. In contrast, the marginal modelling approach implicitly accounts for random effects, treats the intra-visit correlations as a nuisance and focuses the modelling only on population averages. They have appealing asymptotic properties, are robust against model misspecification and, as there are no explicit random effects, are easier to specify than LME models. However, they only focus on population-averaged inferences or predictions, typically require iterative algorithms and assume large samples.

Recently, Bernal-Rusiel et al. (2013a) proposed the use of LME models to analyse longitudinal neuroimaging data, but only on a small number of regions of interest or biomarkers, Chen et al. (2013) and Bernal-Rusiel et al. (2013b) extended the use of the LME models to mass-univariate settings. In particular, Bernal-Rusiel et al. (2013b) proposed the use of a spatiotemporal LME method based on a parcellation of the brain into homogeneous areas; in each area, they model the full spatiotemporal covariance structure by assuming a common temporal covariance structure across all the points and a simple spatial covariance structure. Skup et al. (2012) and Li et al. (2013) proposed to use marginal models to analyse neuroimaging longitudinal data. Specifically, Skup et al. (2012) proposed a Multiscale Adaptive Generalised Method of Moments (MA-GMM) approach which combines a spatial regularisation method with a marginal model called Generalised Methods of Moments (GMM; Hansen, 1982; Lai and Small, 2007) and Li et al. (2013) proposed a Multiscale Adaptive Generalised Estimating Equations (MA-GEE) approach which also combines a spatial regularisation method, but with a marginal model called Generalised Estimating Equations (GEE; Liang and Zeger, 1986). Thanks to their appealing theoretical asymptotic properties, the two latter methods seem very promising for analysing longitudinal neuroimaging data. Nevertheless, like the LME models, they require iterative algorithms, which make them slow, and – due to the fact that they rely on asymptotic theoretical results – their use may be problematic in small samples.

In this paper, we propose an alternative marginal approach. We use a simple OLS model for the marginal model (i.e. no subject indicator variables) to create estimates of the parameters of interest. For standard errors of these estimates, we use the so-called Sandwich Estimator (SwE; Eicker, 1963) to account for the repeated measures correlation. The main property of the SwE is that, under weak conditions, it is asymptotically robust against misspecification of the covariance model. In particular, this robustness allows us to combine the SwE with a simple OLS model which has no covariance model. Thus, this method is easy to specify and, with no need for iterative computations, is fast and has no convergence issues. Moreover, the proposed method can deal with unbalanced designs and heterogeneous variances across time and groups (or even subjects; more below on this). In addition, note that the SwE method can also be used for cross-sectional designs where repeated measures exist, such as fMRI studies where multiple contrasts of interests are jointly modelled, or even for family designs where subjects from the same family cannot be assumed independent. Nevertheless, like the MA-GMM and MA-GEE methods, the SwE method relies on asymptotic theoretical results, guaranteeing accurate inference only in large samples. Therefore, we also review and propose small sample adjustments that improve its behaviour in small samples.

The remainder of this paper is organised as follows. Starting from the LME model and its implied marginal model, we introduce the SwE method in its standard form. Then, we review and propose different adjustments to the SwE in order to improve its behaviour, mainly in the case of small samples. Finally, we assess the SwE method with intensive Monte Carlo simulations in a large range of settings and, more particularly, by analysing real brain images acquired as part of the Alzheimer's Disease Neuroimaging Initiative (ADNI; Mueller et al., 2005).

Methods

The Linear Mixed Effects model and the marginal model

Using the formulation of Laird and Ware (1982), the LME model for individual i is

$$y_i = X_i\beta + Z_i b_i + \epsilon_i \tag{1}$$

where y_i is a vector of n_i observations for individual $i = 1,2,\ldots,m$, β is a vector of p fixed effects which is linked to y_i by the $n_i \times p$ design matrix X_i, b_i is a vector of r individual random effects which is linked to y_i by the $n_i \times r$ design matrix Z_i, and ϵ_i is a vector of n_i individual error terms which is normally distributed with mean 0 and covariance Σ_i. The individual random effects b_i are also normally distributed, independently of ϵ_i, with mean 0 and covariance D. Typically, the p fixed effects might include an intercept per group, a linear effect of time per group, a quadratic effect of time per-group or per-visit measures effects like, in the case of Alzheimer's Disease, the MMSE (Mini-Mental State Examination) score. The r random effects usually include a "random intercept" for each subject (modelled by a constant in Z_i) and may also include a "random slope" for each subject.

Instead of posing a model for each subject consisting of (common) fixed and (individual) random components, we can fit a model with only fixed components and let the random components induce structure on the random error. This is the so-called marginal model, which, for subject i, has the form

$$y_i = X_i\beta + \epsilon_i^* \tag{2}$$

where the individual marginal error terms ϵ_i^* have mean 0 and covariance V_i. Typically, the covariance is taken to be unstructured, but if data arise as per the LME model specified above, then $V_i = \Sigma_i + Z_i D Z_i$. We will denote by X the grand design matrix, the $n \times p$ stacked matrix of the m X_i's, where $n = \sum_i n_i$ is the total number of observations.

In LME models, the randomness of the data is modelled by both the random effects b_i and the error terms ϵ_i. The random effects b_i have an important impact on the variance modelling and have to be chosen carefully. This makes LME models quite difficult to specify in practice. In contrast, in the marginal model, all the randomness is treated as a