Received: 28 March 2022 | Revised: 7 September 2022 | Accepted: 19 September 2022

DOI: 10.1002/hbm.26105

RESEARCH ARTICLE

WILEY

Use of multi-perturbation Shapley analysis in lesion studies of functional networks: The case of upper limb paresis

Shay Ofir-Geva[1,2] | Isaac Meilijson[3] | Silvi Frenkel-Toledo[4] | Nachum Soroker[1,2]

[1]Department of Neurological Rehabilitation, Loewenstein Rehabilitation Medical Center, Raanana, Israel

[2]Department of Rehabilitation Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[3]School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel

[4]Department of Physical Therapy, Ariel University, Ariel, Israel

Correspondence
Shay Ofir-Geva, Department of Neurological Rehabilitation, Loewenstein Rehabilitation Medical Center, 278 Ahuza, Raanana 43100, Israel.
Email: shinofir@gmail.com

Abstract

Understanding the impact of variation in lesion topography on the expression of functional impairments following stroke is important, as it may pave the way to modeling structure–function relations in statistical terms while pointing to constraints for adaptive remapping and functional recovery. *Multi-perturbation Shapley-value analysis* (MSA) is a relatively novel game-theoretical approach for multivariate lesion-symptom mapping. In this methodological paper, we provide a comprehensive explanation of MSA. We use synthetic data to assess the method's accuracy and perform parameter optimization. We then demonstrate its application using a cohort of 107 first-event subacute stroke patients, assessed for upper limb (UL) motor impairment (Fugl-Meyer Assessment scale). Under the conditions tested, MSA could correctly detect simulated ground-truth lesion-symptom relationships with a sensitivity of 75% and specificity of ∼90%. For real behavioral data, MSA disclosed a strong hemispheric effect in the relative contribution of specific regions-of-interest (ROIs): poststroke UL motor function was mostly contributed by damage to ROIs associated with movement planning (supplementary motor cortex and superior frontal gyrus) following left-hemispheric damage (LHD) and by ROIs associated with movement execution (primary motor and somatosensory cortices and the ventral brainstem) following right-hemispheric damage (RHD). Residual UL motor ability following LHD was found to depend on a wider array of brain structures compared to the residual motor ability of RHD patients. The results demonstrate that MSA can provide a unique insight into the relative importance of different hubs in neural networks, which is difficult to obtain using standard univariate methods.

KEYWORDS

hemiparesis, hemispheric asymmetry, lesion-symptom mapping, multi-perturbation Shapley value analysis, multivariate analysis, stroke

Shay Ofir-Geva and Isaac Meilijson contributed equally to this study.

1 | INTRODUCTION TO MULTI-PERTURBATION SHAPLEY-VALUE ANALYSIS AND LESION-SYMPTOM MAPPING

The large majority of previous studies that aimed to reveal the impact of lesion configuration on residual motor or cognitive functions post-stroke used univariate lesion-symptom mapping (LSM) methods. Such methods ignore the fact that processing of information relevant to motor or cognitive functions is accomplished by the orchestrated activity of large-scale networks, where the strength of connectivity between different network hubs is important and is subject to practice effects (Koch et al., 2016). For example, in the widely used voxel-based LSM (VLSM) method (Bates et al., 2003; Rorden & Karnath, 2004) the impact of damage to each voxel of the normalized brain is analyzed as a separate unit, unrelated to damage elsewhere, by comparison of behavioral scores of patients affected in that voxel to the scores of patients who are not affected there. Unlike univariate LSM, methods employing multivariate LSM may be better suited to explore network-mediated functions, due to their ability to account for inherent dependencies between network hubs for which univariate methods are blind (Mah et al., 2014; Pustina et al., 2018). A growing number of multivariate LSM methods have been applied in recent years (Forkert et al., 2015; Ivanova et al., 2021; Pustina et al., 2018; Smith et al., 2013; Yourganov et al., 2016; Zavaglia et al., 2015; Zhang et al., 2014). These LSM methods usually require relatively large cohorts for optimal performance, thus limiting their applicability (Ivanova et al., 2021; Sperber et al., 2019). Additionally, recent works which applied synthetic ground-truth simulations have shown that multivariate LSM methods may exhibit the same drawbacks of univariate LSM methods (Sperber et al., 2019) and surprisingly, may not outperform univariate LSM methods in detection of network hubs (Ivanova et al., 2021).

