

RESEARCH

Open Access



# ENIGMA-Viewer: interactive visualization strategies for conveying effect sizes in meta-analysis

Guohao Zhang<sup>1</sup>, Peter Kochunov<sup>2\*</sup>, Elliot Hong<sup>2</sup>, Sinead Kelly<sup>4,5</sup>, Christopher Whelan<sup>6,7</sup>, Neda Jahanshad<sup>3</sup>, Paul Thompson<sup>3</sup> and Jian Chen<sup>1\*</sup>

## Abstract

**Background:** Global scale brain research collaborations such as the ENIGMA (Enhancing Neuro Imaging Genetics through Meta Analysis) consortium are beginning to collect data in large quantity and to conduct meta-analyses using uniformed protocols. It becomes strategically important that the results can be communicated among brain scientists effectively. Traditional graphs and charts failed to convey the complex shapes of brain structures which are essential to the understanding of the result statistics from the analyses. These problems could be addressed using interactive visualization strategies that can link those statistics with brain structures in order to provide a better interface to understand brain research results.

**Results:** We present ENIGMA-Viewer, an interactive web-based visualization tool for brain scientists to compare statistics such as effect sizes from meta-analysis results on standardized ROIs (regions-of-interest) across multiple studies. The tool incorporates visualization design principles such as focus+context and visual data fusion to enable users to better understand the statistics on brain structures. To demonstrate the usability of the tool, three examples using recent research data are discussed via case studies.

**Conclusions:** ENIGMA-Viewer supports presentations and communications of brain research results through effective visualization designs. By linking visualizations of both statistics and structures, users can gain more insights into the presented data that are otherwise difficult to obtain. ENIGMA-Viewer is an open-source tool, the source code and sample data are publicly accessible through the NITRC website ([http://www.nitrc.org/projects/enigmaviewer\\_20](http://www.nitrc.org/projects/enigmaviewer_20)). The tool can also be directly accessed online (<http://enigma-viewer.org>).

**Keywords:** Interactive visualization, Meta-analysis, Effect size, Diffusion tensor imaging, Comparative studies

## Background

Large scale harmonization of image processing protocols across different studies around the world and the extraction of effect sizes across reliably extracted regions of interest, allows for a common framework through which results can be compared, and combined through unbiased meta-analyses as performed in the ENIGMA (Enhancing

Neuro Imaging Genetics through Meta-Analysis) consortium [1]. These advancements offer essential opportunities for brain scientists to produce credible findings through meta-analysis, a method that combines data cohorts collected worldwide to obtain the statistical power otherwise unavailable from a single cohort, in order to find cross-modality data associations that influence brain structures [2]. Cohort studies boost power to detect associations. Seminal accomplishments with promising results in imaging-genomics associations have accelerated scientific discoveries in areas such as schizophrenia [3], bipolar disorders [4], and other neurodegenerative diseases [5].

\*Correspondence: [pkochunov@mprc.umaryland.edu](mailto:pkochunov@mprc.umaryland.edu); [jichen@umbc.edu](mailto:jichen@umbc.edu)

<sup>2</sup> Maryland Psychiatric Research Center, University of Maryland, Baltimore, 55 Wade Ave, 21228 Baltimore, MD, US

<sup>1</sup> Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, 1000 Hilltop Circle, 21250 Baltimore, MD, US  
Full list of author information is available at the end of the article

In meta-analysis, one important task is to interpret effect size, a statistical measure that can be broadly defined as any statistic that quantifies the degree to which sample results diverge from the expectations in the null hypothesis. Computing effect size is important because if effect sizes are stable across studies or even generalizable over some variations in design or analysis, the results are replicable. That is, effect size is a statistical tool for meta-analysis that quantitatively synthesizes effects across different studies. Ranking brain measures in order of their effect sizes for case-control differences can unearth brain measures on the basis of both the stability of the brain volume measures (so-called heritability [6]) and their relevance in the disease being studied [7].

Comparing effect sizes is, however, a multi-variant issue, not only because scientists must choose studies carefully to ensure consistency of protocol use, but also because the variety of cohorts has made it possible to dig more deeply into and disentangle the sources (medication-related geographical or demographics and genetic factors, e.g. [8]) of variations that could explain why brain differences vary across studies and different phenotypes.

As new analytical results are produced and lead to increased data dimensionality and size, the bottleneck to human understanding is not only limited to data mining and computational approaches, but also to human limited memory capacity. Presenting and interpreting effect sizes and locating regions across studies can be obscured due to the complexity of interpreting the multivariate information space and the problems inherent in presenting rich datasets on a two-dimensional (2D) computer screen. Synthesizing new information for new discoveries and comparison with past results is cognitively demanding. The bandwidth of discovery will be bounded by the characteristics of human perception, and hence the quest for visualization has commenced in the brain sciences, as evidenced by recent reviews and by research on the vital role of visualization in the analysis of multimodal neuroimaging data [9–11].

Our long-term goals include making analytics results derived from the ENIGMA pipeline accessible to the neuroscience community at large and assisting brain scientists in seeing patterns in massive multimodal computational solutions, as well as encouraging effective communication and collaborative activities through visual means to convey our results to the general public. Here we present ENIGMA-Viewer (Fig. 1), an interactive visualization tool to let users explore multimodality brain data to compare effect sizes and associated brain anatomical structures and genomics factors. This work makes several contributions:

- A series of design strategies for spatial and non-spatial data integration in the context of meta-analysis of brain imaging and genetics.

- Consideration of the brain science domain and tools to aid multivariate comparison studies.
- Priority in integrating different imaging modalities to compare results and locate important information.

### Related work

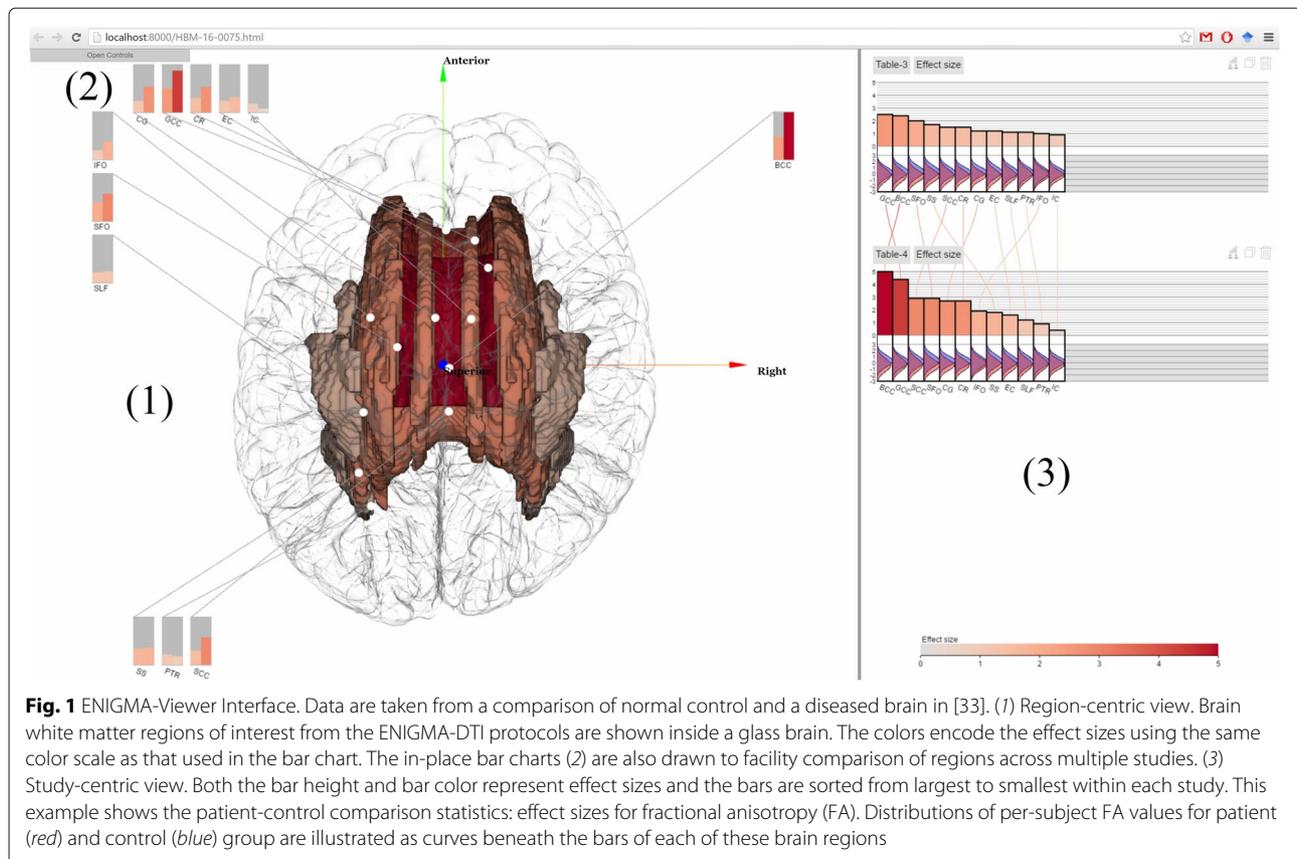
Both neuroscience and visualization scientists have worked extensively on visualizing brain datasets. This section reviews related work in visualization, related data analysis, and multi-modality data visualization.

