

# BRAIN COMMUNICATIONS

## REVIEW ARTICLE

# Interpreting and validating complexity and causality in lesion-symptom prognoses

 Mohamed L. Seghier<sup>1,2</sup> and Cathy J. Price<sup>3</sup>

This paper considers the steps needed to generate pragmatic and interpretable lesion-symptom mappings that can be used for clinically reliable prognoses. The novel contributions are 3-fold. We first define and inter-relate five neurobiological and five methodological constraints that need to be accounted for when interpreting lesion-symptom associations and generating synthetic lesion data. The first implication is that, because of these constraints, lesion-symptom mapping needs to focus on probabilistic relationships between Lesion and Symptom, with Lesion as a multivariate spatial pattern, Symptom as a time-dependent behavioural profile and evidence that Lesion raises the probability of Symptom. The second implication is that in order to assess the strength of probabilistic causality, we need to distinguish between causal lesion sites, incidental lesion sites, spared but dysfunctional sites and intact sites, all of which might affect the accuracy of the predictions and prognoses generated. We then formulate lesion-symptom mappings in logical notations, including combinatorial rules, that are then used to evaluate and better understand complex brain-behaviour relationships. The logical and theoretical framework presented applies to any type of neurological disorder but is primarily discussed in relationship to stroke damage. Accommodating the identified constraints, we discuss how the 1965 Bradford Hill criteria for inferring probabilistic causality, *post hoc*, from observed correlations in epidemiology—can be applied to lesion-symptom mapping in stroke survivors. Finally, we propose that rather than rely on *post hoc* evaluation of how well the causality criteria have been met, the neurobiological and methodological constraints should be addressed, *a priori*, by changing the experimental design of lesion-symptom mappings and setting up an open platform to share and validate the discovery of reliable and accurate lesion rules that are clinically useful.

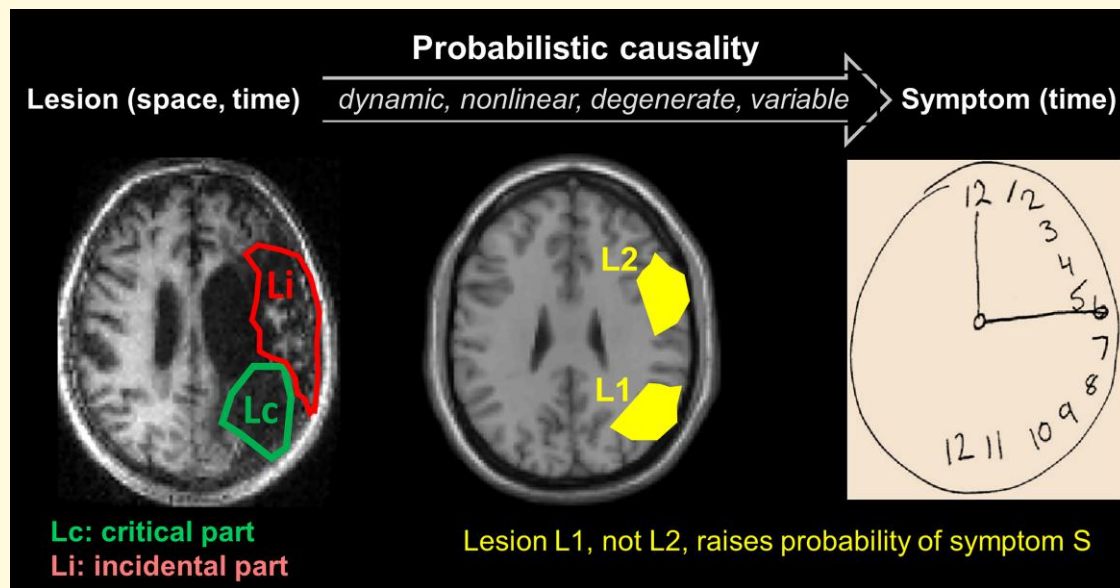
- 1 Department of Biomedical Engineering, Khalifa University of Science and Technology, Abu Dhabi, UAE
- 2 Healthcare Engineering Innovation Center (HEIC), Khalifa University of Science and Technology, Abu Dhabi, UAE
- 3 Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, London, UK

Correspondence to: Mohamed Seghier  
Department of Biomedical Engineering  
Khalifa University of Science and Technology  
PO BOX: 127788, Abu Dhabi, UAE  
E-mail: mseghier@gmail.com

**Keywords:** lesion-symptom mapping; prognoses; brain-behaviour relationships; probabilistic causality; logic

**Abbreviations:** BH = Bradford Hill; L = lesion; Lc = critical part of a lesion; Li = incidental part of a lesion; LSM = lesion-symptom mapping; MC = methodological constraint; NC = neurobiological constraint; S = symptom

## Graphical Abstract



## Introduction

Here, we review a range of issues related to causality and validity in lesion-symptom mapping (LSM) and how such issues can be addressed when using neuroimaging data to generate valid and unbiased predictions of outcome and recovery following brain damage (i.e. lesion-symptom prognoses). Causation in LSM has been discussed and examined in previous reports (e.g. Godefroy *et al.*<sup>1</sup>, Toba *et al.*<sup>2</sup>, Siddiqi *et al.*<sup>3</sup> and Sperber<sup>4</sup>) with an increasing appreciation of the importance of understanding the neurobiological mechanisms that influence how lesions produce symptoms. Validity, however, has been overlooked in this literature despite the fact that valid reasoning is critical for understanding the direction of causal effects and formulating predictions and prognoses.

Our paper builds on, and integrates, several lines of previous work. For example, Siddiqi *et al.*<sup>3</sup> offered a multi-criteria continuum for assessing causality in findings from a range of brain imaging techniques and reported that LSM with naturally occurring lesions (e.g. stroke lesions) has moderate causality strength (a score of 4.5 out of 6.5 along a six-criterion scale) compared to other approaches, such as targeted brain stimulation. Other researchers have highlighted the graded nature of causation in complex biological systems where multiple factors contribute to the occurrence and severity of an effect.<sup>5</sup> More generally, it has been proposed (e.g. Eells<sup>6</sup>, Mitroff and Silvers<sup>7</sup>, Sobel<sup>8</sup> and Sprenger<sup>9</sup>) that it might be more intuitive to measure the strength of causal associations probabilistically; i.e. how likely it is that an effect will be associated with a cause.

Here, we go beyond prior reports by focusing on the logic and validity of causality in LSM. We consider the

neurobiological and methodological factors that interact to produce a particular LSM outcome, unravel some inherent constraints that lessen the strength of causality in LSM findings, adapt prior criteria for assessing causality to make them relevant to the interpretation of lesion-symptom relationships and emphasize the need for probabilistic lesion-symptom prognoses. Although the theoretical framework discussed here applies to any type of neurological events that can cause a behavioural change, the majority of our examples concern the effect of stroke damage.

The organization of the paper is as follows: section 'Causality and complexity in lesion-symptom mapping (LSM)' describes how the deduction of causality in LSM is confounded by multiple levels of complexity that result in uncertainty and necessitate the formulation of probabilistic rather than binary associations. Section 'The logic of combinatorial lesion-symptom mapping and its validity' discusses how causal relations can be made stronger if supplemented by the application of logic and a deeper appreciation of the factors that determine validity. Section 'Inferring causality, *post hoc*, along a multi-criteria continuum' considers how causality can be deduced, *post hoc*, by applying a multi-criteria framework. Section 'Implications for future LSM studies' discusses potential implications for future LSM studies.