Most multivariate LSM methods apply machine-learning (ML) principles in an effort to delineate the importance of each brain volume (either voxel or region-of-interest [ROI]) to the observed behavioral score (Sperber et al., 2019). Alternatively, some authors have suggested a game-theoretical approach to the problem (Kaufman et al., 2009; Malherbe et al., 2018, 2021; Toba et al., 2017, 2020; Zavaglia et al., 2015, 2016). Multi-perturbation Shapley-value analysis (MSA or MPA) is a game-theoretical method that treats each discrete element in a network as a player in a coalitional game and the network performance when all elements are intact as the game's worth. The Shapley value (SV; Shapley, 1953) is defined as the unique fair division of the game worth among the players (Keinan et al., 2004). MSA was used to analyze primitive biological neural circuits of the lamprey and *Caenorhabditis elegans* (Kaufman et al., 2005; Keinan et al., 2004), genetic biochemical pathways (Kaufman et al., 2004), behavior of autonomous agents driven by artificial neural networks (Keinan et al., 2006) and reversible deactivation effects in cats (Kaufman et al., 2005; Zavaglia & Hilgetag, 2016). MSA was used also to study the impact of lesion configuration on the expression of spatial neglect, a symptom complex emerging from damage to the attention network (Kaufman et al., 2009; Malherbe et al., 2018; Toba et al., 2017, 2020) and also to study the impact of lesion configuration on a general measure of neurological impairment (Malherbe et al., 2021; Zavaglia et al., 2015). One prominent advantage of MSA over other univariate and multivariate methods was the ability to detect the importance of regions with a lower rate of lesion involvement (Kaufman et al., 2009). Moreover, due to its theoretical framework, MSA is specifically suited for studying the cerebral organization of complex functional systems that rely on large-scale neural networks.

Although MSA was applied for LSM purposes in several studies so far (Kaufman et al., 2009; Malherbe et al., 2018, 2021; Toba et al., 2017, 2020; Zavaglia et al., 2015), different authors used different approaches to critical data preparation steps and statistical inference, thus limiting reproducibility of the results. Additionally, the method was not validated using a ground-truth simulation study, as was done recently for other univariate and multivariate LSM methods (Inoue et al., 2014; Ivanova et al., 2021; Mah et al., 2014; Pustina et al., 2018; Sperber et al., 2019; Sperber & Karnath, 2017; Xu et al., 2018; Zhang et al., 2014). Ground-truth simulation is probably the best available tool for determining validity and accuracy of LSM methods, as well as for finding out their optimal parameters (see, e.g., Zhang et al., 2014; Pustina et al., 2018; Sperber & Karnath, 2018).

The aim of the current paper is to present in detail a theoretically revised approach for the application of MSA in lesion studies and to demonstrate the merits of this approach. First, we discuss methodological constraints in MSA application for LSM, together with our suggested solutions. Next, we measure the method's performance using simulated ground-truth data, assessing the accuracy of the method and choosing the best-performing parameters. We then apply MSA on a data set derived from a recent VLSM study of our group, which addressed the neural correlates of upper limb (UL) motor impairment (Frenkel-Toledo et al., 2019). This data set is used as a test case, to demonstrate how the application of MSA unravels important differences in the characteristics of the reorganizing motor network in the right and left cerebral hemispheres. Finally, we discuss the possible contribution of MSA to LSM research. The code and the data supporting this article are freely available (https://github.com/ShayOfir/MSA).

2 | METHODS

2.1 | MSA for LSM—Theoretical background

This section is arranged as follows. First, we provide some basic definitions of MSA for LSM. Next, the SV and its computation are described, including a novel formula with improved characteristics. Then, essential requirements for performance prediction are explained and a novel predictor is described, which enables use of graded-lesion data, without the need for dichotomization. Finally, we discuss the issue of reference central values, the computation of error estimates via bootstrap and the notion of "calibration." See Table 1 for a summary of key issues in research applying MSA and solutions offered.

TABLE 1 Issues in MSA research and their suggested solutions

	Issue	Previous solutions	Possible issues with previous solutions	Suggested solution	Section
1	Analyzing a network with high number of regions is impossible with full-information MSA	Estimated MSA using random permutations (Toba et al., 2020; Zavaglia et al., 2016)	Increased error and decreased reproducibility	Bound-perturbation method with potentials formula	2.1.2
2	Shapley-value computation requires binary working states while lesion-data are graded	Dichotomization of lesion data (Kaufman et al., 2009); Dichotomization of behavioral score (Malherbe et al., 2018; Toba et al., 2017)	Decreased power due to data loss	A K-NN-related predictor which can be applied on binary working states but is trained using graded lesion-data	2.1.3
3	Shapley values obtained via predicted MSA tend to regress toward the mean	No solutions suggested		Setting the reference value for Shapley values as the average and not zero; Use calibration procedure	2.1.4
4	The distribution of Shapley values is unknown, hence hypothesis testing regarding Shapley value is not feasible	Calculate confidence intervals across different tasks (Kaufman et al., 2009) Use leave-one-out (Jackknife) method (Toba et al., 2017) Use bootstrap resamples (Malherbe et al., 2021; Toba et al., 2020)	Does not allow true inferential statistics Increased bias	Apply bootstrap procedure; Beware of polymodal distribution: select the "hump" in which the overall-sample Shapley-value turn out	2.1.5

Abbreviations: K-NN, K-nearest neighbors; MSA, multi-perturbation Shapley-value analysis.