### Brain data visualization

Many brain data visualization tools have addressed important issues in conveying single modality imaging techniques. In diffusion tensor magnetic resonance imaging (DTI) data visualization, Laidlaw et al. designed multivariate tensor field visualization at every voxel using creative artistic rendering [12]. Other powerful techniques have used tensor glyphs to convey tensor shapes [13], or non-photorealistic rendering to resolve complex spatial depth perception [14] as well as validating studies in the large display uses [15] and rendering solutions [16, 17]. Functional brain network (fMRI) visualizations have showed bundling 3D trajectories can support functional network understanding in both 3D [18] or 2D connectivity studies in matrix views [19].

Despite these creative solutions and technical advances, none of the work to our knowledge has exploited features and interpreting results across multiple modalities and multiple datasets, except our own work by Novak et al. [10] and Zhang et al. [20]. A main difference between single and cohort analyses is that single images become unimportant and statistical results comparing cohorts can lead to valuable understanding of associations between brain regions and diseases. Kehler et al. have laid out important design challenges in multi-modality multi-faceted data visualization in the broad medical imaging areas related to comparative studies as well as possible solutions in the use of multiview visualization to represent multidimensional data [21]. Our current work follows the multiview solutions to let scientists visually synthesize results from different views.

Other work most closely related to ours is Novak et al.'s EnigmaVis. That work lets scientists make quick comparisons among new and existing DTI-GWAS (genome-wide association study) queries through a powerful web interface [10]. This pioneering study is significant because it supports quick hypothesis confirmation through comparisons and lets brain scientists explore studies and examines results before their own study is conducted. However, that tool generates fixed images and only supports limited interactivity. Our design advances visual exploration by supporting interactive data exploration



especially not only for presenting and combining different imaging and measurements results but also for comparative visualization between modalities. Our design is very different from that of EnigmaVis in that no prior hypotheses or knowledge of prior studies is required to explore the prior studies in an interactive environment. Brain scientists can load and compare their data. We believe our solution can have great potential to support opportunistic discovery and may enable scientists to more easily and interactively investigate broader scientific questions.

#### Integrating spatial and non-spatial data

Our solution to comparative effect sizes is related to spatial and non-spatial data integration to assist data analysis. Our choices of visualization is mostly driven by data types, which is similar to the design rationale in Keefe et al. [22], where they visualize quantitative parameters using non-spatial data visualization to avoid inaccurate judgment of three-dimensional measurement. Wang and Tao also defines the integration of spatial and non-spatial data visualization [23], as well as Chen, Pyla, and Bowman in three-dimensional interface design [24].

#### Scientific background and data source

This section describes the background that motivates our visualization design, followed by description of the data used in the visualization.

#### Introduction to the goals in ENIGMA DTI-GWAS data analysis

The ENIGMA consortium aims to enable image-genetics discoveries by examining reproducibility, heritability, and association with diseases through analyzing brain imaging measures and genotypes [1]. The goal is to address the most fundamental questions in neuroscience by linking brain brain measures to human well being. Some of the most intriguing questions include: *what are the effects of aging, degenerative disease and psychiatric illness on the living brain? How do brain measures relate to cognition and behavior? Do brain measures predict our risk for disease, or give prognoses for those who are ill?* [1]

The method is meta-analysis, a quantitative statistical analysis of several separate but similar experiments or studies using pre-agreed covariates in order to test the pooled data and examine the effectiveness of the results [1]. Subsequently, the  $p$ -values and regression coefficients

are combined by weighting the results based on the sample size of each contributing cohort. Meta-analysis is not only important for brain white-matter analysis, but it has been the only way to find credible genetic traits of brain disorders with sufficient statistical power to achieve significant effects greater than  $p < 10^{-8}$ .

Great advancements in related fields have laid foundations for making cohort-comparison possible by addressing challenging technical problems in multiple areas. These include creating common ENIGMA template [2], harmonization of protocols to synthesize data captured with different protocols [25], generating tract-based spatial statistics skeletonization [26], regions of interests (ROI) extraction [27], and SOLAR statistics [28]. Meta-analyses have also identified the stabilities of brain volume measures (or heritability) in sub-cortical (containing regions associated with human function) and cortical regions across twins, genders, and geolocations [29].

A common workflow in performing meta-analysis is first to follow pre-determined protocols to obtain desirable imaging modalities (here DTI) and genomics in the population under investigation. Tract or voxel-based analyses and associated metrics measures (e.g., fractional anisotropy (FA) or water diffusion and cortical thickness) sensitive to the neuro-degenerations are then derived. Effect sizes in DTI studies are quantitatively compared.

#### Data

**Brain imaging data** The 3D brain imaging dataset labeled total 48 white matter structures in the JHU white matter atlas [30]. The brain volume in this atlas has  $182 \times 218 \times 182$  voxels measured at the resolution of  $1 \times 1 \times 1$  millimeters. We extract the surface mesh for each white matter region from the atlas using marching-cubes [31]. For cortical regions, we use cortical meshes from FreeSurfer. Since the FreeSurfer and the JHU atlases are different, the 70 FreeSurfer cortical regions are transformed to by matching the atlases using the linear transformation function in FSL [32].

**Statistical analysis data** The statistical data used in the program are from recent studies from ENIGMA group. For these analyses, effect sizes are reported as overall Cohen's *d* values for case/control effects and some studies also report Z-scores for quantitative effects (such as FA values for white matter studies) from linear regressions of individual subjects. An example data is from the study on the heterochronicity of white matter development and effect of aging in schizophrenia [33]. That study computes effect size values for 12 affected brain white matter regions contained in the JHU atlas.

#### Task analysis

The first goal of this study is to characterize the problems being addressed by the brain scientists as visualization tasks.

#### Procedure

The task analysis was achieved by working closely with brain scientists, as well as by literature review. Each scientist was interviewed to gather sufficient information on their workflow tasks and goals. Each participant also used our prototype tool of ENIGMA-Viewer and suggested action steps and desirable outcomes. To collect the resulting feedback, we have asked them to answer the following questions: What kinds of questions do you anticipate exploring using the visualization tool? What would you like to achieve using visualization in general, communication or seeing patterns? Why do the state-of-the-art tools, such as AFNI [34], FSL [35], DtiStudio [36], not address your needs? How would you like the data to be depicted and represented? Should the data be visualized in 2D or 3D? How would you like to interact with and explore the datasets?

#### Task list

Neuroscientists are interested in detecting trends and viewing overall data distribution as well as individual regions of interest. The most important tasks are related to (1) comparing similarities and differences in different disorders or in disease and control conditions; (2) comparing effect sizes in meta-analysis to find the truly significant brain regions and associated genetics factors; (3) studying the most important genetics association with these brain regions to establish the DTI-GWAS association; (4) identifying brain regions with high and low heritability. Each of these domain tasks can be abstracted to the fundamental analytic tasks presented by Amar, Eagan, and Stasko [37] and Schulz et al. [38], as listed in Table 1.

#### Methods

This section presents our main contribution, i.e., the design decisions made in the ENIGMA-Viewer to address all those users' tasks.

#### Overview of the design considerations

The possibilities for encoding and interacting with the data mentioned in Related work section are vast. Our encodings and layout draw upon existing idioms, and our task framework suggests that more novelty is required. We investigate visual design options through our experience of working on interface layout, discussion among the team of co-authors and following good design principles. We have also designed interaction techniques so that results from one data type and modality can guide comparative analysis of another in a unified interface level.