## Causality and complexity in lesion-symptom mapping (LSM)

LSM typically examines causal brain-behaviour relationships between a lesion L (focal damage or a multisite pattern)

and one or more symptoms *S* from observational data. Lesion *L* is considered causal if, with all else held constant, its presence or absence affects the probability of observing a symptom *S*. Put another way, a direct causal relationship in LSM means that a change in *L* is sufficient to produce a change in *S*, without the operation of intermediate causes. A comprehensive framework is required to examine, interpret and validate the conditional statement  $L \rightarrow S$  (where the arrow  $\rightarrow$  indexes the direction of cause-and-effect). This section discusses how the application of  $L \rightarrow S$  rules, and conclusions about causality, are confounded by multiple levels of complexity. Below, we identify five levels of complexity that arise from neurobiological constraints (NCs 1–5) and five levels of complexity that arise from methodological constraints (MCs 1–5) when making casual inferences at the group level.

**NC-1: spatial dependency**—In humans, LSM deals with naturally occurring lesions that result from a particular pathology (Fig. 1) rather than an experimental intervention (as in animal models or with targeted stimulations). It is not possible to fully or tightly control for the location and extent of naturally occurring lesions because pathological lesions are constrained by other factors such as the vascular architecture of the brain,<sup>10–12</sup> whether a stroke is ischaemic or haemorrhagic,<sup>13,14</sup> and the type of neurodegeneration.<sup>15</sup> When naturally occurring lesions are large or distributed, only a small part of the damaged region (lesion) may be causing the symptoms. We refer to this as the critical part of the lesion (*Lc*) and distinguish it from other parts of the lesion (*Li*) that are incidental to the symptoms of interest (Fig. 2). Separating *Lc* from *Li* is not a trivial problem because *Li* may be significantly correlated with symptoms when damage to *Li* is tightly correlated with damage to *Lc*.<sup>11</sup> Spatial dependency refers to the correlation between the occurrence of *Lc* and the occurrence of *Li*.<sup>12</sup> It results in the spurious correlation of *Li* with *S*, which contributes false positive effects to the results.

**NC-2: distributed neural systems**—It is well appreciated that motor and cognitive functions are supported by distributed sets of interacting brain regions that are referred to as neural systems, neural networks or neural circuits.<sup>16,17</sup> If each function relies on multiple brain regions, then damage to any part of the neural system could theoretically disrupt function and cause the same symptoms (Fig. 3A). This can result in false negative results in group-level LSM experiments.<sup>18</sup> For example, in a simple case where two different brain regions (*X* and *Y*) support the same function, patients identified by symptoms only may include a mix of those who have [*X* intact/*Y* damaged] and [*Y* intact/*X* damaged]—which results in high error variance, in both *X* and *Y*, when using univariate (e.g. voxel based) LSM.

**NC-3: time-dependency**—Even after a sudden neurological event, lesions continue to change over time in both the short-term<sup>19</sup> and long-term.<sup>20,21</sup> Symptoms also change with time in line with functional recovery (e.g. after stroke/neurosurgery)

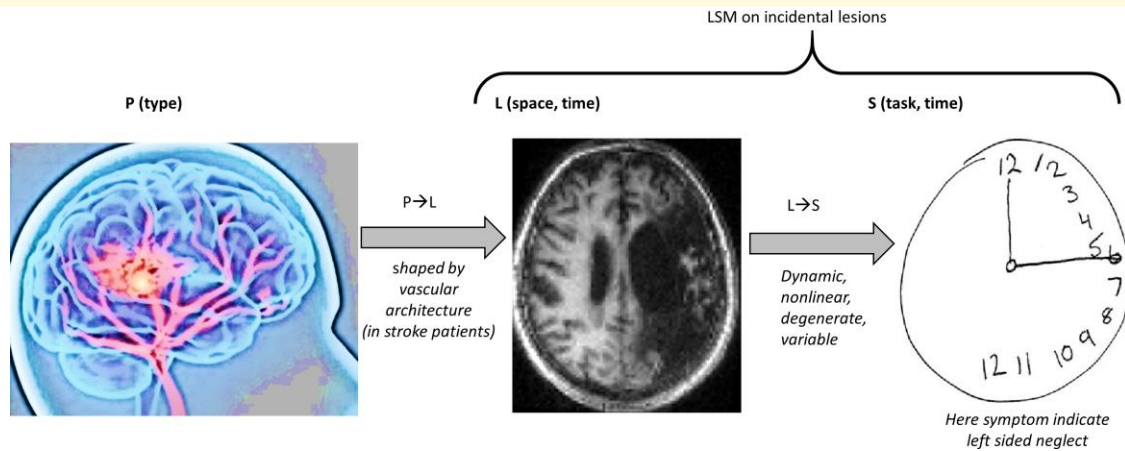
or functional decline in neurodegenerative disorders. The consequence of time-dependency in *L* and/or *S* is that significant *L*-*S* mappings may be observed at one timepoint but not another (Fig. 3B). In the case of stroke, for example, significant *L*-*S* mappings are more likely to be observed early post-stroke than after time has allowed for substantial plasticity, reorganization and recovery.<sup>22–24</sup> Indeed, an increasing number of studies are showing that brain reorganization and recovery occur even years after stroke.<sup>25–29</sup> Depending on the question and degree of recovery/decline, time-dependency will result in both false negative and false positive results in both univariate and multivariate LSMs.

**NC-4: degeneracy in structure–function relationships**—Lesion-symptom relationships may vary across patients when the same ability (function) can be sustained by different sets of brain regions (neural systems) in different individuals or at different times in the same individual. We refer to this type of pre-morbid inter-patient variability as ‘degeneracy’ and argue that it underpins one of the mechanisms that enable recovery of function after brain damage.<sup>30,31</sup> The implication for group-level LSM (univariate or multivariate) is that degeneracy can result in false negative results because damage to the same region will be associated with symptoms in some patients but not others, depending on pre-morbid preferences for one neural system versus another.

**NC-5: inter-individual variability in topography**—Even after spatial normalization to a common template, there will be inter-individual variability in the topography of cortical regions and white matter tracts. For example, a particular brain region in one patient will not necessarily fall within a template parcellation for that region (see discussion in Bryce *et al.*<sup>32</sup> and Moghimi *et al.*<sup>33</sup>). As with degeneracy, this variance could lead to false negative LSM results. This type of inter-individual variability is different from degeneracy (NC-4) that can arise not because of variance in the location of an anatomical structure, or its function, but in differences in how the functions are being recruited to perform a task either within patients in different contexts or between individuals.

