NeuroImage: Clinical 19 (2018) 1–13

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Predicting the pattern and severity of chronic post-stroke language deficits from functionally-partitioned structural lesions

Ajay D. Halai*, Anna M. Woollams, Matthew A. Lambon Ralph*

Neuroscience and Aphasia Research Unit, Division of Neuroscience & Experimental Psychology, Faculty of Biology, Medicine and Health, University of Manchester, UK

ARTICLE INFO

Keywords:
Aphasia
Prediction
Principal component analysis
Regression
Stroke

ABSTRACT

There is an ever-increasing wealth of knowledge arising from basic cognitive and clinical neuroscience on how speech and language capabilities are organised in the brain. It is, therefore, timely to use this accumulated knowledge and expertise to address critical research challenges, including the ability to predict the pattern and level of language deficits found in aphasic patients (a third of all stroke cases). Previous studies have mainly focused on discriminating between broad aphasia dichotomies from purely anatomically-defined lesion information. In the current study, we developed and assessed a novel approach in which core language areas were mapped using principal component analysis in combination with correlational lesion mapping and the resultant '*functionally-partitioned*' lesion maps were used to predict a battery of 21 individual test scores as well as aphasia subtype for 70 patients with chronic post-stroke aphasia. Specifically, we used lesion information to predict behavioural scores in regression models (cross-validated using 5-folds). The winning model was identified through the adjusted R^2 (model fit to data) and performance in predicting holdout folds (generalisation to new cases). We also used logistic regression to predict fluent/non-fluent status and aphasia subtype. Functionally-partitioned models generally outperformed other models at predicting individual tests, fluency status and aphasia subtype.

1. Introduction

Left hemisphere stroke often results in disrupted speech and language processes (aphasia). Under the single umbrella term of 'aphasia' there are considerable variations in patients' language and cognitive presentation, in both the pattern and severity of impairment to different language activities (e.g., comprehension, naming, reading, writing, speech, etc.). The consequence of this significant diversity is that individual patients will need very different types of intervention and clinical management (e.g., patients with primary comprehension or phonological deficits). By utilising fMRI in healthy participants (Price, 2010, 2012) and voxel-lesion symptom mapping (Bates et al., 2003) in aphasic patients, cognitive and clinical neuroscience has made considerable strides in mapping language performance and the underpinning cognitive mechanisms to different brain regions. Despite being a crucial step for clinical application, the reverse mapping – using neuroimaging results to predict individual aphasic profiles or types – has only been attempted by a limited number of studies. The key aim of this investigation, therefore, was to embark on using new methods to generate lesion-based models which are able to predict both the detailed language profile of individual patients as well as their aphasia classification. For clarity, in this study we use prediction-based inference to determine how neural data can predict the current behavioural status using a k-fold cross validation approach. Future studies will be able to test whether similar models can offer accurate prediction in the temporal sense (using neural data to predict future behaviour). Indeed, the chronic stroke lesion is apparent long before patients' long-term language and cognitive abilities have stabilised (the partial, gradual recovery that most patients demonstrate extends to at least nine to twelve months post onset). Accordingly, accurate lesion-based prediction models would have considerable clinical utility, including improvements in the type of information that can be offered to patients and carers, enhanced clinical management planning, and appropriate patient stratification to treatment plans.

Studies using neural lesion information to predict behavioural outcomes have yielded inconsistent results. For example, earlier studies reported little advantage of using lesion information in improving predictions (Hand et al., 2006; Johnston et al., 2002; Johnston et al., 2007; Lazar et al., 2008; Willmes and Poeck, 1993). In contrast, more recent studies have found that models, designed to predict a single

* Corresponding authors at: Neuroscience and Aphasia Research Unit, Division of Neuroscience & Experimental Psychology, Faculty of Biology, Medicine and Health, University of Manchester, 3.19 Zochonis Building, Brunswick Street, Manchester M13 9PL, England, UK.
E-mail addresses: ajay.halai@manchester.ac.uk (A.D. Halai), matt.lambon-ralph@manchester.ac.uk (M.A. Lambon Ralph).

https://doi.org/10.1016/j.nicl.2018.03.011
Received 8 May 2017; Received in revised form 28 February 2018; Accepted 13 March 2018
Available online 16 March 2018

feature of aphasic performance or aphasia type, can be improved by including lesion information (Hope et al., 2013; Saur et al., 2010; Schiemanck et al., 2006; Thijs et al., 2000; Yourganov et al., 2015). For example, Hope et al. (2013) developed a predictive model using basic demographic information (age, gender, etc.) and structural lesion information obtained from a high-resolution T1-weighted image (lesion size and atlas-based lesions) to predict a composite speech production score (and its constituent individual speech test results), with the winning model containing time post-onset, lesion volume and 35 atlas-based predictors. They showed that the model could predict patients' composite speech production score over the first 200 months post-stroke. In addition, the same group used anatomical regions to predict 22 subtests scores of the Comprehensive Aphasia Test (Swinburn et al., 2005) for mono-/bi-lingual patients (Hope et al., 2015). Another study used ridge regression in order to predict behavioural scores across seven domains (left/right motor, language, attention bias, verbal memory and spatial memory) in acute stroke cases (< 2 weeks) (Corbetta et al., 2015) - though, language was identified in a broad sense and thus the study did not allow for predictions of specific language deficits. Other groups have used support vector machines (SVM) trained on atlas-based lesion parcellations to predict six out of ten pairwise binary contrasts between aphasia subtypes at above chance levels (Yourganov et al., 2015). Saur et al. (2010) also used SVMs in order to predict patients' chronic outcome status (a binary classification; good/bad) as well as the type of improvement from the acute to chronic stage (good/bad). They found that age and a composite language recovery score (LRS) achieved above chance classification (62%). It is important to note that this particular study made use of fMRI measurements and showed that the fMRI data within targeted language areas improved prediction accuracy improved significantly (~86%), suggesting that functionally- as well as neuroanatomically-partitioned maps might be critical in improving predictive models. Furthermore, we also know that white matter connectivity (or disconnection) plays an important role in understanding behavioural deficits (Catani and ffytche, 2005; Catani et al., 2005). A recent study has shown that damage to white matter pathways that converge into a bottleneck, for example in the posterior temporal lobe, are critical in predicting multiple behavioural deficits such as speech fluency, naming and auditory semantic decisions (Griffis et al., 2017).