**Table 1** DTI-GWAS Task List. Our ENIGMA-Viewer is designed for these analytical tasks

Index	Task	Example questions
1	Viewing distribution, consistency, or inconsistency	What is the most affected brain region under a certain disorder? What is the overall data distribution of the effect size in all studies? Are the results consistent or very different?
2	Detect trends	What is the distribution of cortical thicknesses and FA values? How does effect size vary among studies?
3	Find association	What are the disorder (brain regions) and genetics correlates to risk? How does brain structural change associated with behavioral risks and changes and vulnerability? How do geographical factors affect disease expression in the brain?
4	Locate extremes	What are the significant brain regions mostly affected by diseases?
5	Find local relationships	What are the differences between studies in terms of their effect size?
6	Compare different disorders	What are the common and different effects in the brain networks between or among multiple disorders?
7	Compare disorders by regions	What are the regions affected or unique to a certain disorder?

Data belonging to different types can be visually linked through interaction.

We use *juxtaposition*, which places effect sizes, 3D anatomical regions, and artificial GWAS side-by-side in small-multiples displays similar to that of Chen et al. [39]. We also use *superposition* which the effect sizes and 3D anatomical regions are overlaid in the same frame of reference, following the comparative visualization classification by Gleicher et al. [40] and Karnick et al. [41].

#### Visual data fusion

Visual data fusion intermixes different facets of scientific data in a single view using a common frame of reference. In our program, effect sizes in different study cohorts and 3D anatomical regions can be grouped and presented in the space of the 3D glass brain. This visualization addresses scenarios of use in which a brain scientist wants to focus on finding associations of effect sizes in one or more regions of interests. For example, a brain scientist can load new and existing studies and then inspect trends and differences among studies visually. Another example use is to study multiple closely proximate brain regions of cortical and sub-cortical regions. Data from these spatial locations and multiple effect sizes can be discussed together. When the brain scientists' task is to search for associative relationships between different studies in a common region of interest, this visual fusion would be appropriate to let the user focus simply on one view to obtain all information.

#### Focus+context

Focus+context visualization supports both focused and detailed views as well as context for navigation purpose.

Effect sizes of each cohort are displayed in small multiples using bar charts ordered by effect size magnitudes. Since the effect sizes vary across studies, using uniform-scale bar charts would render smaller effect sizes too small to be distinguishable visually. Our solution is to color the magnitudes of the effect sizes. This strategy introduces dual encoding to encode the magnitudes of effect sizes: the bars use length with the most precise magnitude discrimination, while colors encourage pattern finding to locate extreme effect size magnitudes in different cohorts. The diverging color map is perceptually linear and the zero mark appears where the two colors intersect at 0 to represent the least significant effect size. Positive effect sizes are mapped to red and negative effect sizes are mapped to blue. In this way, users can obtain at a glance the most significant brain regions by searching for the most saturated red or blue regions. We plot the FA distributions between the patient and control cohorts in order to show the FA differences. Here we can observe that the control cohort has higher average FA values than the diseased ones in all regions.

#### Reduce context switching cost

The cost of context switching in visualization [42, 43] is one drawback of the small-multiples display in which bar charts are placed side-by-side. For searching for association between studies, the viewer must constantly switch the viewpoint between studies to look for relevant information in other views. To reduce the cost, our current method is to use "information scent" [44, 45], nuances added to the display to help the user construct visual associations. The edges are "scented" using the color representing the effect size magnitude

in the neighboring view so that the user need not visually trace the edge to learn the magnitude in the other view (Fig. 2).

The second way to reduce the context switching cost and to facilitate comparison of common regions is to use the stacked bar chart (Fig. 3). Effect sizes in the same brain region belonging to different cohorts are stacked together and horizontally, the cohorts are ordered by the effect size magnitudes in the bottom cohort. This view facilitates both between and within effect sizes of the same and different brain regions and saves space. It is also easy to find region choice discrepancies between or among studies because some studies many include more regions than others.

Our design follows importance-driven interactions. If the screen space is not enough to show all bars for all regions, we keep the important ones, e.g. those with large effect size, unchanged and make less important ones smaller and in context. This scaling mechanism makes the larger effect size regions visually salient. The user can directly interact with the views to rescale the size of the bar charts. Figure 4 shows an example where the bars with the effect sizes lower than 0.4 are toggled to have one fifth of normal bar width and their labels hidden.

#### View reconfiguration

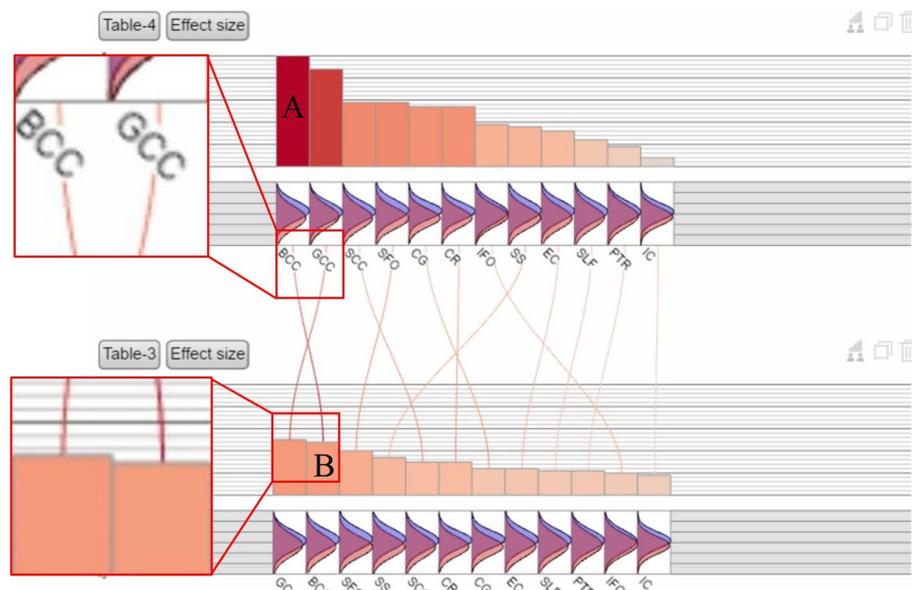
Our tool supports a set of interaction techniques: linking and brushing, zooming, panning, and view recon-

figuration. The viewer can manually select interesting brain regions under study in the effect size bar charts and examine the spatial location in the 3D view via brushing [46]. Multiple regions of interest can be selected and visualized and also linked to the artificial Manhattan plot (Fig. 5).

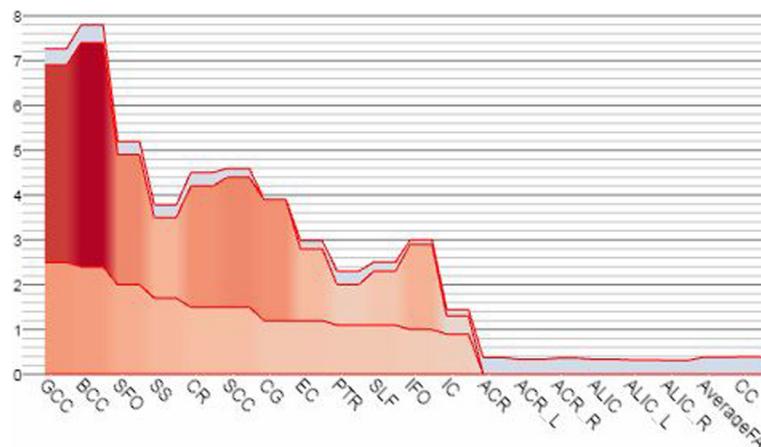
Our tools also support drag-and-drop operations to facilitate inter- and intra-study comparison. The bar charts can be dragged and dropped next to other bar chart or to the spatial view. Dragging-and-dropping a bar chart next to other bar charts can be used to rearranged the layout of multiple bar charts, which could make comparison between different studies easier. The user can also drag the bars from the right-side bar chart to the 3D glass brain regions. This action results in the display of a region-centric comparison chart. Brain regions currently being selected will be shown. This design provides a region-specific comparison mechanism.

#### Multimodality visualization

Our visualization supports multimodality visualization in that multiple attributes of brain regions can be visualized together. As can be seen from Fig. 6, the Manhattan and QQ plots are linked to the 3D brain regions. Figure 7 also shows that the plot modality and the chart modality are both linked to the 3D view. This makes it convenient to visualize multiple attributes of brain regions in the 3D view.