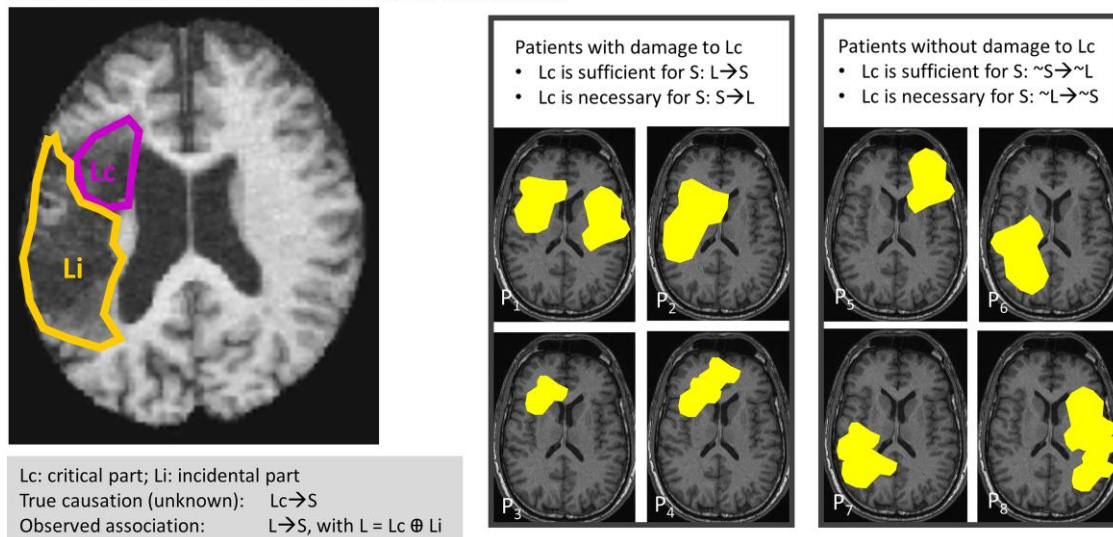
**MC-1: inaccurate lesion definition**—In currently available brain scans, there is uncertainty in the definition of which parts of the brain are lesioned (noted *L*) or not lesioned (noted  $\sim L$ , where  $\sim$  is the logical negation ‘not’). Consequently, some intact voxels might be misclassified as lesioned ( $\sim L$  misclassified as *L*), and some abnormal voxels might be omitted from the lesion definition (*L* misclassified as  $\sim L$ ). One way to minimize uncertainty in the definition of *L* and  $\sim L$  is to use continuous lesion definitions<sup>18,34,35</sup> rather than the more commonly used binary lesion definitions (*L* or  $\sim L$ ) adopted by most LSM studies. In addition to uncertainty in the definition of *L*, inaccuracies can also arise from suboptimal spatial warping of lesioned brains.<sup>36,37</sup>

**MC-2: functional dysfunction**—The absence of a visible/detectable lesion does not necessarily indicate healthy normally



**Figure 1 Conditional relationships between pathology, lesion and symptoms.** An illustration of the causal relationships between pathology (P), lesion (L) and symptoms (S). The mapping from P to L is complex, shaped by the brain's vascular architecture and dependent on space and time; type of pathology P can for instance be neurosurgery, ischaemic or haemorrhagic stroke, etc. The mapping from L to S is dynamic (varies with recovery/decline), nonlinear, governed by the rules of degeneracy and variable across patients and depends on many non-lesion confounders.

Spatially-dependent lesion parts can be associated with symptom



**Figure 2 Hypothetical spatially-dependent lesion-symptom relationships.** Left: each lesion is typically an aggregate of many small subregions or thousands of voxels, and it can be expressed at different spatial granularity. Here, the lesion site for a given patient is illustrated as a combination of (i) the critical part (Lc; a small polygon outlined in purple) that caused symptom S; and (ii) other incidental parts (Li; a large polygon outlined in orange) that did not cause the symptom but co-occurred with Lc (e.g. due to the vascular organisation of the brain). The full lesion is described as  $[Lc \oplus Li]$  where  $[\oplus]$  is the exclusive OR (xor) operator—meaning that Li is not part of Lc for symptom S. Li may, nevertheless, cause other types of symptoms and be further divided into Lc and Li for each of these other symptoms. The true causal relationship  $Lc \rightarrow S$  is not known as the researcher can only assess the following relationship:  $[Lc \oplus Li] \rightarrow S$ . Right: critically, Lc cannot be distinguished from Li without establishing how different parts of the full lesion site  $[Lc \oplus Li]$  are associated with symptoms in many other patients. To identify 'sufficient' brain structures, LSM (in large samples) is required to identify a candidate Lc region (set of voxels) that satisfies two criteria as follows: (i) ALL patients with damage to Lc have symptom S; i.e.  $[Lc_1 \rightarrow S] \wedge [Lc_2 \rightarrow S] \wedge [Lc_3 \rightarrow S]$ , etc.; or (ii) ALL patients with no symptom S do not have lesion Lc; i.e.  $[\sim S \rightarrow \sim Lc_5] \wedge [\sim S \rightarrow \sim Lc_6] \wedge [\sim S \rightarrow \sim Lc_7]$ , etc. To identify 'necessary' brain structures, a candidate Lc region needs to satisfy two criteria as follows: (i) ALL patients with symptom S have damage to Lc; i.e.  $[S \rightarrow Lc_1] \wedge [S \rightarrow Lc_2] \wedge [S \rightarrow Lc_3]$ , etc.; or (ii) ALL patients with no damage to Lc do not have symptom S; i.e.  $[\sim Lc_5 \rightarrow \sim S] \wedge [\sim Lc_6 \rightarrow \sim S] \wedge [\sim Lc_7 \rightarrow \sim S]$ , etc. It is expected that Lc is highly similar across patients with S (Lc<sub>1</sub>, Lc<sub>2</sub>, Lc<sub>3</sub> ... etc.), with differences within typical ranges of inter-individual variability in brain anatomy, whereas Li (Li<sub>1</sub>, Li<sub>2</sub>, Li<sub>3</sub> ... etc.) is expected to be highly variable across patients. Ultimately, the goal is to establish that: (i)  $Lc \rightarrow S$  is true; and (ii)  $\sim S \rightarrow \sim Lc$  is true, and  $Li \rightarrow S$  is false. This is not possible in a single patient or a small sample of patients.



The awareness of the challenges associated with inferring causality is not new in the LSM literature. Seven decades ago, Lashley<sup>49</sup> observed in animal models that behavioural deficits following brain lesions are frequently reversible, by alteration of input, context or additional brain lesions, implying a substantial degree of interaction between the distributed processing systems. The field at that time was converging to the following conclusion: the nervous system displays a level of interconnectedness and interdependence that makes it impossible to presume that a missing function is attributable to the removal of circuitry necessary for that function.

Grobstein<sup>50</sup> appraised the work of Lashley and others and concluded that: *'it seems fairly clear that damage to one part of the nervous system is altering the characteristics of a distant part so that circuitry adequate to form a particular function is not expressed in behaviour'* and therefore *'lesion analysis can be used to prove that a circuit is adequate to support a behavior but not that it is necessary'*. In terms of the neurobiological constraints discussed above, Grobstein's conclusions are correct when there is time-dependency (NC-3), degeneracy (NC-4) or context-dependency (NC-5) in the LSM. However, there are also cases when it is possible to deduce the necessity of brain regions/circuits, for example, when the mapping between L and S is not dependent on time and does not vary across individuals. In line with Grobstein's caution when inferring the 'necessity' of a brain region, we raise four additional points: first, if symptoms after a lesion only persist until surviving circuitry adapts to support recovery of intact behaviour, do we still refer to the lesioned brain region as 'necessary'? Second, although a symptom/functional impairment may persist after small lesions to the brainstem, or corticospinal tract, naturally occurring lesions resulting in persistent impairments (i.e. are not subject to time or context dependency or degeneracy) are more likely to be those that are very large, making it difficult to determine which part of the lesion is necessary (NC-1 above). Third, concluding that a region is necessary for a 'function' does not imply that a lesion is necessary for a 'symptom' because when a neural circuit includes multiple regions, each of which is necessary (distributed neural systems), symptoms can be caused by multiple different lesion sites (NC-2 above). Finally, the term 'necessary' can be misleading when applying formal logic (as we do in the next sections) because identifying which regions are necessary does not imply that all other regions are not necessary.<sup>51</sup> Given these challenges with the term 'necessary', we opt here to use the term 'critical' that simply implies that the lesion will have an effect on functional processing that might result in symptoms at some point after onset.