2.1.1 | Basic definitions

A brain hemisphere is divided into K regions that may be subject to clinically identifiable damage $Z = (Z_1, Z_2, \ldots Z_K)$ where $0 \le Z_i \le 1$ is the proportion of damage to region i. Let then $X_i = 1 - Z_i$ be the working state of region i. The subject performs a task and obtains a score $0 \le Y \le M$, where 0 corresponds to total failure and $M > 0$ is the score under no damage. The empirical data consist of the vectors $(X; Y) = (X_1, X_2, \ldots, X_K; Y)$ of n subjects. The question under study is the assessment of the importance of every region to the performance of the subjects in the task.

Consider a hypothetical situation in which every X_i is either 0 or 1 and for each such binary working-state pattern, there is a virtual subject with this pattern as working state and given score $v(A)$, where A is the set of regions i with $X_i = 1$. The number of virtual subjects is thus 2^K, the score $v(\varphi)$ of the vector φ of 0's is 0 and the score of the vector of 1's is M. As introduced by Keinan et al. (2004) and further analyzed and developed in Keinan et al. (2006), this formulation lends itself to a game-theoretic interpretation of a coalition game, where regions i are players, the list of working regions in a binary pattern is a coalition A and the score is the value $v(A)$ of this coalition. The players can be political parties, agents in a production chain, etc. The question under study is how to assign an importance value to each player. Specifically, how to share the total revenue M between the players in something that can be called a fair manner.

Such an allocation x_v (adding up to M) is said to be efficient in the language of game theory. The allocation is symmetric if it assigns the same share $x_v(i) = x_v(j)$ to symmetric players (i and j are symmetric if $v(S \cup \{i\}) = v(S \cup \{j\})$ for every coalition S containing neither), and it is monotone if whenever $w(S \cup \{i\}) - w(S) \ge v(S \cup \{i\}) - v(S)$ for all coalitions S not containing player i, $x_w(i) \ge x_v(i)$ (the player's share is higher in the game with payoff w).

A fundamental theorem in game theory claims that the SV is the only allocation that is efficient, symmetric and monotone (Young, 1985). This is a good case for adopting it in the context of brain damage, ignoring other allocation values such as Banzhaf (Banzhaf, 1965) or the nucleolus (Schmeidler, 1969), that satisfy different sets of axioms. There have been diverse attempts in game theory to define a SV for fuzzy coalitions (Butnariu, 1980; Tsurumi et al., 2001), corresponding quite accurately to graded damage, but these conceptual methods are beyond numerical implementation or statistical analysis.

As in many real-world problems, the performance of the 2^K multi-perturbed configurations is unknown (Keinan et al., 2004). In this case, one can train a ML predictor using the available configurations to predict the performance of the unknown configurations before applying MSA to yield the so-called predicted SV. Such method was applied successfully outside LSM context, where the elements could be either active or inactive (Kaufman et al., 2005; Keinan et al., 2004; Keinan et al., 2006). However, the damage in the context of LSM is graded, requiring some method to predict performance of binary working states from graded data. As will be described further, this is a fundamental difference between MSA for LSM and other uses of MSA, as it introduces nontrivial prediction error to the resultant SVs (Zavaglia et al., 2015).

2.1.2 | Computation of the SV: Three different formulas

For a player i and a coalition A to which this player belongs, define the contribution $\Delta_i(A)$ of this player to the coalition as the score $v(A)$ of

the coalition minus the score $v(A - \{i\})$ of the resulting reduced coalition when our player is removed from it. Under a natural symmetry axiom, these contribution values can be compactly summarized by the K averages $C_1(i)$, $C_2(i)$, …, $C_K(i)$, where $C_j(i)$ is the equal-weights average of $\Delta_i(A)$ taken over all the $\binom{K-1}{j-1} = \frac{(K-1)!}{(j-1)!(K-j)!}$ coalitions A of size j containing i. The definition via *contributions* of the SV $\gamma(i)$ is the regular (equal-weights) average of the K contributions $C_j(i)$, for $j = 1, 2, \ldots, K$.

There is an alternative and more common definition via *permutations* of the SV, in terms of all $K!$ permutations of the names of the K players. The contribution of player i in a permutation is defined as its contribution to the coalition composed of all names written on or to the left of player i. The SV $\gamma(i)$ of player i is the simple average over all permutations, of the contributions of player i to these permutations. It is clear that the number of player permutations with player i in position j is the number $(K-1)!$ of permutations of the other players, regardless of j. This explains the equivalence of the two definitions.