**Fig. 2** Scented edges to reduce the context switching cost. Here *curved lines* connect the corresponding regions between two studies (*top and bottom* ones). The color on each curve varies gradually in the way that the color in the top bar uses the color in the bottom one and vice versa. In this way, a viewer does not need to trace the link to compare study differences



**Fig. 3** Effect sizes are stacked together to support comparisons among cohorts in different brain regions. Horizontally, the cohorts are ordered by the effect size magnitudes in the bottom cohort. A viewer can easily find differences between studies

**Implementation**

ENIGMA-Viewer is implemented in Google WebGL and JavaScript and can be executed on major web-browsers such as Safari, Firefox, and Internet Explorer, without requiring any third-party software or add-ons. No account or authorization is required to use ENIGMA-Viewer and users are encouraged to email the developers with all comments and suggestions.

The sub-cortical geometry data is extracted from sub-cortical white matter atlas using marching cubes algorithm [31]. To ensure a fast loading, the mesh is only extracted when a region is selected in the 3D view. We only stored the atlas volume to reduce the data to be loaded to the browser.

**Results**

In this section we show three examples [33, 47, 48] of real world applications of this tool. The following work uses

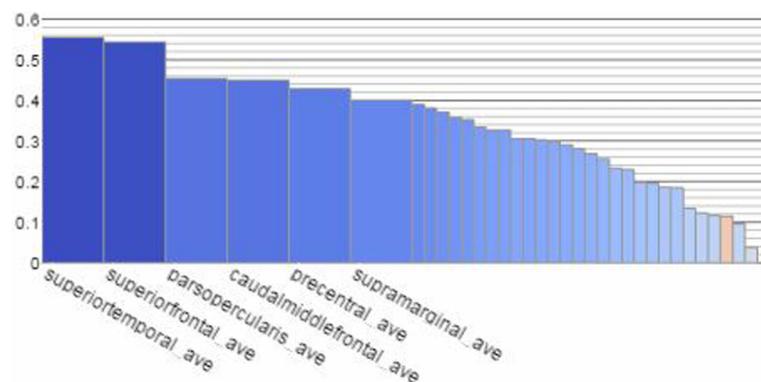
real data which are from ENIGMA group and includes both white matter and cortical gray matter comparisons.

**White matter comparisons**

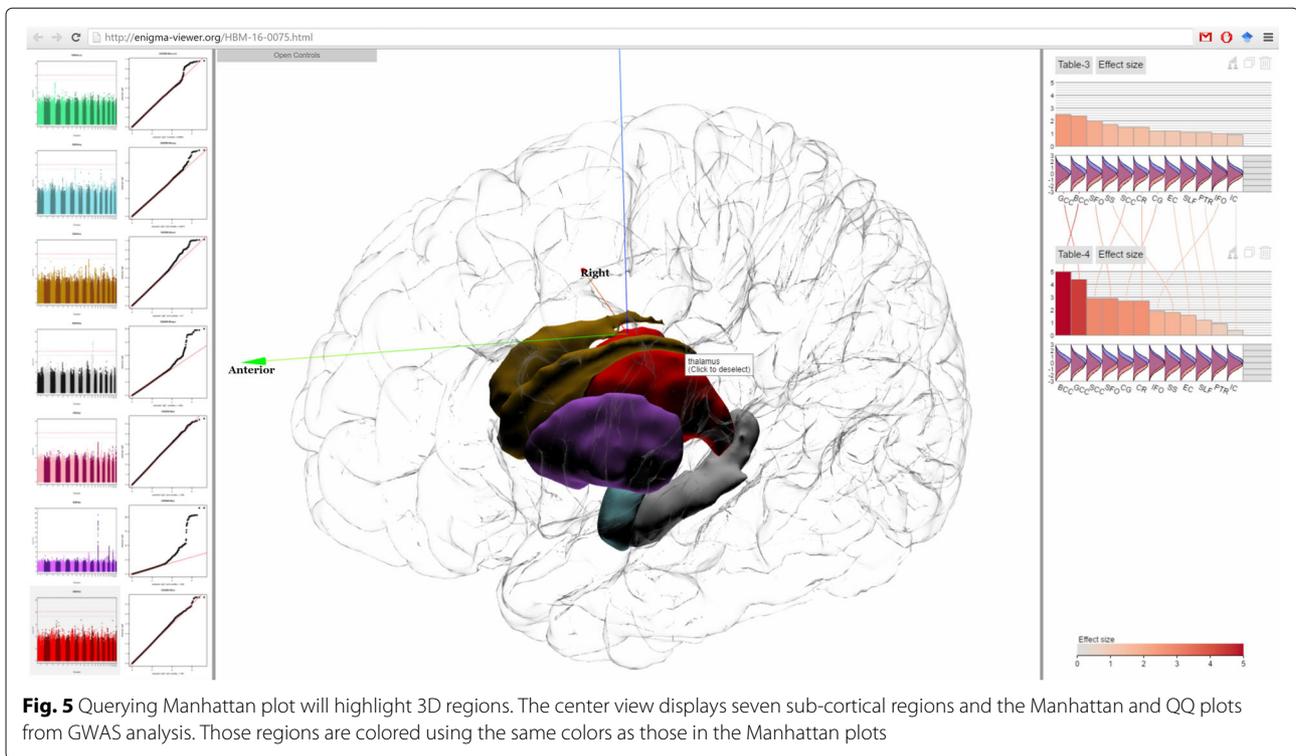
In Fig. 1, values from two result tables from a recent work [33] are displayed. In this study, DTI images of cohorts of schizophrenia patients ( $n=177$ ) and controls ( $n=249$ ) are compared to test if differences in the trajectories of white matter tract development influenced patient-control differences in FA and if specific tracts showed exacerbated decline with aging.

The top chart, named Table-3, shows the effect sizes of impact of diagnosis on white matter FA values. The bottom chart, named Table-4, shows the effect sizes of patient-control FA value decline (unit/year).

The scented lines reduce the mental cost of context switching when viewing the two bar charts from different tables. The two charts contain the same set



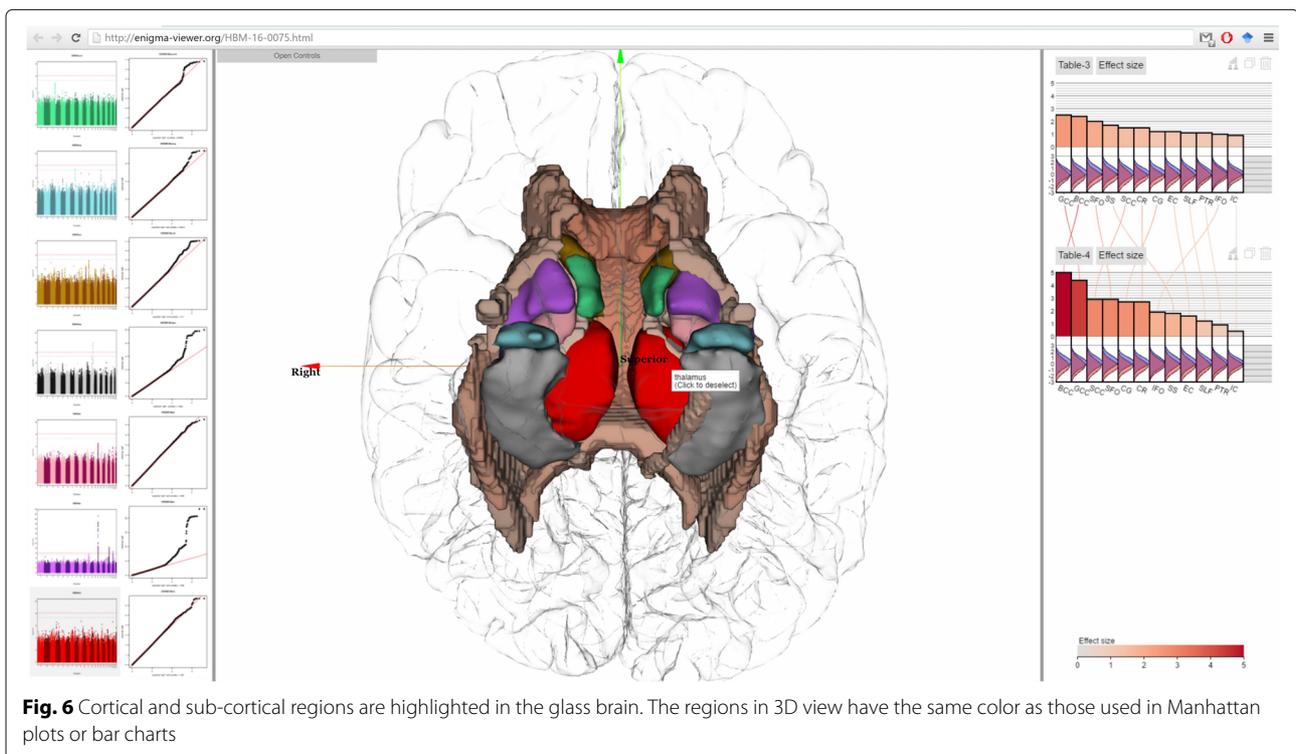
**Fig. 4** Bars can be re-scaled to make whole dataset visible in one view while leaving regions with large effect size enough screen space. In this example, bars with effect size smaller than 0.4 have their width narrowed to one fifth of normal width