In summary, this section has defined multiple levels of neurobiological and methodological complexity in LSM that weaken the strength of the causal inferences that can be made from observational data. Causality can be defined as deterministic or probabilistic,<sup>52,53</sup> but deterministic or true causation ( $L \rightarrow S$ , L causes S) is extremely difficult to prove with LSM, if not impossible, due to the constraints outlined above. We conclude this section with the theoretical

and practical implications for dealing with complexity in LSM and generating synthetic data.

Like others, we propose that LSM is best represented in terms of probabilistic causal relations<sup>6–8</sup> that substitute the statement 'L always causes S' with 'L's occurrence increases the probability of S', i.e.  $p(S|L) > p(S|\sim L)$ . Practically, this means that a lesion is considered more relevant to causality if it raises the probability of S.<sup>9</sup> Given that naturally occurring lesions include multiple (critical and incidental) parts, the strength of causality needs to be considered separately for each part. This has been discussed in other domains under the framework of causal contribution<sup>54</sup> that implies that a given site  $L_1$  is a stronger cause of S than site  $L_2$  if and only if  $L_1$  makes S more expected than  $L_2$ . The implication for LSM is to consider L as a set of different parts with variable causality strength, varying from very strong in case of  $L_c$  to very low in case of  $L_i$  of Fig. 2.

The probabilistic nature of brain-behaviour relationships, where the strength of causal associations varies with different parts of the lesion, also has implications for testing LSM with synthetic lesion data. Many studies<sup>11,55,56</sup> have investigated the robustness of different LSM approaches using simulated data, considering the influence of spatial dependency (NC-1) and/or distributed neural systems (NC-2). The greater challenge will be to generate synthetic lesion data that simultaneously reflects all five neurobiological constraints and all five methodological constraints discussed in this section. By specifying these 10 criteria, we hope to motivate the generation of future synthetic lesion data that provides realistic explanations to black box findings.

Overall, by highlighting the constraints on LSM, our intention is to motivate users of LSM to (i) refrain from overstating 'true' causations; (ii) qualify the use of 'necessity'; and (iii) adopt more pragmatic (and probabilistic) ways of understanding and inferring causality. We have also highlighted the need to distinguish between different types of L ( $L_c$ ,  $L_i$ , and  $\sim L$ ) and how the relationship between L and S ( $L \rightarrow S$ ) should ideally be expressed as  $L(\text{space,time}) \rightarrow S(\text{time})$ , see Fig. 1.

## The logic of combinatorial lesion-symptom mapping and its validity

Framing LSM in terms of logical symbols and notations is useful for clarifying, simplifying, validating and rigorously analysing the inferences that can be deduced from complex relationships. Logic can also help to identify and avoid fallacious reasoning, such as circular arguments. It can be used *post hoc* to evaluate existing observations and theoretical brain-behaviour relationships, and it is also expected to be useful for the *a priori* design of new experiments and for programming data-driven AI approaches to LSM. After reviewing the logical deductions that can be drawn from LSM, we highlight examples of when the application of logic is useful

for clinical prognoses and formalizing theoretical frameworks.

Following the laws of conditional proof, LSM has typically been used to infer a ‘causal’ rule between lesion  $L$  and symptom  $S$ . The conditional statement,  $L \rightarrow S$ , means that the presence of a lesion is sufficient to cause the presence of a symptom. It can be reframed in three ways: the logically equivalent contrapositive statement is expressed as  $\sim S \rightarrow \sim L$  (if no symptom, then no lesion); the inverse is expressed as  $\sim L \rightarrow \sim S$  (if no lesion then no symptom); and the converse is expressed as  $S \rightarrow L$  (a symptom infers the presence of a lesion—which can only be the case when  $L$  is necessary to cause  $S$ ).

In the context of distributed neural systems, when multiple distributed regions work together to support the same function (neurobiological constraint 2 in the previous section), the knowledge of all the possible lesion sites that have the potential to cause a symptom is not necessary to validate the conditional statement ( $L \rightarrow S$ ) or its contrapositive ( $\sim S \rightarrow \sim L$ ). The statement  $[L \rightarrow S]$  simply implies that  $L$  is one part of a distributed system associated with a symptom. Likewise, the statement  $[\sim S \rightarrow \sim L]$  simply implies that all critical lesion sites that have the potential to cause a symptom have been preserved. It is more difficult, however, to validate the converse  $[S \rightarrow L]$  and inverse  $[\sim L \rightarrow \sim S]$  statements, because the presence of a symptom doesn’t imply the presence of a specific lesion site—it only implies damage somewhere in the distributed system. Likewise, the absence of a lesion site of interest ( $\sim L$ ) doesn’t imply the absence of a symptom ( $\sim S$ ) because the symptom could be caused by another lesion site. Therefore, in order to validate the statements  $S \rightarrow L$  and  $\sim L \rightarrow \sim S$ , we need to (i) identify all possible lesion sites associated with a symptom; and (ii) define  $\sim L$  as the preservation of all lesion sites associated with a symptom. This can be framed in Boolean terms using the operator conjunction ( $\wedge$ ) to imply that ‘all are present’; and the operator disjunction ( $\vee$ ) to imply that either one or another is present. For example, if a neural system includes regions  $X$  and  $Y$ , the (disjunctive) expression  $L_x \vee L_y$  is used to indicate that damage to either  $X$  or  $Y$  (or both) can yield the symptom. In other words, the statement  $S \rightarrow L$  can be re-expressed as  $[S \rightarrow (L_1 \vee L_2 \vee L_3 \vee \dots L_n)]$ ; and the statement  $\sim L \rightarrow \sim S$  can be re-expressed as  $[(\sim L_1 \wedge \sim L_2 \wedge \sim L_3 \wedge \dots \sim L_n) \rightarrow \sim S]$ .

Degeneracy (neurobiological constraint 4 in the previous section) can also be expressed in Boolean terms. For example, when a function of interest can be supported by either region  $X$  or region  $Y$ , the (conjunctive) expression  $[(L_x \wedge L_y) \rightarrow S]$  can be used to indicate that the ‘lesion’ associated with a symptom covers two regions,  $X$  and  $Y$ , and lesions to both are required to cause the symptom. Hence, the absence of a symptom implies that at least one region is preserved and able to maintain function. In our example of a two-region system ( $X$  and  $Y$ ), this contrapositive can be expressed as  $[\sim S \rightarrow (\sim L_x \vee \sim L_y)]$ . In other words, only  $X$  or only  $Y$  is damaged:  $[(L_x \wedge \sim L_y) \rightarrow \sim S]$  or  $[(L_y \wedge \sim L_x) \rightarrow \sim S]$ . It is also possible to combine different logical operators

when describing degeneracy in distributed system, for example  $[(L_1 \vee L_2) \wedge (L_3 \vee L_4) \rightarrow S]$  indicates that symptoms will be observed if there is damage to  $L_1$  or  $L_2$  in combination with damage to  $L_3$  or  $L_4$ . Another interesting example describes paradoxical lesions<sup>57</sup> when, in the context of  $L_x$ , a symptom is not observed if  $L_y$  is also present  $[(L_x \wedge \sim L_y) \rightarrow S] \wedge [(L_x \wedge L_y) \rightarrow \sim S]$ .