There is yet another equivalent definition via *potentials* (Hart & Mas-Colell, 1989) that plays a decisive role in the current report. Let $v_j(i)$ be the regular average of the values $v(A)$ of all the $\binom{K-1}{j-1}$ coalitions A of size j containing player i, and let V_j be the regular average of the values $v(A)$ of all $\binom{K}{j}$ coalitions of size j. The contributions $C_j(i)$ as defined earlier, are:

$$C_j(i) = \frac{\binom{K-1}{j-1} v_j(i) - \left(\binom{K}{j-1} V_{j-1} - \binom{K-1}{j-2} v_{j-1}(i)\right)}{\binom{K-1}{j-1}} = v_j(i) + \frac{j-1}{K-j+1} v_{j-1}(i) + \mathrm{REM}_j \quad (1)$$

where the remainder term REM_j does not depend on i (only on perturbation depth j). Recalling the definition via contributions as the regular average of these contributions, and defining the potential-like expressions

$$R(i) = \frac{v_{K-1}(i)}{1} + \frac{v_{K-2}(i)}{2} + \frac{v_{K-3}(i)}{3} + \ldots + \frac{v_1(i)}{K-1}, 1 \le i \le K \quad (2)$$

it is observed that the Shapley vector differs from the vector R by a constant. Thus, definition via potentials

$$\gamma(i) = \frac{M}{K} + R(i) - \mathrm{mean}(R), 1 \le i \le K \quad (3)$$

If contributions of players to coalitions of size smaller that $K - J$ are replaced by zero, the corresponding modified Shapley vector "of depth J" becomes

$$\gamma_J(i) = \frac{M}{K} + R_J(i) - \mathrm{mean}(R_J), 1 \le i \le K, J = 1, 2, \ldots \quad (4)$$

where

$$R_J(i) = \frac{v_{K-1}(i)}{1} + \frac{v_{K-2}(i)}{2} + \frac{v_{K-3}(i)}{3} + \ldots + \frac{v_{K-J}(i)}{J} - \frac{v_{K-J}(i)}{K} \quad (5)$$

For small K values, the three methods are equally feasible. For intermediate K values, the method of contributions requires the evaluation of the values of the 2^K coalitions (feasible number of $2^{13} = 8192$ computations for $K = 13$), while the method of permutations requires working on $K!$ permutations (an impossible $13! = 6 \times 10^9$ for the same $K = 13$). However, for cases in which 2^K is also infeasible to compute, the permutations formula becomes handy, since it enables one to compute an unbiased estimate of the exact SV: the average contribution of each player in a random sample of N of the $K!$ permutations is, loosely speaking, a consistent unbiased estimate of the exact SV (Keinan et al., 2006) (with variance insensitive to K, inversely proportional to N).

The two methods—contributions and permutation-estimation—have been applied in the literature, including in LSM studies (Kaufman et al., 2009; Malherbe et al., 2018, 2021; Toba et al., 2017, 2020; Zavaglia et al., 2015). The formula of potentials (5), a hybrid of the two, involving no randomization, is introduced here for the first time.

In the current application, the values $v(A)$ are not exogenously given as intended in Game Theory but are the result of noisy prediction. We thus favor an algorithm based on the modified SV (5) that evaluates as many summands of R as feasible and replaces all remaining summands by zero. If J summands are used (the depth J), the Shapley vector is approximated by its modified version (5). The rationale is that the $v_{K-j}(i)$ terms tend to decrease to zero as j increases but the standard deviations of coalition value estimates remain stable. Hence, beyond some depth, summands add more noise than signal: If $\widehat{\gamma}_{J-1}$ is essentially the same vector as $\widehat{\gamma}_J$ but the estimated standard errors of R_J exceed those of R_{J-1}, depth $J - 1$ should be preferred to depth J. Additionally, patients with high lesion load (i.e., with many damaged brain regions) are assumed to exhibit a substantially different neurological impairment relative to patients with low lesion load (Keinan et al., 2006), making the consideration of highly perturbed situations redundant. This notion is one of the reasons for the popularity of the correction for total lesion volume in VLSM studies (Sperber & Karnath, 2017). Another issue related to the permutations-estimation method is the low reproducibility of results. When large K is considered, repeated estimated analysis will use a substantially different set of coalitions each time (see Figure S1 in Supplementary materials). Thus, both theory and empirical data suggest that using Formula (5) would yield more accurate and reproducible results.

2.1.3 | Prediction of scores

From this description, it is seen that the graded working state and score data of the subjects must somehow provide an assessment of the score of arbitrarily determined binary working patterns (coalitions). This is a standard problem in ML: given training data consisting

of vector features and a scalar label with values in properly fixed intervals, build a predictor of the score of arbitrary feature vectors, to be evaluated at each of the (binary working pattern) coalitions chosen by the algorithm. Some studies have pre-dichotomized graded working patterns into binary working patterns by thresholding (Kaufman et al., 2009; Zavaglia et al., 2015). Others (Malherbe et al., 2018; Toba et al., 2017) have kept the graded nature of damage, but pre-dichotomized performance scores, so as to use standard ML techniques for classification. However, pre-dichotomization may reduce the amount of information (Zavaglia et al., 2016). Here, we introduce a third approach, which does not require pre-dichotomization, and is thus potentially more accurate.