The present study advances this handful of existing prediction models in two novel and important ways – namely, (a) how patients' lesions are partitioned (before they are used as predictors) and (b) in the nature and detail of what is being predicted. Our approach to both research aims was informed by a new, emerging conceptualisation of the aphasia phenotype and underlying brain systems. There is a long-standing tradition in aphasiology to categorise patients into different aphasia types according to clusters of behavioural deficits (e.g., Broca, Wernicke, conduction, etc.). These classifications provide an approximate descriptive shorthand for communicating and comparing cases across clinics/research institutions, and influencing treatment options (Horn et al., 2005). There is increasing agreement, however, that aphasia classifications have strong limitations because (a) there is considerable variability amongst patients within each category and (b) there are fuzzy boundaries between categories. Indeed, it is often difficult to place patients within a single category, leading to the diagnosis of “mixed aphasia”. An alternative approach moves away from categorisation and clustering towards considering each patient as a point in a multidimensional space, where each dimension corresponds to a primary computational-brain system (Butler et al., 2014; Chase, 2014; Halai et al., 2017). In this conceptualisation, each patient's pattern of aphasia reflects a different weighting of the impairments to these primary systems. Likewise, each language activity (e.g., naming, comprehending, repeating, etc.) is not localised to a single brain region but rather reflects the joint action of the underpinning primary systems (Patterson and Lambon Ralph, 1999; Seidenberg and McClelland, 1989; Ueno and Lambon Ralph, 2013; Ueno et al., 2011). A simple analogy is that of the arrangement of different colour hues (cf. patients) across the red, green and blue (RGB) colour space. Whilst it is possible to demarcate and label (cf. categorise) approximate areas in the space as yellow (e.g., Broca), blue (Wernicke), etc., there are in fact many different kinds of each colour and the boundaries between them are fuzzy. Likewise, when presented with individual hues it is not always obvious which colour category they fall into (e.g., teal, maroon, indigo; cf. how to categorise a patient with mixed aphasia). Thus, like aphasia classifications, colour labels provide approximate albeit limited information about the underlying graded differences. This is sufficient to communicate broad distinctions between cases (e.g., blue vs. yellow; Broca vs. Wernicke) but not finer variations (the overlapping variations of orange vs. yellow; conduction vs. Wernicke). An alternative and more precise approach is to represent each hue (patient) in terms of its position along the RGB dimensions (cf. patients' performance in terms of the underlying primary language-cognitive systems).

With sufficient breadth of assessments (to sample the full spectrum of language activities) and patient numbers, it is possible to use statistical approaches such as principal component analysis (PCA) to uncover the underlying dimensions (Lambon Ralph et al., 2002; Lambon Ralph et al., 2003). Recent applications of this approach have not only recovered the same set of orthogonal dimensions (phonology, semantics, executive skills, speech quanta) but have found that each one is associated with damage to discrete brain regions (Butler et al., 2014; Halai et al., 2017). Importantly, for the present study, very similar or identical behavioural dimensions and lesion correlates have been observed across independent studies both in patients with chronic (Lacey et al., 2017; Mirman et al., 2015a; Mirman et al., 2015b) and acute aphasia (Kümmerer et al., 2013), indicating the robustness of these core underlying factors.

The ramifications of this aphasia conceptualisation on generating prediction models are as follows. In terms of prediction targets, the ultimate aim is to predict the full behavioural profile of each patient from the neuroimaging data. Thus, rather than focussing on individual language activities, in this study we predicted each patient's scores across the full range of assessments. Given the strong tradition of using aphasia classifications, we also generated a predictive classification model but rather than focussing on pairwise discrimination between pairs of aphasia types, we required the model to discriminate *simultaneously* between all major types (thus providing a full albeit coarse-coding of the aphasic multidimensional space). Secondly, in terms of deriving the best predictors for inclusion in these models, we utilised the finding that the core underlying ‘primary’ dimensions (phonology, semantics, etc.) have been associated with discrete lesion correlates. As such, one might expect the status of each of these key regions to be a strong predictor of the patients' performance across the full range of tests. Accordingly, each patient's lesion was functionally-partitioned according to the overlap with these primary language regions and the resultant four component model was used to predict each patient's individual test scores as well as aphasia classification.

2. Materials and methods

2.1. Participants

Seventy post-stroke patients (53 males, mean age ± standard deviation [SD] = 65.21 ± 11.70 years) were recruited in the chronic stage (minimum 12 months post onset; mean = 56.6, SD = 50.17 months). A subset of cases (31/70) was the same as reported in two previous studies (Butler et al., 2014; Halai et al., 2017). The mean years in education was 12.11 (SD = 2.20). All cases were diagnosed with aphasia (using the Boston Diagnostic Aphasia Examination, BDAE), having difficulty with producing and/or understanding speech. No restrictions were placed according to aphasia type or severity (spanning from global to minimal aphasia). All subjects were right handed (premorbidly) using the Edinburgh Handedness Inventory