Spatial dependency (neurobiological constraint 1 in the previous section), when multiple functional regions are damaged by the same lesion, will result in combinations of different behavioural symptoms i.e.  $[L \rightarrow S_1 \wedge S_2 \wedge S_3 \wedge \dots S_n]$ . The precision of this statement can be improved when the lesion can be divided into multiple parts. Ideally, we need to partition the lesion into parts that are critical or incidental to specific symptoms of interest. For example,  $[L_c \rightarrow (S_1 \wedge \sim S_2)]$  and  $[L_i \rightarrow (S_2 \wedge \sim S_1)]$ . When the critical and incidental parts are unknown, under investigation and or being refined, lesions can be divided into different parts using anatomical parcellations.<sup>58,59</sup> However, as anatomical parcellations do not correspond to either functional regions or the spatial distribution of lesions governed by pathology, it may be more informative to segregate lesions in a data-driven way.<sup>18,41,60,61</sup>

Time-dependency (neurobiological constraint 3 in the previous section) can also be factored into combinatorial rules. This is essential to the clinical application of  $[L \rightarrow S]$  rules that are only relevant for predicting outcome after brain injury when they are specified in time. For example, the statement  $[L \rightarrow (S_{T1} \wedge \sim S_{T2})]$  indicates that a lesion site will result in temporary symptoms at time 1 but not time 2 (due to recovery at time 2); the statement  $[L \rightarrow (S_{T1} \wedge S_{T2})]$  indicates that the symptoms will persist at time 2 as well as time 1, and the statement  $[L \rightarrow (S_{T2} \wedge \sim S_{T1})]$  indicates that symptoms will emerge at T2 (e.g. decline in neurodegenerative disorders).

Our example of a clinical application concerns  $[\sim L \rightarrow \sim S]$  rules in pre-surgical mapping for the treatment of epilepsy and brain tumours. Here, the question is whether cortical stimulation (a temporary induced lesion) to site  $L$  causes symptom  $S$ . If so, the surgeon may avoid resection ( $\sim L$ ) to minimize post-surgical impairment ( $\sim S$ ). Critically, however, preserving  $L$  ( $\sim L$ ) cannot be used to imply the absence of symptoms when other areas are being resected. Inferring  $\sim S$  will only be possible once the full set of possible  $L$  sites associated with  $S$  is known, and the following statement is validated  $[\sim L_1 \wedge \sim L_2 \wedge \sim L_3 \wedge \dots \sim L_n \rightarrow \sim S]$ . A second example of a clinical application for lesion-symptom rules is when the presence of a symptom is used to predict the presence of a lesion  $[S \rightarrow L]$ . For example, after the discovery that damage to Broca’s area caused speech production difficulties (Broca’s aphasia), some researchers assumed that Broca’s aphasia was the consequence of damage to Broca’s area. However, all that can be inferred for symptoms associated with distributed systems is that some part of the system has been damaged. Neuroimaging data is required to indicate which part of the system is damaged (unless damage to different parts can be inferred from unique combinations of symptoms).

**Table 1** Logic rules and recovery predictions for each of the theoretical ‘modes’ for brain–behaviour relationships

Godefroy <i>et al.</i> <sup>1</sup> and Toba <i>et al.</i> <sup>2</sup>		Logic rules, validity and recovery predictions (this study)	
Mode	Definition	Rule (logic)	Recovery prediction
The unicity mode	One single lesion (L) causes the symptom (S).	$L \rightarrow S$	Recovery potential is very limited. This mode mainly applies to lesions along primary or sensory pathways.
The equivalence mode	Lesions to either structure x ( $L_x$ ) or y ( $L_y$ ) cause symptoms (S).	$(L_x \vee L_y) \rightarrow S$	Recovery speed depends on the learning capacity of the spared (undamaged) regions.
The association mode	Both $L_x$ and $L_y$ must occur to cause symptoms (S).	$(L_x \wedge L_y) \rightarrow S$	Slow recovery when both lesions are present. No or mild deficit when only one lesion is present.
The summation mode	Symptoms (S) are observed after $L_x$ or $L_y$ (equivalence) but are worse when $L_x$ and $L_y$ co-occur (association).	$[(L_x \wedge L_y) \rightarrow S] \wedge [(\sim L_x \vee \sim L_y) \rightarrow \sim S]$	Slow recovery when both lesions are present. Faster recovery when only one lesion is present.
The mutual inhibition/masking summation mode	$L_y$ paradoxically reduces S caused by $L_x$ .	$[(L_x \wedge \sim L_y) \rightarrow S] \wedge [(L_x \wedge L_y) \rightarrow \sim S]$	Recovery from $L_x$ is faster when $L_y$ is induced by brain stimulation.

The first two columns are definitions of ‘modes’ taken from Godefroy *et al.*<sup>1</sup> and Toba *et al.*<sup>2</sup>

Finally, framing LSM in logical combinatorial rules provides a powerful and grounded way to test existing theoretical ‘modes’ describing different types of brain–behaviour relationships. Table 1 represents the brain modes previously considered.<sup>1,2</sup> We have supplemented these modes and their definitions by adding: (i) the equivalent logical rules for each mode to allow researchers and clinicians to frame their hypotheses and test the validity of combinatorial lesion rules when assessing lesion–symptom relationships; and (ii) the implications for recovery. Overall, the tools of logic can be extremely helpful to ensure valid inferences from LSM by examining the soundness of lesion rules and the associated brain–behaviour modes.

## Inferring causality, *post hoc*, along a multi-criteria continuum

A nuanced approach for considering causality is to assess it along a multi-criteria continuum. For example, the Bradford Hill framework includes nine criteria to consider when inferring causality from correlation as follows: (i) strength; (ii) consistency; (iii) specificity; (iv) temporality; (v) biological gradient; (vi) plausibility; (vii) coherence; (viii) experiment; and (ix) analogy; see below. This approach has been widely adopted in epidemiology to shift experimental inference from correlation to causation. Theoretical applications of this framework have been adapted for the assessment of causality in correlations between functional connectivity and post-stroke recovery<sup>62</sup> and between repetitive head impacts and chronic traumatic encephalopathy.<sup>63</sup> Traditional LSM does not meet all these criteria (see Siddiqi *et al.*<sup>3</sup> for a full discussion). Here, we re-frame the same set of nine Bradford Hill criteria (referring to them as BH-1 to BH-9) for assessing causality in lesion–symptom

relationships, inter-relating them with the five neurobiological constraints discussed in the first section above.

### BH-1: a high correlation between brain lesion and symptoms, across patients (strength)

An essential first step in the identification of a potentially causal relationship is to demonstrate a high correlation between brain lesion and symptoms, across large samples of patients. This helps to focus on areas with low inter-patient variability avoiding the challenge of high inter-patient variability (NC-4 and NC-5 in the first section above). Although the strength of a lesion–symptom relationship can be used as a proxy for the soundness of causal relationships, this is limited. High correlations may be spurious, particularly in small sample sizes that are vulnerable to outliers and not representative of the population<sup>64</sup> or when there is spatial dependency of damage<sup>12</sup> across functionally unrelated brain regions (NC-1 in Section I). Conversely, weak correlations may still reflect a causal relationship, for example, when patients are tested at different timepoints during their recovery or decline (NC-3 in Section I) and/or when the same symptom can be generated by different lesion sites, within the same neural circuit (NC-2 in Section I).

