Lesion-symptom mapping with NIHSS sub-scores in ischemic stroke patients

Deepthi Rajashekar,1,2,3 Matthias Wilms,2 M Ethan MacDonald,2,4,5 Serena Schimert,2,4 Michael D Hill,2,3,4,6,7 Andrew Demchuk,2,7 Mayank Goyal,2,7 Sean P Dukelow,3,4 Nils Daniel Forkert2,3,4,8

ABSTRACT

Background Lesion-symptom mapping (LSM) is a statistical technique to investigate the population-specific relationship between structural integrity and post-stroke clinical outcome. In clinical practice, patients are commonly evaluated using the National Institutes of Health Stroke Scale (NIHSS), an 11-domain clinical score to quantitate neurological deficits due to stroke. So far, LSM studies have mostly used the total NIHSS score for analysis, which might not uncover subtle structure–function relationships associated with the specific sub-domains of the NIHSS evaluation. Thus, the aim of this work was to investigate the feasibility to perform LSM analyses with sub-score information to reveal category-specific structure–function relationships that a total score may not reveal.

Methods Employing a multivariate technique, LSM analyses were conducted using a sample of 180 patients with NIHSS assessment at 48-hour post-stroke from the ESCAPE trial. The NIHSS domains were grouped into six categories using two schemes. LSM was conducted for each category of the two groupings and the total NIHSS score.

Results Sub-score LSMs not only identify most of the brain regions that are identified as critical by the total NIHSS score but also reveal additional brain regions critical to each function category of the NIHSS assessment without requiring extensive, specialised assessments.

Conclusion These findings show that widely available sub-scores of clinical outcome assessments can be used to investigate more specific structure–function relationships, which may improve predictive modelling of stroke outcomes in the context of modern clinical stroke assessments and neuroimaging.

Trial registration number NCT01778335.

INTRODUCTION

Lesion-symptom mapping (LSM) is a population-specific statistical technique that can be used to investigate the relationship between the post-stroke structural integrity of brain regions and the severity of patient outcome measured by clinical assessments of functional and neurological deficits.1 The resulting LSM displays regions of the brain that are critically associated with the clinical outcome of interest. LSM analyses not only improve the current understanding of the function of different brain areas but also help to identify potential brain regions of interest to be studied in predictive models of stroke outcomes.2 Any improvements in quantifying the relationship between regions of the brain and individual aspects of brain function are likely to improve the precision of stroke outcome prediction. Ultimately, an improved understanding of the structure–function relationship can assist patient-specific management and rehabilitation.

There are several factors that influence the analysis of the relationship between the structural damage caused by the stroke and the observed clinical deficits among patients. These include, but are not limited to: (1) the size of the sample dataset,3 (2) the spatial distribution of the lesions in the study population, (3) the time-point of clinical assessment, (4) the imaging modality (as a proxy for image resolution and specificity), (5) statistical tests employed (assumptions and limitations), (6) confounding clinical variables, and (7) the methods to correct for multiple comparisons.4 While most of these aspects have been investigated in detail, the LSM literature has focused predominantly on specialised and detailed measures of stroke-induced deficits such as the aphasia and spatial neglect batteries5–10 of clinical assessment. However, complex and highly specific assessments are not commonly performed in the clinical routine and are naturally limited to surviving patients who have milder deficits. In contrast, the National Institutes of Health Stroke Scale (NIHSS) is a composite clinical score to quantitate neurological deficits that is administered routinely for clinical stroke assessment. While various LSM studies have been performed using the NIHSS score, only the gross score has been used for this purpose.11,12 By design, composite neurological assessments do not have a uniform weight for the different domains of brain function they assess. For example, the NIHSS is unevenly weighted in that motor and sensory function contribute 23 points of the overall assessment, while level of consciousness accounts for 7 points, language and vision/
oculomotor control contributes 5 points each and spatial neglect is allotted 2 points of the overall score. Thus, the NIHSS has a ceiling effect in its responsiveness, despite its excellent internal consistency and inter-rater reliability. Moreover, the scale is weighted towards the dominant hemisphere because language function has greater representation than hemispatial neglect and as a result left hemisphere strokes, on average, score 5 more points on the scale than right hemisphere strokes. Thus, using the total NIHSS score for LSM analyses might not reveal critical brain regions specific to functions that have less weight towards the total score.

The aim of this study is to investigate if NIHSS sub-score information can be used to identify critical brain regions associated with more specific brain functions, which are normally identified using more specific and complex assessment scores.