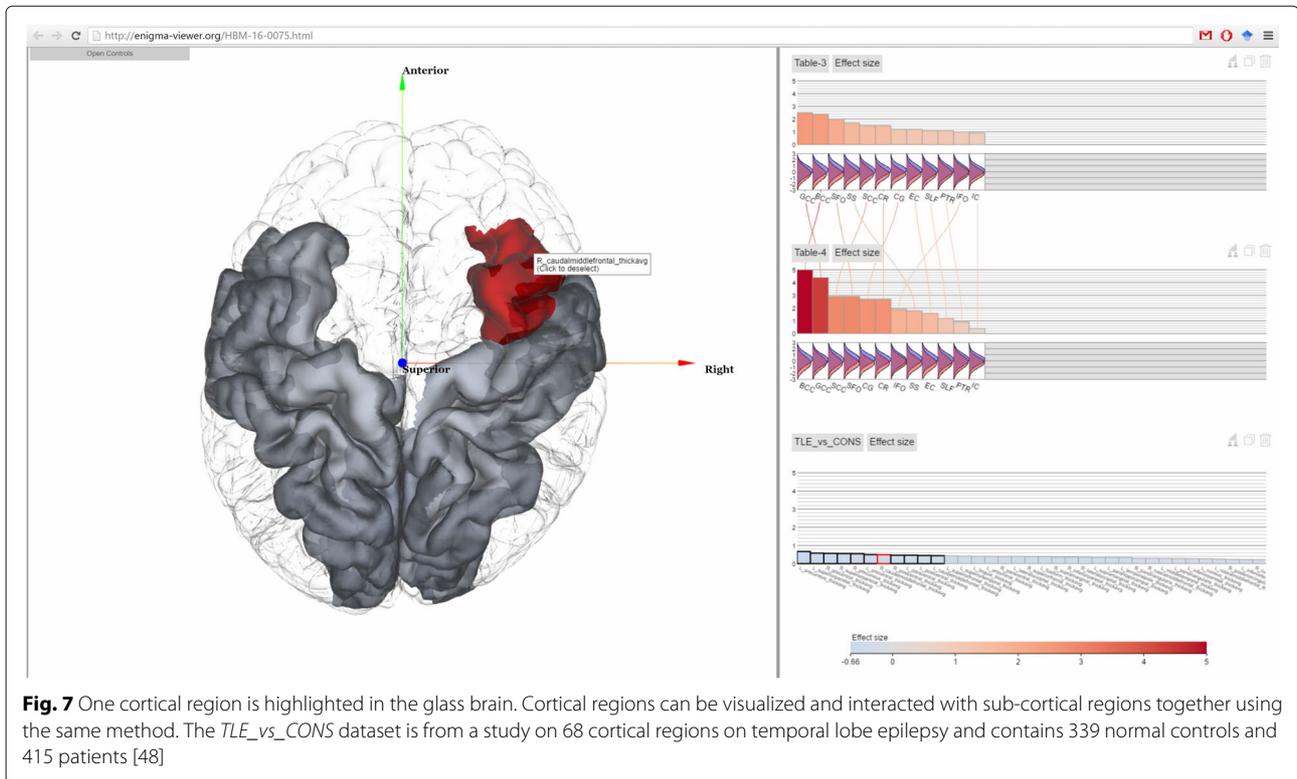


of white matter regions but each region has different effect sizes in two charts. It can be seen from the lines connecting two charts that the rankings of effect sizes are different but BCC (Body of corpus callosum) and GCC (Genu of corpus callosum) are the two regions

that show the highest patient-control difference in both charts.

The visual fusion of statistics data and 3D structure data enable users to further exam the spatial distributions of these statistics. The user can drag the tables onto





the spatial view, which shows the color encoded brain white matter structures. This, along with the in place charts, immediately reveals that the regions showing the highest effect sizes are in the middle-frontal areas, especially for data depicted in Table-4. Compare to those in Table-3, the middle-frontal areas still show the highest effect sizes, but they are not as outstanding as they are in Table-4.

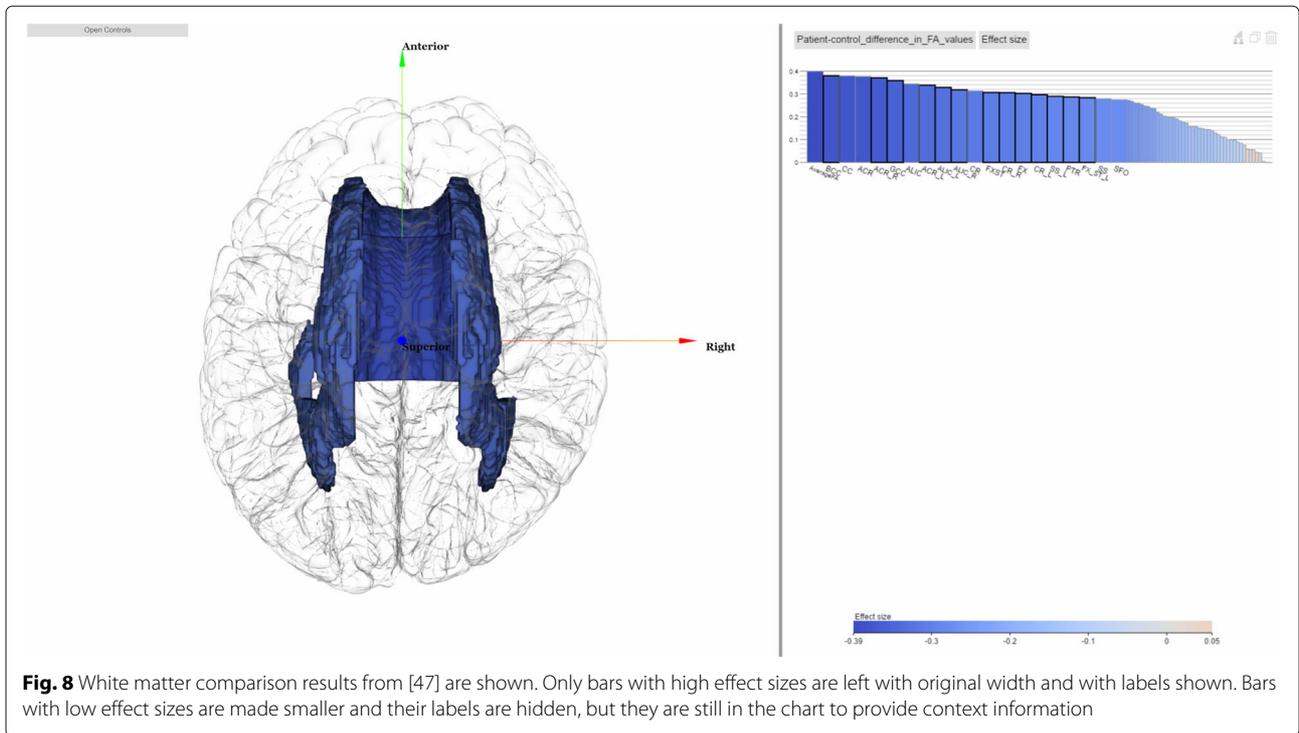
Figure 8 shows a meta analysis on brain white matter in order to identify brain regions with FA differences between schizophrenia patients and controls. The dataset comprises 30 cohorts from 14 countries totalling 2391 healthy controls and 1984 individuals with schizophrenia [47]. In the visualization, to make sure regions with more importance, i.e. higher effect sizes, are visible in the bar chart, we use focus+context method to make enough space for bars representing these regions as well as their labels. For regions with less importance, the bars can be made narrow but still in context. The focus+context design allow the brain scientists to focus on a few brain regions while keeping other regions easily accessible when needed.