Distance between working-state patterns

The current report concentrates on the original graded working-state data (normalized so that in each individual brain region, the maximal damage over all subjects in the sample is 100%) and defines the (squared) distance $d(x, y)$ between two working-state vectors $x = (x_1, x_2, ..., x_K)$ and $y = (y_1, y_2, ..., y_K)$ as

$$d(x,y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_K - y_K)^2 \quad (6)$$

whether these two are graded or binary. Of course, this particular choice can be modified, as can the prediction method described in the next paragraph. As is well known, it is generally difficult to collect clinical $(X; Y)$ data, and sample sizes are necessarily much smaller than the desirable sizes from Statistics considerations. With this in mind, we feel reticent to apply standard black-box prediction packages, in favor of simple, intuitive ideas. Furthermore, in usual ML applications, training data are representative of real data and predictors serve as interpolators. Training on graded data and applying on binary data are extrapolation, and the predictor should be addressed to this function. Future applications, with larger data banks, may whet the appetite for more complex methodology, but the interpolation–extrapolation difference should be taken into account.

The proposed predictor

As stated above, data consist primarily of the graded working-state vector $x(j)$ and score $y(j)$ of subjects $j = 1, 2, ..., n$. We add two pseudo-subjects: the binary working-state vector of all 1's and score M of the intact pseudo-subject, and that of all 0's and score 0 of the totally lesioned pseudo-subject. Let $nn = n + 2$.

Fix a number $b > 0$. For a binary working-state pattern xx, define the raw weight $w_1(j)$ of subject j as

$$w_1(j) = \exp\{-b * d(xx, x(j))\}, 1 \le j \le nn \quad (7)$$

and the normalized weight $w_2(j) = w_1(j)/\Sigma_1^{nn} w_1(i)$ adding up to 1 over the augmented sample. The working-state vectors $x(j)$ are a cloud of nn points in the K-dimensional unit cube, and these points have respective masses $w_2(j)$. As such, the center of gravity of the cloud is the point

$$CG = w_2(1) \times (1) + w_2(2) \times (2) + ... + w_2(nn) \times (nn) \quad (8)$$

Consider the line through CG and the 0-corner of the hypercube. This line describes all the working-state patterns with subject weights proportional to w_1 (and w_2). We single out the point $xxx = c\Sigma_1^{nn} w_1(j) \times (j)$ on this line for which $\sum xxx = \sum xx$, that is, with the same total working state as the pattern xx. Clearly, $c = \Sigma_i xx^{(i)} / \Sigma_i \Sigma_j w_1(j) \times^{(i)} (j)$. Finally, let $w_3 = cw_1$. *Just as* $\Sigma w_3(j) \times (j) = xxx$, we let the predictor of the performance at xx be

$$PR(xx) = w_3(1)y(1) + w_3(2)y(2) + ... + w_3(nn)y(nn) \quad (9)$$

As can be observed, the constant b is the only parameter of the prediction model (there is also the perturbation depth J). SV evaluation turned out to be quite insensitive to the choice of b, which was fixed as $b = 15$ (see Appendix among supplementary materials for more details on b-value tuning).

2.1.4 | Calibration of SVs

The K coordinates of the Shapley vector add up to M, so the average SV is fixed at $\frac{M}{K}$. More (less) important regions have SV exceeding (below) $\frac{M}{K}$. The typical brain region has positive Shapley contribution. It is tempting to claim that regions with positive SV have a beneficial effect on performance while those with negative SV hinder performance (Toba et al., 2019). This claim would be correct if the Shapley vector had been computed as designed in Game Theory, based on the true value $v(A)$ of each binary pattern coalition. Not so if these values are estimated through a ML predictor trained on the noisy data of the sample of subjects. As all averaging techniques, the Shapley vector behaves like Regression lines, where the reported slope of the relationship between dependent and independent variables (or feature and label) is mellowed by the correlation coefficient. Thus, the order of the coordinates of the Shapley vector is really what the study determines, but their magnitude is less informative or unequivocal. If the estimated Shapley vector is represented as:

