NeuroImage 124 (2016) 1208–1212

Contents lists available at ScienceDirect

NeuroImage

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

The PLORAS Database: A data repository for Predicting Language Outcome and Recovery After Stroke☆

Mohamed L. Seghier, Elnas Patel, Susan Prejawa, Sue Ramsden, Andre Selmer, Louise Lim, Rachel Browne, Johanna Rae, Zula Haigh, Deborah Ezekiel, Thomas M.H. Hope, Alex P. Leff, Cathy J. Price *

Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London, UK

ARTICLE INFO

Article history:
Accepted 29 March 2015
Available online 14 April 2015

ABSTRACT

The PLORAS Database is a relational repository of anatomical and functional imaging data that has primarily been acquired from stroke survivors, along with standardized scores on a wide range of sensory, motor and cognitive abilities, demographic details and medical history. As of January 2015, we have data from 750 patients with an expected accrual rate of 200 patients per year. Expansion will accelerate as we extend our collaborations. The main aim of the database is to Predict Language Outcome and Recovery After Stroke (PLORAS) on the basis of a single structural (anatomical) brain scan that indexes the stereotactic location and extent of brain damage. Predictions are made for individual patients by indicating how other patients with the most similar brain damage, cognitive abilities and demographic details recovered their language skills over time. Predictions are validated by longitudinal follow-ups of patients who initially presented with speech and language difficulties. The PLORAS Database can also be used to predict recovery of other cognitive abilities on the basis of anatomical brain scans. The functional imaging data can be used to understand the neural mechanisms that support recovery from brain damage; and all the data can be used to understand the main sources of inter-subject variability in structure–function mappings in the human brain. Data will be made available for sharing, subject to: funding, ethical approval and patient consent.

Introduction

Our aim is to Predict Language Outcome and Recovery After Stroke (PLORAS; Price et al., 2010b) on the basis of a single structural (anatomical) brain scan that pinpoints the location and extent of brain damage (see [http://www.ucl.ac.uk/ploras]). The major challenge in this endeavor is the well-known problem of inter-patient variability in how a stroke damages the brain and, more critically, how the same lesion (i.e. site of brain damage) can have inconsistent effects on cognitive abilities (including speech) in different patients. In order to provide reliable predictions for a new patient with aphasia (i.e. speech difficulties), we aim to (i) find out how other aphasic patients with very similar lesions recovered over time; (ii) quantify the variance in the language abilities of these lesion-matched patients; and (iii) identify the main sources of this variance. We should then be able to give a new patient with aphasia an indication of the most likely time course of recovery, with an estimate of certainty in this prediction based on how many selected patients the prediction is made on and the consistency of recovery from aphasia in these patients. This approach is similar to that used when generating a prognosis for patients with cancer; doctors usually rely upon statistics based on data from large groups of cancer patients whose situations are most similar to that of the patient.

The PLORAS Database is a repository of information about the speech and language abilities of hundreds of stroke patients who may or may not have suffered from aphasia (currently >750, though this number is growing all the time). It started in 2008 (Price et al., 2010b) and has attracted long term funding in the UK from the Wellcome Trust since 2007, the Stroke Association since 2014 and the Medical Research Council since 2015. At a minimum, the data available for each patient includes: (i) a structural brain scan, (ii) the results of a cognitive and speech and language assessment (from the Comprehensive Aphasia Test (CAT); Swinburn et al., 2004), (iii) demographic information (age, gender, handedness, time post-stroke, etc.), and (iv) information about their stroke and other co-morbidities. The structural brain images are used to index the degree of damage or preservation at each voxel in the brain (Seghier et al., 2008). By integrating this information with knowledge of the patients' speech, language and cognitive abilities over time, we plan to generate a clinical tool that will provide data-driven predictions of language outcome and recovery after stroke (Hope et al., 2013, 2015), see Fig. 1. This multivariate data-driven

☆ ***Neuroimage Special Issue***: Sharing the wealth: Brain Imaging Repositories in 2015.