**MATERIALS AND METHODS**

**Datasets**

The datasets used in this study were sampled from the multicentre, double-blind, randomised controlled trial ESCAPE, which included patients who had an acute ischaemic stroke with an occlusion of the intracranial carotid artery or proximal middle carotid artery, who obtained a follow-up MRI or CT scan and had a baseline NIHSS >5. Any patients with severe white matter hyperintensities, bilateral strokes or remote haemorrhagic transformations were excluded from this secondary analysis. Of the 315 patients enrolled in the original trial, 180 patients met the inclusion criterion for this retrospective study. Due to the exploratory nature of this secondary analysis of the data, no power calculation was performed. This subset consisted of 84 men, had an average age of 69.1±13.3 years, average lesion size of 91.3±79.9 mm³ and a median 48-hour NIHSS of 9 (IQR: 4–17.5). Since the data from the trial was acquired in multiple centres with different imaging modalities (non-contrast CT and fluid-attenuated inversion recovery MRI), the scanning protocols varied by site. The in-slice spatial resolution of the NCCT datasets varied from 0.36×0.36 mm² to 0.55×0.55 mm², while the slice thickness was in the range of 2–10 mm. The follow-up FLAIR datasets were acquired with an in-slice spatial resolution of 0.43×0.43 mm² to 1.4×1.4 mm² and a slice thickness ranging from 3 mm to 7.5 mm. Online supplemental table S1 summarises the patient characteristics per imaging modality.

The clinical assessment of interest for this study is the 48-hour NIHSS with the sub-score information for each of the 11 NIHSS domains. Since these 11 domains are not mutually exclusive and stroke-induced deficits co-occur, two plausible categorisations were considered in this work. The first scheme aligns with the design of the NIHSS evaluation alone and is referred to as ‘score-based grouping’. The second scheme was developed based on the subjective expertise of the collaborating, local stroke neurologists and rehabilitation experts. This design was motivated by clinical aspects of stroke presentation, considering the neuroanatomy of stroke, and frequently co-occurring deficits among patients with a stroke in the anterior circulation and is referred to as ‘anatomy-based grouping’ in the remainder of this article. For instance, in the anatomy-based grouping, the level of consciousness components that pertain to understanding and responding to commands (1b and 1c) are grouped under the language category. The summary of sub-score categorisation schemes is presented in table 1 and follow the definition of the 11-item modified NIHSS scale.

**Lesion-symptom mapping**

**Preprocessing**

First, lesions were segmented in all datasets by an experienced observer using ITK-SNAP. After this, individual patient scans were non-linearly registered to the age-appropriate MIPLAB FLAIR and NCCT atlas with cost function masking of the segmented lesion using the ANTs toolkit. Registration was performed in three stages, employing rigid, affine and non-linear transformations using mutual information as the similarity metric. The resulting composite transformation was then applied to the binary lesion masks to map the native lesions to the atlas space. This registration of all lesions to a common co-ordinate space corrects for differences in image acquisition allowing for voxel-level statistical analysis. All registrations were visually assessed for quality control of the mapping of individual scans to the MIPLAB atlas, and all subsequent computations were done in the MIPLAB atlas space.
imaging-based adaptation of the Canonical Correlation Analysis (CCA) method, which is a commonly employed dimensionality reduction technique for multimodal data. In this context, the CCA and, therefore, the SCCAN can be understood as extensions of the well-known principal components analysis to two sets of variables.

In our setup, SCCAN allows computing a linear projection of the lesion data so that Pearson’s correlation between the projected lesion data and the outcome score of interest is maximised. This enables an estimation of the population-level relationship between these two entities, thus, defining the structure–function relationships from stroke-induced injury. Formally, SCCAN as used here optimises the following function:

$$\omega = \max_{\omega} \left( L^T \omega, s \right) \text{such that } \sum_i \| \omega_i \| \leq \epsilon \text{ and } \forall \| \omega_i \| \geq 0.$$

Here, $L$ is a random vector representing the lesion data, $s$ is the random variable associated with the score of interest, and $\omega$ is a vector of optimal weights. The two constraints are necessary to enforce sparseness and positivity, which will ensure that non voxel is negatively associated with the score. In summary, the SCCAN adapts the CCA framework to the sparse lesion data to maximise the correlation between the linear combination of voxels and the outcome score. This results in a weight for each voxel location with higher weights indicating a critical association of that voxel with the outcome score. The result of one run of SCCAN is a map of voxel weights normalised between 0 and 1, referred to as ‘eloquence score’ in the remainder of this article.

**Experiment**

A separate LSM analysis was then conducted for each of the six NIHSS categories from the score-based and anatomy-based groupings and the total NIHSS score, resulting in 13 unique LSMs, using the same 180 patients in all 13 analyses. To ensure that isolated large vessel occlusions do not dominate the lesion–symptom relationship, voxels that were not lesioned in at least 10% of the dataset were excluded, which is a default setting in LESYMAP. The SCCAN corrects for lesion volume as a covariate in the sparse regression model with a significance threshold of 0.05 and sparseness value of 0.045 (default).