$\text{Shapley} = \frac{M}{K} + (\text{Shapley} - \frac{M}{K})$, the true Shapley vector is $\frac{M}{K} + \text{FACTOR} \times (\text{Shapley} - \frac{M}{K})$, where *FACTOR* is high if the predictor is weak. The identification of such a factor is the calibration meant by the title of this section, common practice in Statistics, supported by a vast literature (Brown, 1993; Dawid, 1982; Hastie & Tibshirani, 1998). It is reasonable to assume that for each particular task or performance measure, only a few regions have decisive impact on the task, whether beneficial or inhibitory. If this is the case, then the coordinates below $\frac{M}{K}$ of the true Shapley vector should display a modal value at zero. This mode should manifest itself as a local mode in the empirical distribution of the coordinates of the estimated Shapley vector. *FACTOR* could then be determined so as to correct this modal value to be zero. This is the sense in which calibration is applied in this study. The full calibration algorithm is available as MATLAB code in

https://github.com/ShayOfir/MSA. The resultant calibrated SVs (cSVs) closely approximate ground-truth coefficients, supporting their use for inference of absolute (not only relative) importance (see Figure S2; supplementary materials). It should be noted, however, that calibration is not possible when this localization assumption does not apply (e.g., when performance is extremely noisy or highly distributed among the regions), since no clear modal value will emerge below the average.

2.1.5 | Statistical inference for SVs

Once the SV γ of each brain region has been estimated (by $\widehat{\gamma}$), some valid method should provide a standard error of this estimate, for the purpose of testing hypotheses on whether contributions are different enough from each other, and providing confidence intervals for the true, latent SVs. There are two common nonparametric methods to estimate standard errors, the leave-one-out Jackknife procedure (Efron, 1981; Quenouille, 1949, 1956) and the Bootstrap procedure (Efron, 1981; Efron & Tibshirani, 1998).

Although the jackknife leave-one-out resampling procedure was applied in previous MSA-LSM studies (Toba et al., 2017), it has several limitations. In essence, Jackknife calls for the evaluation of *n* Shapley vector estimates, leaving one subject (*i*) out at a time. These *n* vectors $\widehat{\gamma}_{LOO}^{\{-i\}}$ (leave-one-out missing *i*) are not typical as-if independent realizations of the Shapley vector estimation (as is the case to some extent under the Bootstrap method) but rather small perturbations of the all-sample Shapley vector estimate $\widehat{\gamma}$. Theory dictates that the standard error of $\widehat{\gamma}$ can be consistently estimated by the standard deviation of the leave-one-out sample *amplified* by $\sqrt{n-1}$, that is, of $\sqrt{n-1}\widehat{\gamma}_{LOO}^{\{-i\}}$. Theory is known to apply well in clean cases such as sample averages, but there is experience on some inadequacy of the square-root rule more generally (Efron & Stein, 1981). The Bootstrap resampling procedure is based on the observation that the sample of size *n* is more or less the best surrogate there is for the infinite population. Sampling *n* observations (with replacement) from this surrogate population yields an as-if sample from the population, from which the parameter in question may be estimated.

The theory behind Bootstrap has shown that for some purposes (such as estimating standard errors), the empirical distribution of these resampling parameter estimates approximates their true distribution. The current study fixes the Bootstrap sample size as 1,000 and ignores the two most extreme observations. We follow Tukey's (1980) advice of doing exploratory data analysis before resorting to automation. As will be shown in Section 3, we have observed a most remarkable phenomenon of the Bootstrap SV empirical distributions. These are bimodal for some of the brain regions, generally the most important ones, with one hump where the original-sample Shapley estimate is located, and another hump generally near the location of the mode applied in the calibration process (see SV distributions for ROIs corona radiata (CR), primary motor cortex, primary somatosensory cortex, inferior parietal lobule (IPL) and basal ganglia (BG) in the right-hemispheric damage (RHD) group, and CR, "rest-of-brain" (RoB) in the left-hemispheric damage (LHD) group, Figures S3 and S4, supplementary materials). It is as if in some Bootstrap samples the SV of the brain region expresses itself while in others the region comes out irrelevant.

Ignoring this phenomenon may produce overly conservative results. We thus incorporated an automatic mechanism that identifies the "correct" hump (i.e., the one in which the original-sample SV is located) and computes standard errors using this hump exclusively (making no difference for unimodal bootstrap distributions). This was done by fitting a mixture of two logistic distributions with the same variance and measuring theoretically the valley between the peaks, where the two densities are equal. The standard error is then estimated as the empirical standard error of the Bootstrap sample on the relevant half line to one side of the valley, amplified by a theoretical factor that depends of the Z-score of the valley in the pertinent fitted distribution. If this Z-score is less than 1.5, the regular empirical standard deviation of the entire Bootstrap sample is applied. Figure 1 is a flowchart illustrating an example of MSA application using the revised approach.

