NeuroImage 95 (2014) 136–150

Contents lists available at ScienceDirect

NeuroImage

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

Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: Comparing meta and megaanalytical approaches for data pooling

Peter Kochunov [a,*,1], Neda Jahanshad [b,1], Emma Sprooten [c], Thomas E. Nichols [d,e], René C. Mandl [f], Laura Almasy [g], Tom Booth [h], Rachel M. Brouwer [f], Joanne E. Curran [g], Greig I. de Zubicaray [i], Rali Dimitrova [j], Ravi Duggirala [g], Peter T. Fox [k], L. Elliot Hong [a], Bennett A. Landman [l], Hervé Lemaitre [m], Lorna M. Lopez [h], Nicholas G. Martin [o], Katie L. McMahon [p], Braxton D. Mitchell [q], Rene L. Olvera [r], Charles P. Peterson [g], John M. Starr [h,s], Jessika E. Sussmann [j], Arthur W. Toga [b], Joanna M. Wardlaw [h,n], Margaret J. Wright [o], Susan N. Wright [a], Mark E. Bastin [h,n], Andrew M. McIntosh [j], Dorret I. Boomsma [t], René S. Kahn [f], Anouk den Braber [t], Eco J.C. de Geus [t], Ian J. Deary [h], Hilleke E. Hulshoff Pol [f], Douglas E. Williamson [r], John Blangero [g], Dennis van 't Ent [t], Paul M. Thompson [b,u], David C. Glahn [c]

[a] Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA
[b] Imaging Genetics Center, Institute of Neuroimaging and Informatics, Department of Neurology, Keck School of Medicine of USC, Los Angeles, CA, USA
[c] Olin Neuropsychiatry Research Center in the Institute of Living, Yale University School of Medicine, New Haven, CT, USA
[d] Department of Statistics & Warwick Manufacturing Group, The University of Warwick, Coventry, UK
[e] Oxford Centre for Functional MRI of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, Oxford University, UK
[f] Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, Utrecht, The Netherlands
[g] Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA
[h] Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, The University of Edinburgh, Edinburgh, UK
[i] School of Psychology, University of Queensland, Brisbane, Australia
[j] Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK
[k] Research Imaging Institute, University of Texas Health Science Center San Antonio, San Antonio, TX, USA
[l] Department of Electrical Engineering, Vanderbilt University, Nashville, TN, USA
[m] U1000 Research Unit Neuroimaging and Psychiatry, INSERM-CEA-Faculté de Médecine Paris-Sud, Orsay, France
[n] Brain Research Imaging Centre, Centre for Clinical Brain Sciences, University of Edinburgh, UK
[o] QIMR Berghofer Medical Research Institute, Brisbane, Australia
[p] Centre for Advanced Imaging, University of Queensland, Brisbane, Australia
[q] Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA
[r] Department of Psychiatry, University of Texas Health Science Center San Antonio, San Antonio, TX, USA
[s] Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK
[t] Department of Biological Psychology, VU University, Amsterdam, The Netherlands
[u] Department of Neurology, Pediatrics, Engineering, Psychiatry, Radiology, & Ophthalmology, University of Southern California, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Accepted 4 March 2014
Available online 18 March 2014

Keywords:
Diffusion tensor imaging (DTI)
Imaging genetics
Heritability
Meta-analysis
Multi-site
Reliability

ABSTRACT

Combining datasets across independent studies can boost statistical power by increasing the numbers of observations and can achieve more accurate estimates of effect sizes. This is especially important for genetic studies where a large number of observations are required to obtain sufficient power to detect and replicate genetic effects. There is a need to develop and evaluate methods for joint-analytical analyses of rich datasets collected in imaging genetics studies. The ENIGMA-DTI consortium is developing and evaluating approaches for obtaining pooled estimates of heritability through meta-and mega-genetic analytical approaches, to estimate the general additive genetic contributions to the intersubject variance in fractional anisotropy (FA) measured from diffusion tensor imaging (DTI). We used the ENIGMA-DTI data harmonization protocol for uniform processing of DTI data from multiple sites. We evaluated this protocol in five family-based cohorts providing data from a total of 2248 children and adults (ages: 9–85) collected with various imaging protocols. We used the imaging genetics analysis tool, SOLAR-Eclipse, to combine twin and family data from Dutch, Australian and Mexican-American cohorts into

* Corresponding author.
E-mail address: pkochunov@mprc.umaryland.edu (P. Kochunov).
[1] PK and NJ contributed equally to this work.

http://dx.doi.org/10.1016/j.neuroimage.2014.03.033

one large "mega-family". We showed that heritability estimates may vary from one cohort to another. We used two meta-analytical (the sample-size and standard-error weighted) approaches and a mega-genetic analysis to calculate heritability estimates across-population. We performed leave-one-out analysis of the joint estimates of heritability, removing a different cohort each time to understand the estimate variability. Overall, meta- and mega-genetic analyses of heritability produced robust estimates of heritability.

Introduction

Human brain mapping studies have shown substantial advantages of pooling data across multiple studies (Van Horn et al., 2004). Genetic analyses, particularly genome-wide association studies (GWAS), tend to be limited in statistical power as there is typically a small (<0.5%, Flint and Munafò, 2013) contribution to complex phenotypic variability from individual, common genetic variants. This limitation is especially problematic for imaging genetic studies of human brain. The structure and function of the human brain is greatly influenced by genetics, but the proportion of the variance due to individual differences in the human genome depends on the brain structure and measure assessed (Kochunov et al., 2009, 2010).