### BH-2: lesion load (biological gradient)

A causal relationship between lesion and symptoms is more likely when it can be demonstrated that the severity of symptoms increases with the amount of damage to a critical region—just as the degree of therapy-induced improvement is expected to be related to the amount (dose) of therapy. This has been shown in the many studies that identified: significant relationships between deficit severity and lesion load,<sup>65,66</sup> the degree of damage to critical sites,<sup>60</sup> the lesion impact score to major brain hubs<sup>67</sup> or the degree of



disconnection to key white matter tracts.<sup>58,68</sup> The greater the lesion load, the more likely it is that there is damage to one or more critical parts of the lesion (Lc in Fig. 2). Notably, however, high lesion load does not necessarily indicate more severe symptoms, particularly when a region of interest includes parts that are differentially related to the function (NC-1 in Section I). For example, if only a small part of the lesion is critical to a function (Lc), the rest being incidental co-occurring damage (Li), then high lesion load could be observed without the symptom so long as the critical part (Lc) is spared.

### BH-3: lesion site dependent symptoms (specificity)

Causality is more probable when symptoms can be shown to relate to one lesion site more than another (or one part of a lesion site more than another). This can be expressed as  $p(S|L) > p(S|\sim L)$ , i.e. the conditional probability of observing S is higher when L is present than when L is absent. A second type of specificity occurs when a lesion-symptom relationship has time-dependency (NC-3 in Section I) because symptoms may be observed in the acute phase after a neurological event but not in the chronic phase when recovery is more likely to have occurred.

### BH-4: lesion before symptoms (temporality)

If the lesion caused the symptoms, then the symptoms will occur after the lesion. Put another way, damage always precedes the onset of a symptom. This is easy to deduce after neurological events like stroke but may be less easy to infer in neurodegenerative conditions. It also applies to inferences about recovery, with the expectation that recovery will follow neuroplasticity/brain reorganization irrespective of whether improvement was spontaneous or therapy-induced. Ideally, a demonstration of temporality requires longitudinal, prospective studies (within patients). Alternatively, the time course of recovery can be estimated from assessments staggered over time across patients,<sup>69</sup> or self-reported descriptions of the time course of an individual's recovery from patients, their family or friends.

### BH-5: reproducibility (consistency)

A causal relationship between lesion site and symptoms should be observed across patients, studies, pathologies (e.g. damage caused by stroke or neurosurgery) and investigators. Demonstrating this (i.e. cross-validation) can be across heterogeneous samples and methodologies (indicating a universal relationship) or in groups of patients with specific demographics or treatments when context dependency needs to be controlled (MC-4 in Section I).

### BH-6: indirect reproducibility/consistency (analogy)

This is the demonstration that an association mimics other causal relationships that have already been demonstrated. For example, in the context of distributed neural systems (NC-2 in Section I), the effect of damage to one part of a neural circuit is expected to have a similar effect as damage to a second part of the same neural circuit.<sup>62</sup> Therefore, if a lesion to one part of a white matter tract has previously been proven to be causal, then the evidence required to infer a causal association is slightly lowered, by analogy, when symptoms are observed following damage to another part of the same white matter tract.

### BH-7: experimental manipulation

While it is possible to generate probabilistic lesion-symptom relationships from naturally occurring lesions, causal relationships can be made more compelling if supplemented by external interventions.<sup>70</sup> This can be expressed as follows:  $p(S|ex(L)) > p(S|ex(\sim L))$ , where the  $ex(L)$  means to experimentally intervene on L, (i.e. an active manipulation of L) and the expression overall means that L is a cause of S if and only if an intervention on L changes the probability that S has a particular deficit severity. For example, in the context of spatial dependency (NC-1 in Section I), experimental lesions (e.g. induced surgically in animals or virtually with neurostimulation or pharmacology in humans) might be able to distinguish which part of a lesion is causing the symptoms and which part of a lesion is spuriously associated with the symptoms.

### BH-8: biologically sensible lesion-symptom relationships (plausibility)

A causal relationship is more likely if it fits with prior observations, current scientific knowledge and established theory. For example, there is plausibility when observing visual perceptual impairments following damage to parts of the well-recognised visual processing pathways. This criterion fits well with a more Bayesian view of LSM in which prior knowledge is used to evaluate different models and explanations.

### BH-9: not biologically implausible (coherence)

A causal relationship doesn't necessarily have to replicate established knowledge but it should not profoundly conflict with what is already known from other fields and approaches.<sup>63,71</sup> An example of a lack of coherence would be if LSM associated damage to the visual cortex with auditory rather than visual perceptual impairments. This does not mean that findings that do not align with established knowledge should be ignored. It means that extra effort should be taken when checking the validity of surprising findings

particularly when LSM is based on purely data-driven approaches. ‘Coherence’ has also been defined as ‘the convergence of different approaches that reach similar results’,<sup>3</sup> which blends the ideas of coherence (BH-9) with consistency (BH-5).

## Implications for future LSM studies

Applying the nine Bradford Hill causality criteria to LSM results, *post hoc*, allows us to shift our inferences from correlation to causation.<sup>62</sup> The number of criteria that are met, along with a priority system (e.g. temporality is mandatory, whereas experimental manipulation is rarely possible), allows different degrees of causal inference to be ranked.<sup>3</sup> The goal of LSM is to maximize the likelihood that the causal parts of a lesion (Lc) are dominating the lesion-symptom relationship (Fig. 2).

For clinically reliable lesion-symptom relationships, validation needs to satisfy at least one of the following two criteria, in at least two independent datasets as follows: (i) a high positive predictive value [i.e.  $L \rightarrow S$ , the probability that a patient with lesion L has a positive (abnormal) test result (symptom S)]; and/or (ii) a high negative predictive value (i.e.  $\sim L \rightarrow \sim S$ , the probability that a patient is symptom free at a particular timepoint in recovery following a lesion that spares all the lesion sites that have previously been associated with that symptom). These requirements are not met by current LSM methods because of challenges related to (i) segregating Lc from Li (NC-1 in Section I); and (ii) defining  $\sim L$  that requires *a priori* knowledge of all possible lesion sites that have the potential to cause symptoms (see Sections II and III).

Given these constraints and limitations, it might make sense to take a more pragmatic approach by: (i) relaxing the assumption of a strong causal relationship at every identified region in LSM, and shifting interest towards the discovery of useful lesion rules; and (ii) replacing the emphasis on ‘ALL patients’ meeting a lesion-symptom rule with probabilistic estimates that aim to show ‘the expected proportion of patients’ that will meet a lesion-symptom rule. The higher the probabilistic relationship, across patients, the higher the likelihood of a strong causal lesion; and the more confident we can be in generating accurate prognoses for new patients. Furthermore, it is also essential that LSM findings are integrated with findings from other techniques to ensure convergence on common neuroanatomical circuits.<sup>72–74</sup>

Overall, by unravelling these constraints, we hope to help make sense of LSM findings that do not always converge with findings from other brain mapping techniques (cf. BH-5 and BH-9).

## Conclusion

When assessing causal relationships with LSM, observational data brings many challenges. Lesions (causes) are not

manipulated by the experimenter, and it is not possible to undo the impact of the lesion on the vascular architecture or to control brain reorganization over time. In this paper, we have (i) defined five neurobiological and five methodological constraints on LSM; (ii) examined LSM through the lenses of logic; and (iii) considered how causality can be inferred from correlations through the nine Bradford Hill causality criteria. The Bradford Hill criteria are useful for assessing the strength of causality *post hoc*. We now propose that rather than rely on *post hoc* evaluation of how well the causality criteria have been met, it might be possible to address the neurobiological and methodological constraints, *a priori*, by making the LSM approach more pragmatic. This requires changing the experimental design of LSM, a possibility that will be addressed in future work. As LSM is increasingly used on diverse clinical groups, we call for an open platform to share the identified lesion rules (e.g. as 3D maps of critical lesions in a stereotaxic standard space) so that their explanatory and predictive power can be tested in new samples. The expectations here is that sharing LSM-based lesion factors will speed up the discovery of reliable and accurate lesion rules that are clinically useful.