**Regions of critical association**

The regions critically associated with the category-specific NIHSS sub-scores were identified using the parcellations defined in the AALCAT atlas (https://www.nitrc.org/projects/niistat/). The AALCAT atlas is a fusion of 116 grey matter regions from the AAL atlas and 34 white matter tracts from the CAT atlas provided in the MRIcron toolbox. In this work, the AALCAT atlas was mapped from its native space to the MIPLAB atlas space.

Subsequently, the volume overlaps of the 150 AALCAT brain regions with the LSM for each NIHSS category were calculated. Additionally, the average of all non-zero voxel weights was computed for each brain region resulting in an eloquence score for each region. It is important to note that the same severity of deficit can be caused by lesions of varying volume in different parts of the brain. In other words, a region with a large volume overlap might result in the same deficit as smaller insult in another brain region. Therefore, it is necessary to analyse the LSM results not only by the volume of overlap with a brain region but also by considering the average eloquence score per region. For simplicity of the interpretation and to reduce the impact of false positives, critical regions and tracts are defined in this work as those that are in the highest quartile (Q4) in either volume overlap or eloquence score for each individual LSM result.

**RESULTS**

**Lesion distribution**

The overlap of all individual lesions in this dataset (in the MIPLAB space) is mostly characterised by lesions in the middle cerebral artery territory (see figure 1). Among the grey matter structures, maximum lesion overlap is found in the insular gyrus, basal ganglia (insult in putamen, caudate and pallidum in both hemispheres), superior temporal sulcus, Rolandic operculum and Heschl’s gyrus. The most frequently affected white matter tracts are the inferior occipital frontal fasciculus, internal capsule and uncinate fasciculus. The most frequently affected white matter projections are the corticospinal tract and optic radiations.

**LSM for NIHSS total score**

Although the LSM for the total NIHSS scale (total map) has comparable volumes of critical regions in both hemispheres (54.2% of the total map in the left hemisphere and 45.8% in the right hemisphere), the average eloquence values are relatively higher in the left hemisphere (figure 2). The critical regions identified include the left anterior segment of the arcuate fasciculus (0.37±0.21), left postcentral gyrus (0.34±0.18), left long segment of the arcuate fasciculus (0.32±0.15), left cortico-ponto cerebellar tract...
(0.32±0.15), bilateral arcuate fasciculi (0.3±0.17 in left and 0.19±0.06 in right hemisphere) and the right corticospinal tract (0.19±0.06).

**LSM for sub-scores in score-based grouping**

The LSM results for level of consciousness, language, motor, sensory, vision and neglect are summarised in online supplemental table S2 and illustrated in figure 3. In this scheme of sub-score categorisation, the critical brain regions identified by LSM for level of consciousness and language are only in the left hemisphere, with the language LSM identifying five out of the nine regions found in the level of consciousness LSM results (see figure 3A). The left insular gyrus is the only critical region exclusive to the language LSM result. Both the motor and sensory LSM results (figure 3B) are characterised by a dominance of white matter tracts in the right hemisphere. The vision LSM result (figure 3C) shows clusters of similar volume in both hemispheres. Conversely, the neglect LSM result indicates only right hemispheric tracts.

**LSM for sub-scores in anatomy-based grouping**

The LSM results for domains of level of consciousness, language, motor, sensory, vision and ataxia are summarised in online supplemental table S3 and illustrated in figure 4. In this scheme of sub-score categorisation, the LSM result for level of consciousness mostly identified critical brain regions in the right hemisphere except for the left corticospinal tract (contrary to score-based grouping). The language LSM identified left hemispheric brain regions that are mostly similar to the results for the score-based grouping except for the absence of left anterior segment, arcuate, and addition of corpus callosum (figure 4A). The motor and sensory LSM results (figure 4A) are also similar to the results of the first grouping characterised by a dominance of white matter.
tracts in the right hemisphere. The vision LSM result (figure 4C) shows clusters of similar volume in both hemispheres. Additionally, the LSM result for ataxia also identified primarily right hemispheric brain regions except for the left corticospinal tract.

**Comparison of category-specific maps and total NIHSS map**

Of the seven critical regions identified in the total NIHSS score LSM result, at least two regions were identified in the LSM results for each category according to the score-based and anatomy-based groupings. Figure 5 shows the total NIHSS score LSM result, which appears to be a subset of the union of all the NIHSS category-specific maps referred to as the cumulative map.

**DISCUSSION**

The main finding of this study is that, in a clinical acute ischaemic stroke population, unique LSM can be generated for each sub-domain of the NIHSS score. This approach not only captures the known structure–function relationships with respect to the total score but also reveals category-specific critical associations of brain regions to each domain of the composite assessment (level of consciousness, language, motor, sensory, vision, ataxia and hemispatial neglect) that may not be appreciated when examining the total score alone. Even though the NIHSS is heavily weighted towards motor function, the NIHSS domains that have a rather small representation in the total NIHSS score (such as the hemispatial neglect or ataxia sub-domains) can be mapped using LSM to identify category-specific structure–function relationships, allowing quantitative neuroimaging techniques to make the most of the existing and widely available clinical assessment scores.