* Corresponding author at: Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, 12 Queen Square, London WC1N 3BG, UK.
E-mail address: c.j.price@ucl.ac.uk (C.J. Price).

http://dx.doi.org/10.1016/j.neuroimage.2015.03.083

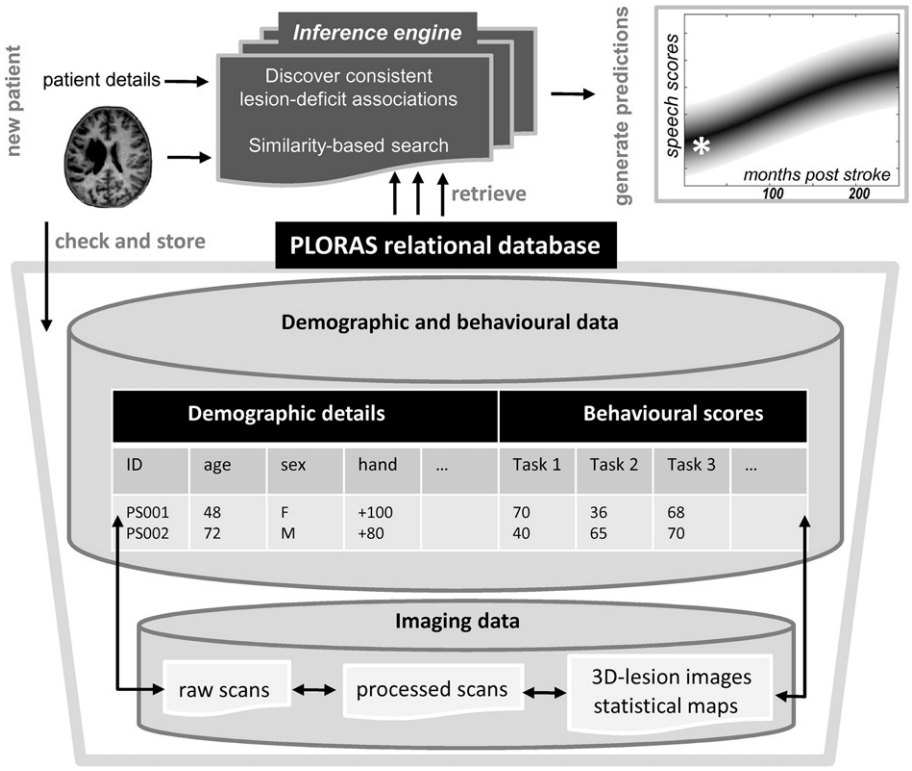

Fig. 1. A schematic illustration of the PLORAS system. Currently we have 750 patients who vary with respect to age (range: 19–99 years), gender, handedness, language skills (87% of all patients are native English speakers), time post-stroke (from a few weeks to 40 years after stroke), lesion location and size (lesion size: from <1 cm^3 to >400 cm^3, lesion side: 59% of patients have lesions in the left hemisphere, 22% in the right hemisphere, 18% in both hemispheres), and language deficits. The bespoke relational database links demographic, behavioral and imaging data in a single data repository. An automated lesion identification technique transforms the brain scan of each patient to a high-resolution stereotactic 3D-lesion image (Seghier et al., 2008). Based on the learnt structure–function–recovery rules from all patients in the PLORAS Database, the inference engine generates a probabilistic recovery curve for the new patient (Hope et al., 2013). The curve is a probability distribution through time, with the mean prediction in black, and borders at 2 standard deviations from the mean (i.e. 95% confidence). The white star illustrates the real speech score of the new patient measured at a given time point after stroke.

approach aligns with other current trends in this era of big data and data mining (Amari et al., 2002; Gee et al., 2010; Marcus et al., 2007; Mueller et al., 2005; Van Horn and Toga, 2009, 2014).