## Acknowledgements

The authors thank Douglas Neville, Thomas Hope, Sharon Geva and David Green for their valuable suggestions on the manuscript.

## Funding

This work was funded by the Khalifa University (grant numbers RC2-2018-022 and FSU-2022-006 to M.L.S.) and was also funded in whole, or in part, by the Wellcome Trust (grant numbers 203147/Z/16/Z, 205103/Z/16/Z and 224562/Z/21/Z to C.J.P.). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Competing interests

The authors report no competing interests.

## Data availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

## References

1. Godefroy O, Duhamel A, Leclerc X, Saint Michel T, Henon H, Leys D. Brain-behaviour relationships. Some models and related

- statistical procedures for the study of brain-damaged patients. *Brain*. 1998;121(Pt 8):1545-1556.
2. Toba MN, Godefroy O, Rushmore RJ, *et al*. Revisiting 'brain modes' in a new computational era: Approaches for the characterization of brain-behavioural associations. *Brain*. 2020;143(4):1088-1098.
  3. Siddiqi SH, Kording KP, Parvizi J, Fox MD. Causal mapping of human brain function. *Nat Rev Neurosci*. 2022;23(6):361-375.
  4. Sperber C. Rethinking causality and data complexity in brain lesion-behaviour inference and its implications for lesion-behaviour modelling. *Cortex*. 2020;126:49-62.
  5. Halpern JY. *Actual causality*. MIT Press; 2016.
  6. Eells E. *Probabilistic causality*. Cambridge University Press; 1991.
  7. Mitroff II, Silvers A. Probabilistic causality. *Technol Forecast Soc Change*. 2013;80(8):1629-1634.
  8. Sobel ME. Causal inference in randomized and non-randomized studies: The definition, identification, and estimation of causal parameters. In: Millsap RE and Maydeu-Olivares A, eds. *The sage handbook of quantitative methods in psychology*. Sage Publications Ltd; 2009:3-22.
  9. Sprenger J. Foundations of a probabilistic theory of causal strength. *Philos Rev*. 2018;127(3):371-398.
  10. Xu T, Jha A, Nachev P. The dimensionalities of lesion-deficit mapping. *Neuropsychologia*. 2018;115:134-141.
  11. Mah YH, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. *Brain*. 2014;137(Pt 9):2522-2531.
  12. Wang Y, Juliano JM, Liew SL, McKinney AM, Payabvash S. Stroke atlas of the brain: Voxel-wise density-based clustering of infarct lesions topographic distribution. *Neuroimage Clin*. 2019;24:101981.
  13. Salvadori E, Papi G, Insalata G, *et al*. Comparison between ischemic and hemorrhagic strokes in functional outcome at discharge from an intensive rehabilitation hospital. *Diagnostics (Basel)*. 2020;11(1):38.
  14. Chu CL, Chen YP, Chen CCP, *et al*. Functional recovery patterns of hemorrhagic and ischemic stroke patients under post-acute care rehabilitation program. *Neuropsychiatr Dis Treat*. 2020;16:1975-1985.
  15. Gorno-Tempini ML, Hillis AE, Weintraub S, *et al*. Classification of primary progressive aphasia and its variants. *Neurology*. 2011;76(11):1006-1014.
  16. Boes AD, Prasad S, Liu H, *et al*. Network localization of neurological symptoms from focal brain lesions. *Brain*. 2015;138(Pt 10):3061-3075.
  17. Salvalaggio A, De Filippo De Grazia M, Zorzi M, Thiebaut de Schotten M, Corbetta M. Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. *Brain*. 2020;143(7):2173-2188.
  18. Gajardo-Vidal A, Lorca-Puls DL, Crinion JT, *et al*. How distributed processing produces false negatives in voxel-based lesion-deficit analyses. *Neuropsychologia*. 2018;115:124-133.
  19. Caglayan AB, Beker MC, Sertel Evren E, *et al*. The unconventional growth factors cerebral dopamine neurotrophic factor and mesencephalic astrocyte-derived neurotrophic factor promote post-ischemic neurological recovery, perilesional brain remodeling, and lesion-remote axonal plasticity. *Transl Stroke Res*. 2023;14(2):263-277.
  20. Seghier ML, Ramsden S, Lim L, Leff AP, Price CJ. Gradual lesion expansion and brain shrinkage years after stroke. *Stroke*. 2014;45:877-879.
  21. Naeser MA, Palumbo CL, Prete MN, *et al*. Visible changes in lesion borders on CT scan after five years poststroke, and long-term recovery in aphasia. *Brain Lang*. 1998;62(1):1-28.
  22. Saur D, Lange R, Baumgaertner A, *et al*. Dynamics of language reorganization after stroke. *Brain*. 2006;129(Pt 6):1371-1384.
  23. Desowska A, Turner DL. Dynamics of brain connectivity after stroke. *Rev Neurosci*. 2019;30(6):605-623.
  24. Busby N, Hillis AE, Bunker L, *et al*. Comparing the brain-behavior relationship in acute and chronic stroke aphasia. *Brain Commun*. 2023;5(2):fcad014.
  25. Valdés Hernandez MDC, Grimsley-Moore T, Chappell FM, *et al*. Post-stroke cognition at 1 and 3 years is influenced by the location of white matter hyperintensities in patients with lacunar stroke. *Front Neurol*. 2021;12:634460.
  26. Soros P, Teasell R, Hanley DF, Spence JD. Motor recovery beginning 23 years after ischemic stroke. *J Neurophysiol*. 2017;118(2):778-781.
  27. Ciceron C, Sappey-Marinié D, Riffó P, *et al*. Case report: True motor recovery of upper limb beyond 5 years post-stroke. *Front Neurol*. 2022;13:804528.
  28. Hope TMH, Leff AP, Prejawa S, *et al*. Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. *Brain*. 2017;140(6):1718-1728.
  29. Ballester BR, Maier M, Duff A, *et al*. A critical time window for recovery extends beyond one-year post-stroke. *J Neurophysiol*. 2019;122(1):350-357.
  30. Seghier ML, Price CJ. Interpreting and utilising intersubject variability in brain function. *Trends Cogn Sci*. 2018;22(6):517-530.
  31. Price CJ, Friston KJ. Degeneracy and cognitive anatomy. *Trends Cogn Sci*. 2002;6(10):416-421.
  32. Bryce NV, Flournoy JC, Guassi Moreira JF, *et al*. Brain parcellation selection: An overlooked decision point with meaningful effects on individual differences in resting-state functional connectivity. *Neuroimage*. 2021;243:118487.
  33. Moghimi P, Dang AT, Do Q, Netoff TI, Lim KO, Atluri G. Evaluation of functional MRI-based human brain parcellation: A review. *J Neurophysiol*. 2022;128(1):197-217.
  34. Taylor KI, Stamatakis EA, Tyler LK. Crossmodal integration of object features: Voxel-based correlations in brain-damaged patients. *Brain*. 2009;132(Pt 3):671-683.
  35. Tyler LK, Marslen-Wilson W, Stamatakis EA. Dissociating neurocognitive component processes: Voxel-based correlational methodology. *Neuropsychologia*. 2005;43:771-778.
  36. Ripolles P, Marco-Pallares J, Miro J, *et al*. Analysis of automated methods for spatial normalisation of lesioned brains. *Neuroimage*. 2012;60(2):1296-1306.
  37. Pappas I, Hector H, Haws K, Curran B, Kayser AS, D'Esposito M. Improved normalization of lesioned brains via cohort-specific templates. *Hum Brain Mapp*. 2021;42(13):4187-4204.
  38. Guggisberg AG, Koch PJ, Hummel FC, Buettefisch CM. Brain networks and their relevance for stroke rehabilitation. *Clin Neurophysiol*. 2019;130(7):1098-1124.
  39. Bonilha L, Nesland T, Rorden C, Fillmore P, Ratnayake RP, Fridriksson J. Mapping remote subcortical ramifications of injury after ischemic strokes. *Behav Neurol*. 2014;2014:215380.
  40. Foulon C, Cerliani L, Kinkingnehun S, *et al*. Advanced lesion symptom mapping analyses and implementation as BCBtoolkit. *Gigascience*. 2018;7(3):1-17.
  41. Zhao Y, Halai AD, Lambon Ralph MA. Evaluating the granularity and statistical structure of lesions and behaviour in post-stroke aphasia. *Brain Commun*. 2020;2(2):fcaa062.
  42. Ramsey LE, Siegel JS, Lang CE, Strube M, Shulman GL, Corbetta M. Behavioural clusters and predictors of performance during recovery from stroke. *Nat Hum Behav*. 2017;1:0038.
  43. Sperber C, Nolingberg C, Karnath HO. Post-stroke cognitive deficits rarely come alone: Handling co-morbidity in lesion-behaviour mapping. *Hum Brain Mapp*. 2020;41(6):1387-1399.
  44. Alawieh A, Zhao J, Feng W. Factors affecting post-stroke motor recovery: Implications on neurotherapy after brain injury. *Behav Brain Res*. 2018;340:94-101.
  45. Rajashekar D, Wilms M, Hecker KG, *et al*. The impact of covariates in voxel-wise lesion-symptom mapping. *Front Neurol*. 2020;11:854.