It is important to consider that of all critical regions in the NIHSS total score LSM results, the left right arcuate and corticospinal tracts also occur in the motor, sensory, and neglect LSM results. Contrary to this, the left postcentral gyrus was not identified as critical in any NIHSS category (of both score-based and anatomy-based groupings) but had a high average
eloquence score in the total NIHSS LSM result. This indicates that regardless of the granularity of the target score, the structure–function relationship with respect to the total NIHSS scale is a cumulative effect from each of the category-specific LSMs.

A less obvious observation is that the category-specific LSM results that use NIHSS sub-score information identify additional brain regions that are associated with the NIHSS score. These additional regions include the bilateral insula, Rolandic operculum, Heschl’s gyrus, posterior segment of the arcuate fasciculus, right internal capsule, caudate, pallidum, anterior segment of the arcuate fasciculus, superior temporal gyrus, arcuate fasciculus, supramarginal gyrus, cortico-ponto cerebellum, inferior occipitofrontal fasciculus, and corpus callosum.

Heschl’s gyrus was shown to be critical to level of consciousness and language based on the score-based grouping and only to language in anatomy-based grouping (figures 3A and 4A). This finding is concordant with prior work examining consciousness through the auditory pathways, correlating the volume of Heschl’s gyrus with the ability for auditory comprehension and language learning, with a particular emphasis on its role in distinguishing pitch. The association of Heschl’s gyrus with language comprehension is also reported in prior work in the LSM literature. The arcuate fasciculus’ association with hemisphere, the arcuate fasciculus is important for linguistic efficiency, and language-based groupings of the NIHSS domains used in this work are by no means an exhaustive categorisation scheme. Multiple grouping strategies are plausible and largely determined by study design. However, the important inference here is that regardless of the grouping scheme, utilising sub-score information reveals additional structure–function relationships that are clinically meaningful in comparison to using the total NIHSS alone (figure 5).

In a more general context, the proposed framework is also relevant in stroke recovery and rehabilitation studies where the clinical assessments cover a wider range of post-stroke deficits and measure patient recovery precisely and repeatedly, allowing for a longitudinal study design. The framework could also be adapted to other established impairment outcome measures such as the Montreal Cognitive Assessment or those derived from rehabilitative therapies that treat localised impairment. Overall, computational neuroimaging research of this nature has the potential to uncover novel information about neural correlates of stroke-induced deficits that could help re-structure the clinical evaluation of patients who had a stroke and improve predictive models of long-term stroke outcome.

**CONCLUSION**

In summary, this study shows that the LSM from the total NIHSS score captures a cumulative effect of all the NIHSS categories and not just the post-stroke motor deficits. The results also indicate that using sub-score information reveals function-specific aspect of structure–function relationships seems to be observable already early in the standard stroke care using clinical assessments like the NIHSS sub-scores that are readily available, resulting in a possibility of an improved clinical decision-support framework.
critical brain regions that are not identified using the NIHSS total score alone for LSM analysis. Thus, using sub-scores of widely available assessment scores for LSM analyses is feasible and can improve our understanding of clinical stroke outcome.

Author affiliations
1Biomedical Engineering Graduate Program, Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada
2Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
3Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada
4Department of Clinical Neurosciences, University of Calgary, Calgary, Alberta, Canada
5Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, Canada
6Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
7Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
8Alberta Children’s Hospital Research Institute, University of Calgary, Calgary, AB, Canada

Acknowledgements We thank the Hotchkiss Brain Institute at the University of Calgary (in partnership with Alberta Health Services) for all the support and encouragement in fostering collaborative brain research.

Contributors DR—conceptualisation, data curation, formal analysis, investigation, methodology, validation, visualisation, roles/writing—original draft, software, gaurantor; MW—formal analysis, methodology, writing—review and editing, validation; MEM—writing—review and editing, visualisation; SS—writing—review and editing; MDH—conceptualisation, data curation, writing—review and editing; AMD—resources; MG—resources; SPV—writing—review and editing; NDF—conceptualisation, project administration, supervision, writing—review and editing, funding acquisition.

Funding This work was funded by the Heart and Stroke Foundation of Canada Grant in aid (G-17-0018368), the Canada Research Chairs program, and the River Fund at Calgary Foundation.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This is a secondary study that pools retrospective data from the ESCAPE trial. All the datasets used in this study were made available only after complete anonymisation.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplemental information. Not applicable.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

Acknowledgements

Deepthi Rajashekar http://orcid.org/0000-0003-2009-764X
Nils Daniel Forkert http://orcid.org/0000-0003-2556-3224

REFERENCES