One of the unique strengths of the PLORAS Database is that all patients are assessed in exactly the same way, and contribute scores for each subtest of the CAT, irrespective of whether or not they have a diagnosed aphasia deficit, their lesion site, time post-stroke or other demographic information. This provides a more detailed and standardized set of data to that collected in the clinical setting where testing is typically tailored to the patient's deficit. Our plan is to link all types of data (demographic, behavioral and imaging) in a format that is easy to store, maintain, interrogate, search, extract, transfer, process and share. Users can access full details about how data were collected, transformed and analyzed. Heterogeneous data from different sites and contributors can be added along with longitudinal studies of the same patient to complement the cross sectional studies. Having large data samples provides the opportunity to understand, characterize and model inter-subject variability in brain function (Miller and Van Horn, 2007), which will help us to better understand inter-patient variability in recovery.

In the last five years, we have used the PLORAS Database to: (i) search for the optimal (necessary) combination of lesions that impairs a given ability (Price et al., 2010a); (ii) learn complex multivariate lesion–symptom associations in a data-driven way (Hope et al., 2013); (iii) test the predictive power of identified compensatory pathways for a given language skill (Seghier et al., 2014); and (iv) test whether critical lesion sites vary with other non-lesion factors (e.g. age, gender and language skills (Hope et al., 2015)). The PLORAS methodology and database is also applicable for predicting deficits other than aphasia; for instance, to study the neural underpinnings of other functions that are captured by the CAT (e.g. sensory processing and memory) and how these functions break down and recover in stroke patients.

Last but not least, we also collect functional magnetic resonance imaging (fMRI) data from a subset of patients who have good language abilities despite relatively large left or right hemisphere strokes. The fMRI data allow us to identify which of the preserved brain regions are functioning normally or abnormally and how speech and language abilities can be recovered after stroke. This may help to refine predictions in the future but is primarily collected for a scientific explanation of how the predictions are working, and of how individual patients recover lost language skills.

In a previous paper (Price et al., 2010b), we set out the conceptual foundations that motivated the PLORAS Database. The current paper describes the relational structure of the database, the types of images that are available, quality control and general procedures. It also provides concrete numbers of the types of patients that are included and information about how to contribute and access the database.

What is the PLORAS Database designed to do?

In 2008, the PLORAS Database (Fig. 1) was set up to look at inter-subject variability in (i) lesion-deficit mappings and (ii) functional localisation and integration in healthy and brain damaged patients. We also need large populations of patient data to test our hypothesis that more than one neural system can support the same cognitive task (Price and Friston, 2002; Seghier et al., 2012; Seghier et al., 2014). This approach involves a combination of functional imaging and multiple lesion–deficit analyses (i.e. multivariate analyses) that systematically

investigate how the effect of damage to one brain region (e.g. Region-X) depends on the presence or absence of damage to other brain regions (e.g. Region-Y) that might compensate for the loss of Region-X (Price and Friston, 2002). For example, if a deficit is only incurred after combined damage to X&Y but not X-only or Y-only, then we hypothesize that (i) the task can be supported by X when Y is damaged, (ii) by Y when X is damaged, and (iii) recovery capacity is poor when both X&Y are damaged.

More recently, the large datasets we have accrued have also allowed us to use data-driven machine learning approaches to make individualized prognoses about language outcome and recovery after stroke (cf. Hope et al., 2013, 2015). The results show that once the main sources of inter-subject variability are understood and modeled, we are able to find very consistent lesion–deficit mappings. We believe that this approach will ultimately help individual patients and their carers make important life decisions (e.g. return to work; type of rehabilitation and required support systems).

In order to support the studies described above, the PLORAS Database has been designed to: provide a single repository for all types of data; with maximum accessibility for future users. For example, the PLORAS Database has a sophisticated and flexible query tool that can select any combination of data (or metadata) from a large number of parameters for the same patient.

PLORAS data

Subjects

Stroke survivors, irrespective of lesion site or size or time post-stroke. In January 2015, 750 patients had contributed to the PLORAS Database, and this is expected to grow at a rate of 200 patients per year with funding currently available until 2018 and recruitment procedures in place with multiple sites across the UK, coordinated by the NIHR Clinical Research Network [http://www.crn.nihr.ac.uk/]. Neurologically normal imaging data from approximately 510 healthy adults are also available for establishing inter-subject variability in brain structure and function. Participants give their consent to how their data can be used, stored and managed. Subject to the provisions of these consents, pseudo-anonymized data can be made available to external researchers (see below).