2.2 | MSA in the case of UL paresis poststroke

2.2.1 | Participants and clinical assessment

MSA was applied on a data set extracted from a cohort of 130 first-event subacute stroke patients for whom we had both normalized lesion data and a quantitative measure of motor impairment in the hemiparetic UL. These patients took part in an earlier study, where VLSM (Bates et al., 2003) was used to determine the impact of lesion topography on UL function (Frenkel-Toledo et al., 2019). Given the greater sensitivity of MSA relative to VLSM, to individual patients' data, we decided to reduce the variance in the cohort that stems from differences in the time of testing after stroke onset and recruited for the current study only patients that have been tested in the later stage of the early subacute period (between 1- and 3-months post-stroke onset). Following the trimming of subjects outside this time window, we ended with a cohort of 107 patients—49 with RHD and 58 with LHD. Time after stroke onset was 58 ± 19 days (average ± standard deviation [SD]). Patients were included in the original (Frenkel-Toledo et al., 2019) study if they did not suffer from previous neurological or psychiatric disorders and their language and cognitive status enabled comprehension of the task requirements.

All patients received standard rehabilitation therapy by therapists unaffiliated with the study who did not have access to the research data. The study was approved by the Ethics Review Board of the Loewenstein Rehabilitation Medical Center and all patients provided a signed informed consent prior to recruitment for the study. Patients' demographic and clinical data are described in Table 2. Twelve patients with RHD and two patients with LHD had spatial neglect, according to the Behavioral Inattention Test (Halligan et al., 1991), and 32 patients with LHD had aphasia, according to the Israeli Loewenstein Aphasia Test (Gil & Goral, 2004). Individual data are