A large number of neuroimaging traits with ever-increasing spatial resolution are becoming increasingly available to describe the regional complexity of brain variability. This presents a daunting challenge where the number of degrees of freedom, in both the imaging and genetics, can be overwhelming for any single imaging-genetic study. Therefore, data pooling strategies are crucial whereby data from multiple large imaging genetics studies can be analyzed together.

Imaging and genetics have both greatly advanced neuroscience in recent years. The two fields have developed in parallel but in the last decade, there was a push to merge them to fully capitalize on their power leading to the development of the new field of imaging genetics. This field emerged as a variation of classical genetic analyses that related diagnostic, clinical and/or behavioral measures to locations and specific variants in the genome. This new field is thought to be able to provide new approaches to characterize, treat and potentially prevent some brain-related disorders. Insight into biological mechanisms that predispose individuals to these types of illnesses holds the promise of yielding potential new therapies and a significant reduction of this considerable burden. Advantages of imaging genetics include the presumed greater biological proximity to genetic variation and the quantitative nature of imaging phenotypes, which ideally suited for partitioning phenotypic variance into variance explained by genetic and environmental factors. Therefore, the statistical power of genetic analysis depends on both the closeness of a phenotype to the action of the gene and the precision of the measurements. Modern MRI offers phenotypic measurements that may provide a more detailed description of the disorder than clinical symptoms or neuropsychological assessments, and many of these measures have high precision and reproducibility. Our experience and that of others indicate that the inter-session, scan-rescan variability of many common imaging measurements can be low, in the range of 1–5% (Agartz et al., 2001; Julin et al., 1997; Kochunov and Duff Davis, 2009; Kochunov et al., 2012b; Lemieux et al., 1999; Lerch and Evans, 2005). Therefore, the imaging genetics approach may help ascertain effects of specific genetic variants on the human brain and may also discover genetic variants associated with neurological or psychiatric illnesses (Braskie et al., 2011; Chen et al., 2012; Glahn et al., 2007, 2010; Hasler and Northoff, 2011; Thompson et al., 2001; van den Heuvel et al., 2013).

In imaging genetic studies, up to a million voxel-based imaging traits may be analyzed. The required correction for multiple comparisons may limit the statistical power for gene discovery, even in the largest individual imaging studies of hundreds or even thousands of subjects. One solution is collaborative data sharing and pooling through consortia such as Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) Consortium (http://enigma.ini.usc.edu). Recent examples highlight the potential of large, meta-analyses of genome-wide association studies (GWAS) to uncover genetic loci that are reliably and consistently associated with MRI-based phenotypes in worldwide datasets, including hippocampal volumes (Bis et al., 2012; Stein et al., 2012), intracranial volumes (Ikram et al., 2012; Stein et al., 2012), and head circumference (Taal et al., 2012). Recently, the ENIGMA-DTI Consortium working group was organized to develop methods to facilitate multi-site approaches to study genetic influences on white matter micro-architecture and integrity, assessed using diffusion tensor imaging (DTI). Here we specifically focus on the fractional anisotropy (FA) as it is the most commonly analyzed scalar parameter extracted from DTI (Basser and Pierpaoli, 1996; Basser et al., 1994). The absolute FA values are sensitive to fiber coherence, myelination levels, and axonal integrity, and have been widely used as an index of white matter integrity (Thomason and Thompson, 2011). FA has emerged as a sensitive index of diffuse abnormalities in many brain disorders including Alzheimer's disease (AD) (Clerx et al., 2012; Teipel et al., 2012); in many studies, it is related to cognitive performance (Penke et al., 2010a, 2010b) and is altered in numerous psychiatric disorders including major depressive disorder (Carballedo et al., 2012) and bipolar disorder (Barysheva et al., 2012; Sprooten et al., 2011). Patient-control differences in FA values are also among the most replicable and consistent neuroimaging findings in schizophrenia (Alba-Ferrara and de Erausquin, 2013; Friedman et al., 2008; Kochunov et al., 2012a; Mandl et al., 2012; Nazeri et al., 2013; Perez-Iglesias et al., 2011).

The goal of the ENIGMA-DTI Working Group is to develop generalizable analyses, methods, and techniques for extraction and combined genetic analysis of phenotypes from DTI data collected from imaging groups around the world, regardless of the imaging acquisition or specific population under study. Its overall aim is to discover genetic factors influencing or related to white matter architecture. The first step towards this goal was the development of homogenization protocols to reliably extract phenotypic measurements from data collected with different imaging equipment and parameters (Jahanshad et al., 2013). The next step is to evaluate different statistical approaches for data pooling and specifically compare meta and megaanalytical techniques to choose one approach that yields the greatest improvements in the power of genetic discovery while accommodating for potential for genetic diversity among samples. Two specific advantages of data pooling are the increased power for genetic discovery and the genetic diversity of the population sample. Data pooling makes it easier to identify genetic variants that exert only small individual effects (Zuk et al., 2012). However, pooling data may be confounded by variations in data acquisition across datasets, heterogeneities in study population, and other factors. Another limitation to data pooling arises from restrictions that can arise with sharing raw data (both ethical and regulatory), including either phenotypic or genetic information.

Here, we tested and compared the outcomes of three approaches to pool imaging genetic data from five separate cohorts worldwide, that had used various imaging acquisition parameters and population structures for analysis of heritability of the DTI–FA. Of course the ultimate goal is to detect specific variants on the genome that