Other information

Age, gender, handedness, time since stroke, languages spoken, mother tongue, country of birth, language proficiency prior to stroke, any reading difficulties (dyslexia) prior to stroke, education and occupation. We also record whether the patient had any other mental health disorders or loss of hearing and vision.

Structural imaging data

High-resolution structural MRI scans are available for all subjects, acquired from several different 1.5T and 3T scanners at UCL (London) using 3D whole brain T1-weighted sequences. Our choice to limit the essential imaging data to a T1-weighted structural MRI scan is motivated by our desire to produce a quick and easy method for lesion identification and prognoses that can be based on clinically acquired brain scans with minimal data analysis. Additional information could indeed be gained from Diffusion Tensor Imaging (DTI), Perfusion Weighted Imaging (PWI) or resting-state data. We are not currently collecting such data because (1) this type of imaging data is not routinely collected in the clinical setting, and (2) the data we are collecting are already providing a very rich source of information that is yielding extremely reliable and informative results. As of 2015, we will be accepting structural MRI or CT scans from other research centers and hospitals.

Functional MRI data, paradigm 1

Data have been collected from 90 stroke survivors and 194 neurologically normal subjects, scanned in the same 1.5T Siemens scanner, while engaged in 8 different tasks:

(1) semantic associations on pictures of objects;
(2) semantic associations on written object names;
(3) naming objects aloud;
(4) reading written object names aloud;
(5) perceptual decisions on pictures of meaningless object shapes;
(6) perceptual decisions on meaningless Greek letter strings;
(7) saying “1,2,3” aloud to the meaningless object shapes;
(8) saying “1,2,3” to the meaningless Greek letter strings.

Data collection is now complete and has already contributed to multiple publications (c.f. Hu et al., 2010; Josse et al., 2009; Seghier et al., 2010; Seghier and Price, 2011, 2012).

Functional MRI data, paradigm 2

Data collection is on-going. In January 2015, we had scanned 51 stroke patients and 96 neurologically normal subjects engaged in 13 different tasks:

(1) production of simple declarative sentences to describe two interacting objects with 2 nouns and a verb (e.g. “the goat is eating the hat”);
(2) object naming from pictures of 2 objects (e.g. “goat and hat”);
(3) production of one verb (e.g. eating);
(4) semantic decisions on pictures of two objects;
(5) semantic decisions on heard object names;
(6) object naming (one object per picture);
(7) reading object names;
(8) reading meaningless pseudowords;
(9) naming the color of meaningless object shapes;
(10) hearing environmental sounds and naming their (animal and object) source;
(11) repeating heard object names;
(12) repeating heard pseudowords;
(13) hearing a male or female voice humming and naming the gender.

This paradigm is designed to tease apart different types of semantic and syntactic processing (Sanjuán et al., in press) as well as visual, auditory, and different types of phonological processing, articulatory planning, articulation and auditory processing of the spoken response (Hope et al., 2014; Parker Jones et al., 2014).

Image processing

Structural and functional images are stored in their raw format, and after segmentation, normalization and/or smoothing following standard pre-processing in SPM [http://www.fil.ion.ucl.ac.uk/spm/]. We also convert all patient structural scans into fuzzy and binary images that index the degree of anatomical abnormality at each voxel (Seghier et al., 2008), thus providing a 3D representation of the lesion site in a standard stereotactic space. Functional images are also available in the form of whole brain “contrast images” (i.e. statistical maps) which index the difference in BOLD signal between conditions.

Format

All imaging data types have been converted into ANALYZE or NIFTI-1 formats that are compatible with almost all MRI analysis packages.

Behavioral data

Standardized scores from the Comprehensive Aphasia Test (CAT, Swinburn et al., 2004) are available for all patients. Scores for different subtests indicate how well patients can comprehend and produce spoken and written words, as well as a range of other sensory, motor, cognitive and memory functions. Together the scores provide a detailed indication of how patients have been affected by their stroke (Howard et al., 2010). For any patients who have abnormally low scores, we aim to repeat the testing in order to monitor recovery.