### Cortical thickness comparisons

The example in Fig. 9 shows part of the results from a study [48] which contains multiple comparisons. The study pools data from 24 research centres worldwide

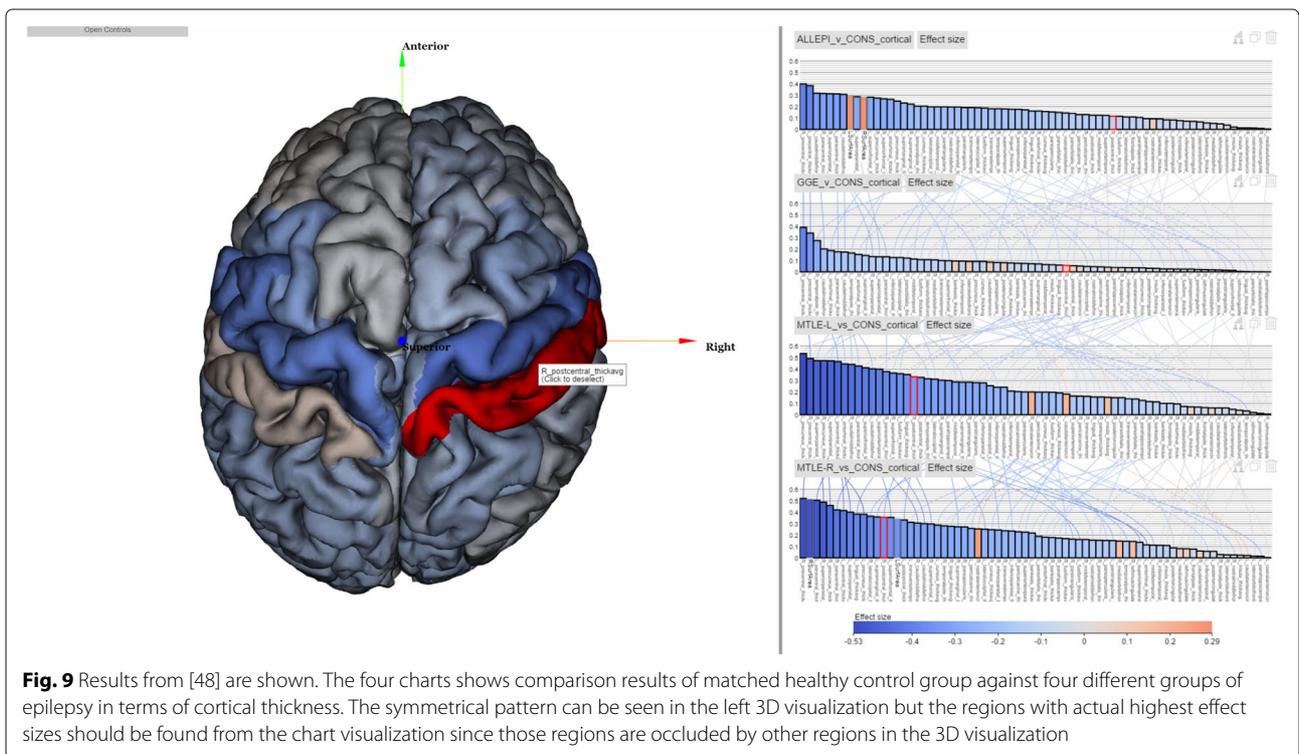
to identify reliable neuroimaging biomarkers in epilepsy. Here the four charts from top to bottom show comparison of gray matter (cortical thickness) between a matched healthy group ( $n=1727$ ) and four epilepsy groups: all types of epilepsy in aggregate (ALLEPI,  $n=2149$ ), genetic generalised epilepsies (GGE,  $n=367$ ), mesial temporal lobe epilepsies left (MTLE-L,  $n=415$ ) and mesial temporal lobe epilepsies left right (MTLE-R,  $n=339$ ). These four charts contain the same 70 cortical regions but they have different effect sizes in different comparisons. From the scented lines we can see that the rankings of regional effect sizes are different among different comparisons.

The linking of statistics data and spatial structure data via highlighting enables users to see information which is otherwise difficult to notice. In Fig. 9, the color encodings of brain cortical meshes show the results from the GGE comparison (the second topmost chart). We can see from the mesh that the result shows seemingly left-right symmetrical pattern, which is difficult to observe with bar chart only representation. On the other hand, if we look at only the brain mesh visualization we may assume that the left and right poster central regions are the two most abnormal regions since they are the most reddish color compared to other regions. However when mouse hovering those two regions' meshes, the linked bars are highlighted (with red boarder) in the



bar charts and we can see that our previous assumption is not true since they are not the ones with highest positive effect sizes in this group but the right banksst (banks of the superior temporal sulcus) regions are. This is hard to observe in the 3D brain mesh visualization

alone because the banksst regions are occluded by other cortical regions and even rotating the brain mesh cannot make this readily obvious. It is thus important to link both statistics data and spatial structure data in one visualization.



## Discussion

This section discusses alternative designs and the logical next steps to improve the usefulness and usability of the visualization design.

### The interpretation of effect size

Visualizing effect sizes only is not enough; neuroscientists must interpret and evaluate effect size for its practical significance and interpret other factors that cause the differences in different studies. Thus, how to interpret effect size is also a crucial question. The common practice here is to use the benchmarks for “small”, “medium”, and “large” effects. However, often this categorization depends on the domain of use and applying existing guidelines directly can be inappropriate.

Because such interpretation often depends on prior effect sizes in the related literature, both the size and nature of the effect should be included in the interpretation to increase the practical significance. Effect size also depends on multiple factors such as the context of the study, the importance of the outcomes, etc. Thus, visualizing the effect sizes and automatic searching and showing related context information to assist effect size interpretation within and between studies is the logical next step in truly helping neuroscientists’ decision-making.

### Limitations of the current study

Though we have attempted to use good visual design principles to guide our study, validation is the next step in creating truly meaningful tools for neuroscientists. One cannot assume that understanding has been gained from data represented graphically merely from the fact that the visualization has been presented; still less can one assume that a specific visual analysis can integrate all the capabilities required for multifaceted, spatially complex data analysis. Perception and cognition are complex and evaluation of visualization approaches for information presentation and interpretation is much needed to validate our design.

Our next step is also to improve data processing, handling, sharing, and collaborations using common infrastructure and data format standards. We plan to follow the data-format protocols defined by the ENIGMA consortium. We also plan to design dynamic and interactive queries among variables to support dynamic data analysis and to maximize flexibility to cross-link or brush-and-link across displays to find data relationships and compare and filter to remove redundancy. We will integrate the computational solutions and construct the entire workflow so that analysis and visualization can be integrated in a single framework, thus easing computing, data exploration, and human understanding of the massive datasets. The linked multiview visualization also provides a solution for brain scientists to understand how the statistics

data provenance [49] in that the factors, such as cohort distributions, that are used to produce the results can be visualized in addition to the final statistics displayed on the 3D brain.

## Conclusions

Meta-analytic thinking would considerably facilitate knowledge accumulation in brain science. In principle, meta-analysis does not overemphasize the outcomes of statistical tests in individual studies; instead, it stresses the need to explicitly design and place studies in the context of prior research. The visualization, reporting, and interpreting of effect sizes are ways to make more explicit comparisons across cohorts in a meta-study. It is particularly beneficial to incorporate prior effect sizes to guide the findings even before a study is conducted. Visualization tools improve accessibility and facilitate quick pattern finding. Our present tool has been used by several teams and has demonstrated the power of visualization to assist reporting and interpreting effect size measures.

### Abbreviations

AFNI: Analysis of Functional NeuroImages; ALLEPI: ALL types of Epilepsy in aggregate; BCC: Body of Corpus Callosum; DTI: Diffusion Tensor Imaging; ENIGMA: Enhancing Neuro Imaging Genetics Through Meta Analysis; FA: Fractional Anisotropy; fMRI: Functional Magnetic Resonance Imaging; FSL: FMRIB (Oxford Centre for Functional MRI of the Brain) Software Library; GCC: Genu of Corpus Callosum; GGE: Genetic Generalised Epilepsies; GWAS: Genome Wide Association Study; JHU: Johns Hopkins University; MTLE-L: Mesial Temporal Lobe Epilepsies-Left; MTLE-R: Mesial Temporal Lobe Epilepsies-Right; NITRC: Neuroimaging Informatics Tools and Resources Clearinghouse; ROI: Region of Interest; SOLAR: Sequential Oligogenic Linkage Analysis Routines

### Acknowledgments

Not applicable.