46. Campbell KL, Tyler LK. Language-related domain-specific and domain-general systems in the human brain. *Curr Opin Behav Sci.* 2018;21:132-137.
47. MacGregor LJ, Gilbert RA, Balewski Z, *et al.* Causal contributions of the domain-general (multiple demand) and the language-selective brain networks to perceptual and semantic challenges in speech comprehension. *Neurobiol Lang (Camb).* 2022;3(4):665-698.
48. Ardila A, Concha M, Rosselli M. Angular gyrus syndrome revisited: Acalculia, finger agnosia, right-left disorientation and semantic aphasia. *Aphasiology.* 2000;14(7):743-754.
49. Lashley KS. The problem of serial order in behavior. In: Jeffries L, ed. *Cerebral mechanisms in behavior.* Wiley; 1951:112-146.
50. Grobstein P. Strategies for analyzing complex organization in the nervous system. I. Lesion experiments. In: Schwartz E, ed. *Computational neuroscience.* MIT Press; 1990:19-37.
51. Yoshihara M, Yoshihara M. 'Necessary and sufficient' in biology is not necessarily necessary—Confusions and erroneous conclusions resulting from misapplied logic in the field of biology, especially neuroscience. *J Neurogenet.* 2018;32(2):53-64.
52. Ehrenstein K. Theories of causation. In: *Causal pluralism in the life sciences. History, philosophy and theory of the life sciences.* Vol. 25. Springer; 2022. p 17-60.
53. Parascandola M. Causes, risks, and probabilities: Probabilistic concepts of causation in chronic disease epidemiology. *Prev Med.* 2011; 53(4-5):232-234.
54. Kaiserman A. Causal contribution. *Proc Aristot Soc.* 2016;116(3): 387-394.
55. Inoue K, Madhyastha T, Rudrauf D, Mehta S, Grabowski T. What affects detectability of lesion-deficit relationships in lesion studies? *Neuroimage Clin.* 2014;6:388-397.
56. Zhang Y, Kimberg DY, Coslett HB, Schwartz MF, Wang Z. Multivariate lesion-symptom mapping using support vector regression. *Hum Brain Mapp.* 2014;35(12):5861-5876.
57. Sajid N, Parr T, Gajardo-Vidal A, Price CJ, Friston KJ. Paradoxical lesions, plasticity and active inference. *Brain Commun.* 2020;2(2): fcaa164.
58. Griffis JC, Metcalf NV, Corbetta M, Shulman GL. Structural disconnections explain brain network dysfunction after stroke. *Cell Rep.* 2019;28(10):2527-2540.e9.
59. Gajardo-Vidal A, Lorca-Puls DL, Team P, *et al.* Damage to Broca's area does not contribute to long-term speech production outcome after stroke. *Brain.* 2021;144(3):817-832.
60. Geva S, Truneh T, Seghier ML, *et al.* Lesions that do or do not impair digit span: A study of 816 stroke survivors. *Brain Commun.* 2021;3(2):fcab031.
61. Kasties V, Karnath HO, Sperber C. Strategies for feature extraction from structural brain imaging in lesion-deficit modelling. *Hum Brain Mapp.* 2021;42(16):5409-5422.
62. Cassidy JM, Mark JJ, Cramer SC. Functional connectivity drives stroke recovery: Shifting the paradigm from correlation to causation. *Brain.* 2022;145(4):1211-1228.
63. Nowinski CJ, Bureau SC, Buckland ME, *et al.* Applying the Bradford Hill criteria for causation to repetitive head impacts and chronic traumatic encephalopathy. *Front Neurol.* 2022;13: 938163.
64. Weichwald S, Peters J. Causality in cognitive neuroscience: Concepts, challenges, and distributional robustness. *J Cogn Neurosci.* 2021;33(2):226-247.
65. Marchina S, Zhu LL, Norton A, Zipse L, Wan CY, Schlaug G. Impairment of speech production predicted by lesion load of the left arcuate fasciculus. *Stroke.* 2011;42(8):2251-2256.
66. Riley JD, Le V, Der-Yeghiaian L, *et al.* Anatomy of stroke injury predicts gains from therapy. *Stroke.* 2011;42(2):421-426.
67. Aben HP, Biessels GJ, Weaver NA, *et al.* Extent to which network hubs are affected by ischemic stroke predicts cognitive recovery. *Stroke.* 2019;50(10):2768-2774.
68. Hope TM, Seghier ML, Prejawa S, Leff A, Price CJ. Distinguishing the effect of lesion load from tract disconnection in the arcuate and uncinate fasciculi. *Neuroimage.* 2016;125: 1169-1173.
69. Hope TMH, Seghier ML, Leff AP, Price CJ. Predicting outcome and recovery after stroke with lesions extracted from MRI images. *Neuroimage Clin.* 2013;2(1):424-433.
70. Pearl J. *Causality: Models, reasoning, and inference.* Cambridge University Press; 2000.
71. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: How data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol.* 2015;12:14.
72. Siddiqi SH, Schaper F, Horn A, *et al.* Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nat Hum Behav.* 2021;5(12):1707-1716.
73. Molenberghs P, Gillebert CR, Peeters R, Vandenberghe R. Convergence between lesion-symptom mapping and functional magnetic resonance imaging of spatially selective attention in the intact brain. *J Neurosci.* 2008;28(13):3359-3373.
74. Jimenez-Marin A, De Bruyn N, Gooijers J, *et al.* Multimodal and multidomain lesion network mapping enhances prediction of sensorimotor behavior in stroke patients. *Sci Rep.* 2022;12(1): 22400.