Sources

We do not currently acquire or provide any data over the web.

Quality control

The PLORAS Team consists of the Chief Investigator (Professor Cathy Price), the Recruitment, Data Management, Research, Systems and other support Teams through whom quality is maintained.

Behavioral data are collected, scored and standardized by a dedicated team of speech pathologists who also enter the demographic information onto the database.

Imaging data are collected using optimized acquisition protocols that provide high image quality in terms of MR contrast and resolution. The acquisition protocols are constantly assessed, quantitatively and qualitatively, for image quality by our Physics and Methods Teams. The quality of the raw images collected in-house is checked by our radiographers prior to being transferred to the Research Team. The quality of imaging data collected off-site is checked by the Research Team who also conducts and checks the results of all image pre-processing, with particular attention to the spatial normalization and lesion identification routines. All raw and processed images are uploaded to the database by a member of the Data Management Team.

The PLORAS dedicated Data Manager ensures that all data types are effectively and securely uploaded to the database, anonymized, stored, backed up, archived, encrypted and shared with the highest standards of confidentiality and minimum risk (see details at [http://www.ucl.ac.uk/informationsecurity/policy]). Our department's expert IT team handles system security and data back-up. All procedures for collecting and analyzing data are reported in published articles.

Accessing the data

Direct Access to the data is password protected and limited to relevant members of the PLORAS Research Team and local collaborators. The level of access is adjusted according to the role of the individual concerned. The Data Manager allocates appropriate levels of data access to different users in line with user needs and Data Protection Legislation. For example, all private information (e.g. name, address and sensitive medical records) are non-disclosive to analysts and are encrypted.

Data sharing with external users is currently limited. However, the data are highly relevant for external researchers who want to (i) cross-validate (or replicate) lesion–symptom mappings in large independent samples, (ii) test the predictive power of current neurological models, or (iii) implement robust spatial warping or tissue segmentation techniques. External users agree to inform the Chief Investigator of any publication that used PLORAS data.

Requests for data from external sources must be addressed to the Chief Investigator (Professor Cathy Price) and are approved subject to (i) evidence of study specific peer reviewed funding, (ii) ethics, (iii) any constraints imposed by the relevant participant consents and (iv) funding and manpower availability. We share pseudo-anonymized imaging data, behavioral scores, demographic details and some medical history, as required by the investigators of the approved study. We do not share personal identifiable information with external researchers. The Chief Investigator will also advise on the level of collaboration that might be required from the PLORAS Research Team. For example, given the complex nature of some of the data (e.g. functional imaging data relating to a particular paradigm), advice from the PLORAS Team in terms of analytical methods is likely to be required if the data are to be used effectively by external researchers. External users should not share data with third parties or provide PLORAS data for commercial or marketing purposes. All such intentions or requirements should be emailed to the Chief Investigator.

Procedures for approved data sharing

The external researcher completes a requisition form, providing exact details of the data needed, the format, the type and any inclusion or exclusion criteria.

The Data Manager:

(1) assesses the request and confirms availability;
(2) extracts relevant data in the required format subject to patient consent, see below;
(3) ensures pseudo-anonymization for confidentiality of all personal and sensitive information, and annotation of all datatypes, with consistent personal identification codes;
(4) transfers data to a temporary secure area;
(5) informs the client on the means of accessing the data;
(6) helps the client in the event of any access difficulties;
(7) deletes data after confirmation of safe receipt;
(8) keeps a record of what information has been provided.

Formatting of behavioral and demographic data can be provided in CSV, DOC, XLS or XML. Structural and functional imaging data are provided in ANALYZE or NIFTI-1 formats that are compatible with almost all MRI analysis packages.

How to contribute to PLORAS

Inclusion criteria

At a minimum, each new patient entry needs to include a CT or MRI anatomical brain scan (for lesion identification), the results from at least some standardised language tests (e.g. name pictures, match speech to pictures) and a range of demographic details and acknowledgment of any co-morbidities and sensory processing loss.

Ethical approval

We require evidence that all procedures for patient recruitment, MRI scanning, and behavioral assessments have been ethically approved. Once received, data from external sources will get the same quality of treatment as our in-house data in terms of confidentiality and data protection.