### Funding

This work and the publication cost were supported in part by US National Science Foundation IIS-1302755, ABI-1260795, EPS-0903234, and DBI-1062057 and by US National Institutes of Health R01 grants EB015611, NS080655, AG040060, EB008432, MH097268, MH085667, MH089722 and MH094343, and grants U01 AG024904 and P41 EB015922. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation and National Institute of Health.

### Availability of data and materials

The ENIGMA-Viewer program and datasets analyzed during the current study are available in the NITRC repository [http://www.nitrc.org/projects/enigmaviewer\\_20](http://www.nitrc.org/projects/enigmaviewer_20) and can also be accessed at <http://enigma-viewer.org>.

### Authors’ contributions

GZ programmed the ENIGMA Viewer. GZ, PK, and JC designed the ENIGMA Viewer. PK, EH, SK, CW, NJ and PT suggested improvements and provided data. GZ and JC drafted the manuscript. All authors read and approved the final manuscript. JC and PK are the corresponding authors.

### Competing interests

The authors declare that they have no competing interests.

### Consent for publication

Not applicable.

**Ethics approval and consent to participate**

Not applicable.

**About this supplement**

This article has been published as part of *BMC Bioinformatics* Volume 18 Supplement 6, 2017: Proceedings of the 3rd International Workshop on Data Mining and Visualization for Brain Science in conjunction with 7th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB'16). The full contents of the supplement are available online at <http://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-18-supplement-6>.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Author details**

<sup>1</sup> Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, 1000 Hilltop Circle, 21250 Baltimore, MD, US. <sup>2</sup> Maryland Psychiatric Research Center, University of Maryland, Baltimore, 55 Wade Ave, 21228 Baltimore, MD, US. <sup>3</sup> Keck School of Medicine, University of Southern California, 1975 Zonal Ave, 90033 Los Angeles, LA, US. <sup>4</sup> Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave, 02215 Boston, MA, US. <sup>5</sup> Psychiatry Neuroimaging Laboratory, Brigham and Women's Hospital, Harvard Medical School, 75 Francis St, 02115 Boston, MA, US. <sup>6</sup> Imaging Genetics Center, Mark and Mary Stevens Neuroimaging and Informatics Institute, University of Southern California, 1975 Zonal Ave, 90033 Los Angeles, LA, US. <sup>7</sup> Department of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, 123 St Stephen's Green, Dublin 2, Ireland.

Published: 6 June 2017

**References**

- Thompson PM, Andreassen OA, Arias-Vasquez A, Bearden CE, Boedhoe PS, Brouwer RM, Buckner RL, Buitelaar JK, Bulaeva KB, Cannon DM, et al. ENIGMA and the individual: Predicting factors that affect the brain in 35 countries worldwide. *NeuroImage*. 2017;145:389–408.
- Jahanshad N, Kochunov PV, Sprooten E, Mandl RC, Nichols TE, Alamy L, Blangero J, Brouwer RM, Curran JE, de Zubicaray GI, et al. Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the ENIGMA-DTI working group. *NeuroImage*. 2013;81:455–69.
- Kochunov P, Jahanshad N, Sprooten E, Nichols TE, Mandl RC, Alamy L, Booth T, Brouwer RM, Curran JE, de Zubicaray GI, et al. Multi-site study of additive genetic effects on fractional anisotropy of cerebral white matter: comparing meta and mega-analytical approaches for data pooling. *NeuroImage*. 2014;95:136–50.
- Hibar D, Westlye LT, van Erp TG, Rasmussen J, Leonardo CD, Faskowitz J, Haukvik UK, Hartberg CB, Doan NT, Agartz I, et al. Subcortical volumetric abnormalities in bipolar disorder. *Mol Psychiatry*. 2016;21(12):1710–716.
- Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM, Toro R, Appel K, Bartecek R, Bergmann Ø, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet*. 2012;44(5):552–61.
- Kochunov P, Jahanshad N, Marcus D, Winkler A, Sprooten E, Nichols TE, Wright SN, Hong LE, Patel B, Behrens T, et al. Heritability of fractional anisotropy in human white matter: a comparison of human connectome project and ENIGMA-DTI data. *NeuroImage*. 2015;111:300–11.
- Medland SE, Jahanshad N, Neale BM, Thompson PM. Whole-genome analyses of whole-brain data: working within an expanded search space. *Nat Neurosci*. 2014;17(6):791.
- Nivard M, Verweij K, Minică C, Treur J, Derks EM, Stringer S, Verweij KJ, Gillespie NA, Vink JM, Minică CC, et al. Connecting the dots, genome-wide association studies in substance use. *Mol Psychiatry*. 2016;21(6):733–35.
- Marcus DS, Harms MP, Snyder AZ, Jenkinson M, Wilson JA, Glasser MF, Barch DM, Archie KA, Burgess GC, Ramaratnam M, et al. Human connectome project informatics: quality control, database services, and data visualization. *NeuroImage*. 2013;80:202–19.
- Novak NM, Stein JL, Medland SE, Hibar DP, Thompson PM, Toga AW. EnigmaVis: online interactive visualization of genome-wide association studies of the enhancing Neuroimaging genetics through meta-analysis (ENIGMA) consortium. *Twin Res Hum Genet*. 2012;15(03):414–8.
- Van Horn JD, Toga AW. Human neuroimaging as a "big data" science. *Brain imaging and behavior*. 2014;8(2):323–31.
- Laidlaw DH, Ahrens ET, Kremers D, Avalos MJ, Jacobs RE, Readhead C. Visualizing diffusion tensor images of the mouse spinal cord. In: *IEEE Proceedings of Visualization*. Los Alamitos: IEEE Computer Society Press; 1998. p. 127–34.
- Schultz T, Kindlmann GL. Superquadric glyphs for symmetric second-order tensors. *IEEE Trans Vis Comput Graph*. 2010;16(6):1595–604.
- Isenberg T. A survey of illustrative visualization techniques for diffusion-weighted MRI tractography. In: *Visualization and Processing of Higher Order Descriptors for Multi-Valued Data*. Cham: Mathematics and Visualization, Springer; 2015. p. 235–56.
- Chen J, Cai H, Auchus AP, Laidlaw DH. Effects of stereo and screen size on the legibility of three-dimensional streamtube visualization. *IEEE Trans Vis Comput Graph*. 2012;18(12):2130–139.
- Penney D, Chen J, Laidlaw DH. Effects of illumination, texture, and motion on task performance in 3D tensor-field streamtube visualizations. In: *IEEE Pacific Visualization Symposium*. Los Alamitos: IEEE Computer Society Press; 2012. p. 97–104.
- Weigle C, Banks D. A comparison of the perceptual benefits of linear perspective and physically-based illumination for display of dense 3D streamtubes. *IEEE Trans Vis Comput Graph*. 2008;14(6):1723–730.
- Böttger J, Schäfer A, Lohmann G, Villringer A, Margulies DS. Three-dimensional mean-shift edge bundling for the visualization of functional connectivity in the brain. *IEEE Trans Vis Comput Graph*. 2014;20(3):471–80.
- Zielasko D, Weyers B, Hentschel B, Kuhlen TW. Interactive 3D force-directed edge bundling. *Comput Graph Forum*. 2016;35(3):127–34.
- Zhang G, Kochunov P, Hong E, Carr H, Chen J. Towards visual mega- and meta-analysis of voxel-based measurement in brain cohorts. In: *Eurographics Visualization Symposium*. Goslar: The Eurographics Association; 2016.
- Kehrer J, Hauser H. Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Trans Vis Comput Graph*. 2013;19(3):495–513.
- Keefe D, Ewert M, Ribarsky W, Chang R. Interactive coordinated multiple-view visualization of biomechanical motion data. *IEEE Trans Vis Comput Graph*. 2009;15(6):1383–390.
- Wang C, Tao J. Graphs in scientific visualization: A survey. *Comput Graph Forum*. 2016;36:263–87.
- Chen J, Pyla PS, Bowman DA. Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In: *IEEE Proceedings of Virtual Reality*. Washington: IEEE; 2004. p. 181–289.
- Hibar DP, Stein JL, Jahanshad N, Kohannim O, Hua X, Toga AW, McMahon KL, de Zubicaray GI, Martin NG, Wright MJ, et al. Genome-wide interaction analysis reveals replicated epistatic effects on brain structure. *Neurobiol Aging*. 2015;36:151–58.
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006;31(4):1487–505.
- Jahanshad N, Kohannim O, Toga AW, McMahon KL, De Zubicaray GI, Hansell NK, Montgomery GW, Martin NG, Wright MJ, Thompson PM. Diffusion imaging protocol effects on genetic associations. In: *IEEE International Symposium on Biomedical Imaging (ISBI)*. Los Alamitos: IEEE Computer Society Press; 2012. p. 944–47.
- Blangero J, Alamy L. Solar: sequential oligogenic linkage analysis routines. Population Genetics Laboratory Tech Rep 6. Southwest Foundation for Biomedical Research. 1996.
- den Braber A, Bohlken MM, Brouwer RM, van't Ent D, Kanai R, Kahn RS, de Geus EJ, Pol HEH, Boomsma DI. Heritability of subcortical brain measures: a perspective for future genome-wide association studies. *NeuroImage*. 2013;83:98–102.
- Wakana S, Jiang H, Nagae-Poetscher LM, Van Zijl PC, Mori S. Fiber tract-based atlas of human white matter anatomy. *Radiology*. 2004;230(1):77–87.
- Lorenson WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. *ACM SIGGRAPH Comput Graph*. 1987;21(4):163–69.

32. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62(2):782–90.
33. Kochunov P, Ganjgahi H, Winkler A, Kelly S, Shukla DK, Du X, Jahanshad N, Rowland L, Sampath H, Patel B, et al. Heterochronicity of white matter development and aging explains regional patient control differences in schizophrenia. *Hum Brain Mapp*. 2016;37(12):4673–688.
34. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29(3):162–73.
35. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004;23:208–19.
36. Jiang H, van Zijl PC, Kim J, Pearlson GD, Mori S. DTIStudio: resource program for diffusion tensor computation and fiber bundle tracking. *Comput Methods Prog Biomed*. 2006;81(2):106–16.
37. Amar R, Eagan J, Stasko J. Low-level components of analytic activity in information visualization. In: *IEEE Symposium on Information Visualization*. Washington: IEEE Computer Society; 2005. p. 111–17.
38. Schulz HJ, Nocke T, Heitzler M, Schumann H. A design space of visualization tasks. *IEEE Trans Vis Comput Graph*. 2013;19(12):2366–375.
39. Chen J, Forsberg AS, Swartz SM, Laidlaw DH. Interactive multiple scale small multiples. In: *IEEE Visualization 2007 Poster Compendium*. Washington: IEEE; 2007.
40. Gleicher M, Albers D, Walker R, Jusufi I, Hansen CD, Roberts JC. Visual comparison for information visualization. *Inf Vis*. 2011;10(4):289–309.
41. Karnick P, Cline D, Jeschke S, Razdan A, Wonka P. Route visualization using detail lenses. *IEEE Trans Vis Comput Graph*. 2010;16(2):235–47.
42. Convertino G, Chen J, Yost B, Ryu YS, North C. Exploring context switching and cognition in dual-view coordinated visualizations. In: *Proceedings of the International Conference on Coordinated and Multiple Views in Exploratory Visualization*. Washington: IEEE; 2003. p. 55–62.
43. Ryu YS, Yost B, Convertino G, Chen J, North C. Exploring cognitive strategies for integrating multiple-view visualizations. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. vol. 47. Thousand Oaks: SAGE Publications; 2003. p. 591–95.
44. Pirolli P. The use of proximal information scent to forage for distal content on the world wide web. *Adaptive Perspectives on Human-Technology Interaction: Methods and Models for Cognitive Engineering and Human-Computer Interaction*. Cambridge: Oxford University Press; 2007, pp. 247–66.
45. Willett W, Heer J, Agrawala M. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Trans Vis Comput Graph*. 2007;13(6):1129–136.
46. Becker RA, Cleveland WS. Brushing scatterplots. *Technometrics*. 1987;29(2):127–42.
47. Kelly S, Jahanshad N, Zalesky A, Kochunov P, Hibar D, Chen J, Agartz I, Alloza C, Andreassen O, Arango C, Banaj N, Bouix S, Bousman C, Brouwer R, Bruggemann J, Bustillo J, Cahn W, Calhoun V, Cannon DM, Carr V, Catts S, Chen J-x, Chen X, Chiapponi C, Cho KIK, Ciullo V, Corvin A, Crespo-Facorro B, Cropley V, De Rossi P, Diaz-Caneja C, Dickie E, Doan NT, Fan F, Faskowitz J, Fatouros-Bergman H, Flyckt L, Ford J, Fouché JP, Fukunaga M, Gill M, Glahn D, Gollub R, Goudzwaard E, Guo H, Gur R, Gur R, Hashimoto R, Hatton S, Henskens F, Hickie I, Hong EL, Horacek J, Howells F, Pol HH, Hyde C, IsaeV D, Whitford T, Jablensky A, Jansen P, Janssen J, Jonsson E, Kahn R, Kikinis Z, Kirra L, Klausner P, Knöchel C, Kubicki M, Kwon JS, Lagopoulos J, Langen C, Lawrie S, Lenroot R, Lim K, López-Jaramillo C, Lyall A, Magnotta V, Mandl R, Mathalon D, McCarthy R, McCarthy-Jones S, McDonald C, McEwen S, McIntosh A, Melicher T, Meshulam-Gately R, Michie P, Mowry B, Mueller B, Newell D, O'Donnell P, Oertel V, Oestreich L, Paciga S, Pantelis C, Pasternak O, Pearlson G, Pereira A, Pineda J, Piras F, Potkin S, Preda A, Rasser P, Roalf D, Roiz-Santiañez R, Pellicano GR, Roos A, Rotenberg D, Satterthwaite T, Savadjevic P, Schall U, Scott R, Seal M, Seidma L, Weickert C, Shenton M, Spalletta G, Spaniel F, Sprooten E, Stäblein M, Stein D, Sundram S, Tan YL, Tan S, Tang S, Temmingh H, Tønnesen S, Tordesillas-Gutierrez D, Vaidya J, Van Haren N, Vargas C, Vecchio D, Velakoulis D, Voineskos A, Voyvodic J, Wang Z, Wang P, Wei D, Weickert T, Westlye LT, Whalley H, White T, Wojcik J, Xiang H, Xie Z, Yamamori H, Yang FD, Yao N, Zhang G, Zhao J, Van Erp T, Turner J, Ehrlich S, Jung L, Thompson PM, Donohoe G. Widespread white matter microstructural differences in schizophrenia across 4,375 individuals: results from the ENIGMA schizophrenia DTI working group. *Mol Psychiatry* (under revision). 2017.
48. Wheland CD, Altmann A, Botia JA, Jahanshad N, Hibar DP, Absil J, Alhusaini S, Alvim MK, Auvinen P, Bartolini E, Berge F, Bernardes T, Blackmon K, Braga B, Caligiuri ME, Calvo A, Carr SJ, Chen J, Chen S, Cherubini A, David P, Domin M, Foley S, França W, Haaker G, IsaeV D, Keller SS, Kotikalapudi R, Owalczyk MA, Kuzniecky R, Langner S, Lenge M, Leyden KM, Liu M, Loi RQ, Martin P, Mascalchi M, Morita M, Pariente JC, Rodríguez-Cruces R, Rummel C, Saavalainen T, Semmelroch MK, Severino M, Thomas RH, Tondelli M, Tortora D, Vaudano AE, Vivash L, Von Podewils F, Wagner J, Weber B, Yao Y, Yasuda CL, Zhang G, Bargalló N, Bender B, Bernasconi N, Bernasconi A, Bernhardt BC, Blümcke I, Carlson C, Cavalleri GL, Cendes F, Concha L, Delanty N, Depondt C, Devinsky O, Doherty C, Focke NK, Gambardella A, Guerrini R, Hamandi K, Jackson GD, Kälviäinen R, Kochunov P, Kwan P, Labate A, McDonald CR, Meletti S, O'Brien TJ, Ourselin S, Richardson MP, Striano P, Thesen T, Wiest R, Zhang J, Zezzani A, Ryten M, Thompson PM, Sisodiya SM. Multicentre analysis of cortical thinning in the common epilepsies implicates microglial activation. *Nature Neuroscience* (under review). 2017.
49. Ragan ED, Ender T, Sanyal J, Chen J. Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes. *IEEE Trans Vis Comput Graph*. 2016;22(1):31–40.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