Data sharing agreement

This was established to acquire further stroke patient data from multiple NHS sites across the UK. This complies with all relevant standards of Information Governance required by the NHS.

Guidelines on how to contribute to PLORAS will be advertised in our website [http://www.ucl.ac.uk/ploras].

Long term plans

Our long-term goal is to develop a tool that can be used by clinicians to predict language outcome and recovery after brain damage. The more patients we have on the database, the more confident we can be in the

accuracy of our predictions. Likewise, longitudinal data is essential for testing the time course of recovery for individual patients. To accommodate our database growth, data storage requirements, ease of use and efficiency are continuously evaluated, updated and maintained. We also have plans to implement standardization procedures for metadata sharing (Keator et al., 2012; Poline et al., 2012), which will allow us to collaborate with other initiatives with similar aims. Ultimately, multicenter, multinational approaches will accelerate data mining of large samples and facilitate data-led clinical translation of neuroimaging results in stroke (Gee et al., 2010).

Acknowledgments

The PLORAS Database has been funded by the Wellcome Trust since 2007 and the Stroke Association since 2014. Many people other than the authors of this paper have acquired data in previous years including Jenny Crinion, Alice Grogan, Thomas Schofield, Katharine Bowers, Erin Carrol, Nicola Wilson, Caroline Ellis, Julie Stewart, Sharon Adjei, Rebecca Lee, Anna Isherwood, Matthew Lawrence, Karine Gazarian, Louise Ruffle, Clare Shakeshaft, and Laura Stewart. All requests regarding data contribution or sharing should be addressed to the corresponding author. Details about the PLORAS Team members can be found in our website [http://www.ucl.ac.uk/ploras].

References

Amari, S., Beltrame, F., Bjaalie, J.G., Dalkara, T., De Schutter, E., Egan, G.F., Goddard, N.H., Gonzalez, C., Grillner, S., Herz, A., Hoffmann, K.P., Jaaskelainen, I., Koslow, S.H., Lee, S.Y., Matthiessen, L., Miller, P.L., Da Silva, F.M., Novak, M., Ravindranath, V., Ritz, R., Ruotsalainen, U., Sebestra, V., Subramaniam, S., Tang, Y., Toga, A.W., Usui, S., Van Pelt, J., Verschure, P., Willshaw, D., Wrobel, A., 2002. Neuroinformatics: the integration of shared databases and tools towards integrative neuroscience. J. Integr. Neurosci. 1, 117–128.

Gee, T., Kenny, S., Price, C.J., Seghier, M.L., Small, S.L., Leff, A.P., Pacurar, A., Strother, S.C., 2010. Data warehousing methods and processing infrastructure for brain recovery research. Arch. Ital. Biol. 148, 207–217.

Hope, T.M.H., Seghier, M.L., Leff, A.P., Price, C.J., 2013. Predicting outcome and recovery after stroke with lesions extracted from MRI images. NeuroImage Clin. 2, 424–433.

Hope, T.M., Prejawa, S., Parker Jones, O., Oberhuber, M., Seghier, M.L., Green, D.W., Price, C.J., 2014. Dissecting the functional anatomy of auditory word repetition. Front. Hum. Neurosci. 8, 246.

Hope, T.M.H., Parker Jones, O., Grogan, A., Crinion, J., Rae, J., Ruffle, L., Leff, A.P., Seghier, M.L., Price, C.J., Green, D.W., 2015. Comparing language outcomes in monolingual and bilingual stroke patients. Brain 138 (4), 1070–1083.

Howard, D., Swinburn, K., Porter, G., 2010. Putting the CAT out: what the Comprehensive Aphasia Test has to offer. Aphasiology 24, 56–74.

Hu, W., Lee, H.L., Zhang, Q., Liu, T., Geng, L.B., Seghier, M.L., Shakeshaft, C., Twomey, T., Green, D.W., Yang, Y.M., Price, C.J., 2010. Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. Brain 133, 1694–1706.

Josse, G., Kherif, F., Flandin, G., Seghier, M.L., Price, C.J., 2009. Predicting language lateralization from gray matter. J. Neurosci. 29, 13516–13523.

Keator, D.B., Helmer, K., Steffener, J., Turner, J.A., Van Erp, T.G., Gadde, S., Ashish, N., Burns, G.A., Nichols, B.N., 2012. Towards structured sharing of raw and derived neuroimaging data across existing resources. NeuroImage 82.

Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L., 2007. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. J. Cogn. Neurosci. 19, 1498–1507.

Miller, M.B., Van Horn, J.D., 2007. Individual variability in brain activations associated with episodic retrieval: a role for large-scale databases. Int. J. Psychophysiol. 63, 205–213.

Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C.R., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., 2005. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimers Dement. 1, 55–66.

Parker Jones, O., Prejawa, S., Hope, T.M., Oberhuber, M., Seghier, M.L., Leff, A.P., Green, D.W., Price, C.J., 2014. Sensory-to-motor integration during auditory repetition: a combined fMRI and lesion study. Front. Hum. Neurosci. 8, 24.

Poline, J.B., Breeze, J.L., Ghosh, S., Gorgolewski, K., Halchenko, Y.O., Hanke, M., Haselgrove, C., Helmer, K.G., Keator, D.B., Marcus, D.S., Poldrack, R.A., Schwartz, Y., Ashburner, J., Kennedy, D.N., 2012. Data sharing in neuroimaging research. Front. Neuroinform 6, 9.

Price, C.J., Friston, K.J., 2002. Degeneracy and cognitive anatomy. Trends Cogn. Sci. 6, 416–421.

Price, C.J., Seghier, M.L., Leff, A.P., 2010a. Predicting language outcome and recovery after stroke: the PLORAS system. Nat. Rev. Neurol. 6, 202–210.

Price, C.J., Crinion, J.T., Leff, A.P., Richardson, F.M., Schofield, T.M., Prejawa, S., Ramsden, S., Gazarian, K., Lawrence, M., Ambridge, L., Andric, M., Small, S.L., Seghier, M.L., 2010b. Lesion sites that predict the ability to gesture how an object is used. Arch. Ital. Biol. 148, 243–258.

Sanjuán, A., Hope, T.M.H., Parker Jones, O., Prejawa, S., Oberhuber, M., Guerin, J., Seghier, M.L., Green, D.W., Price, C.J., 2015. Dissociating the semantic function of two neighbouring subregion in the left lateral anterior temporal lobe. Neuropsychologia (in press).

Seghier, M.L., Price, C.J., 2011. Explaining left lateralization for words in the ventral occipitotemporal cortex. J. Neurosci. 31, 14745–14753.

Seghier, M.L., Price, C.J., 2012. Functional heterogeneity within the default network during semantic processing and speech production. Front. Psychol. 3, 281.

Seghier, M.L., Ramlackhansingh, A., Crinion, J., Leff, A., Price, C.J., 2008. Lesion identification using unified segmentation–normalisation models and fuzzy clustering. NeuroImage 41, 1253–1266.

Seghier, M.L., Fagan, E., Price, C.J., 2010. Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. J. Neurosci. 30, 16809–16817.

Seghier, M.L., Neufeld, N.H., Zeidman, P., Leff, A.P., Mechelli, A., Nagendran, A., Riddoch, J.M., Humphreys, G.W., Price, C.J., 2012. Reading without the left ventral occipito-temporal cortex. Neuropsychologia 50, 3621–3635.

Seghier, M.L., Bagdasaryan, J., Jung, D., Price, C.J., 2014. The importance of premotor cortex for supporting speech production after left capsular–putaminal damage. J. Neurosci. 34, 14338–14348.

Swinburn, K., Porter, G., Howard, D., 2004. Comprehensive Aphasia Test. Psychology Press.

Van Horn, J.D., Toga, A.W., 2009. Is it time to re-prioritize neuroimaging databases and digital repositories? NeuroImage 47, 1720–1734.

Van Horn, J.D., Toga, A.W., 2014. Human neuroimaging as a "Big Data" science. Brain Imaging Behav. 8, 323–331.