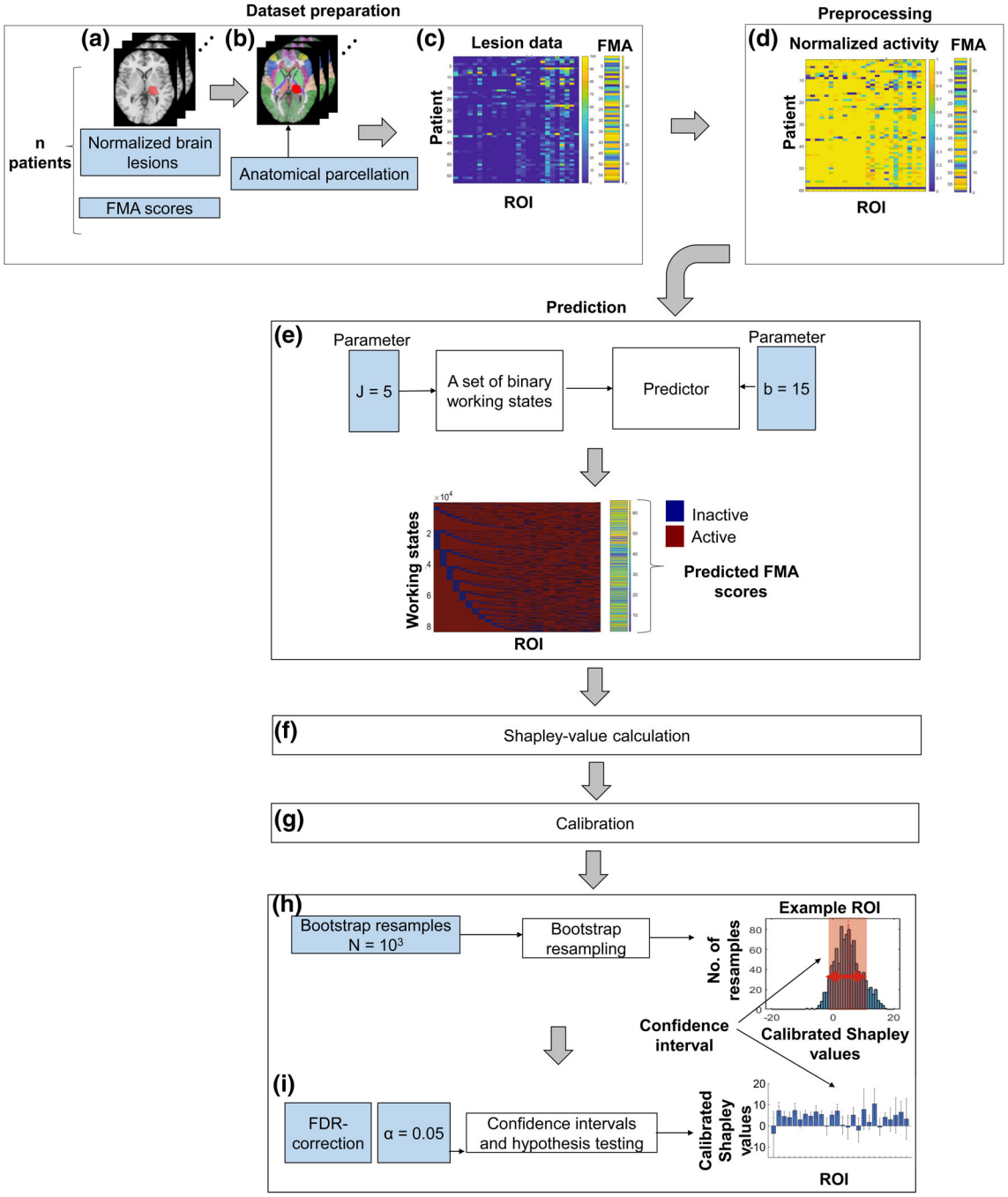

FIGURE 1 Overview of the *Multi-perturbation Shapley-value analysis* (MSA) framework. This is a schematic representation of the main stages in MSA of cerebral structure–function relationships. MSA here is based on the effect of variation in stroke lesion topography on the expression of functional impairment. The figure depicts the stages of MSA as applied in the current study of lesion impact on the expression of paralysis of the contralesional upper limb. The steps are marked with capital letters. Blue rectangles represent inputs to the algorithm and grey rectangles represent the parameters used in the current study, according to their order of application from a to i. (a–c) For each patient, structural damage was manually delineated on high-resolution CT/MRI brain scans followed by normalization into a standard (Montreal Neurological Institute [MNI]) set of templates. Personal tissue damage expressed as a percent of each affected ROI volume of our 26-ROI anatomical parcellation of the brain (see Table S2), based on the automated anatomical labeling (AAL) and white matter (WM) atlases. (d) The score obtained on the Fugl-Meyer Assessment (FMA) scale of the upper extremity represent the level of motor impairment of each patient, thus ending the data preprocessing stage by adding a performance measure to the corresponding anatomical measure in the sample. Two pseudo-subjects (minimal and maximal damage) were generated and appended. (e) A set of binary working states is selected according to perturbation depth (in the current example, only working states with five or less inactive regions). The predictor is used to predict a behavioral score (in this example, FMA score) for each working-state based on the preprocessed data set. The parameter b is an exponential coefficient ($b > 0$) that is set beforehand (in the current a value of 15 was chosen, see supplementary material). (d) Shapley values are computed according to the bound-perturbation formula (Section 2.1.2) and then (g) are calibrated (Section 2.1.4). (h) A set of bootstrap resamples (1000 in the current work) is used to obtain the empirical distribution of Shapley values and (i) for computation of confidence intervals hypothesis testing. FDR, false discovery rate; FMA, Fugl-Meyer Assessment score; ROI, region-of-interest

TABLE 2 Patients' demographic and clinical data

	RHD ($n = 49$)	LHD ($n = 58$)	p-Value
Gender (male/female)	31/18	42/16	.42[a]
Age: mean (SD)	62 (9)	60 (12)	.58[b]
Dominance (R/A/L)	45/1/3	55/0/3	.68[c]
Lesion type (I/H)	32/17	41/17	.70[a]
TAO days: mean (SD)	60 (19)	57 (19)	.41[a]
Lesion volume: mean (SD) (ml)	31.3 (39.7)	21.7 (23.8)	.76[b]
Mean FMA-UE: X/66 (SD)	40 (20)	35 (21)	.22[b]
FMA-UE total class (%) ≤25/26–45/>45	42.9/18.4/38.8	29.3/20.7/50.0	.33[a]

Abbreviations: FMA-UE, Fugl-Meyer Assessment for the upper extremity; R/A/L = right/ambidextrous/left; I/H, ischemic/hemorrhagic (ischemic lesions with hemorrhagic transformation were considered as ischemic); RHD, LHD, right and left hemisphere damage; TAO, time after stroke onset.
[a]Chi-square test.
[b]Mann–Whitney test.
[c]Fisher's exact test.

presented in Table S1 (among the supplementary materials). The evaluation of motor impairment in the hemiparetic UL is based on the standardized Fugl-Meyer Assessment for the upper extremity (FMA-UE; Gladstone et al., 2002). The scale, which emphasizes the capacity to perform isolated proximal and distal UL movements, has proven to be sensitive, reliable and valid for the assessment of UL motor impairment after stroke (Duncan et al., 1983; Platz et al., 2005).

2.2.2 | Imaging and lesion analysis

Follow-up CT ($n = 102$) and T1 weighted MRI ($n = 5$) brain scans (median, interquartile range [IQR]: 42, 32 days from stroke onset) were carefully examined by a physician experienced in the analysis of neuroimaging data (author NS) to ensure that lesion boundaries were clear and traceable and that the scan presents a stable pattern of tissue damage without a mass effect from residual edema. Lesion analyses were performed with the analysis of brain lesions (ABLe) module (Solomon et al., 2007) implemented in MEDx software (Medical-Numerics, Sterling, VA). Lesion delineation was made manually on the digitized CTs. ABLe characterizes brain lesions in MRI and CT scans of the adult human brain by spatially normalizing the lesioned brain into Talairach space using the Montreal Neurological Institute (MNI) template, with 2 mm voxel resolution. Tissue damage was reported according to the normalized automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) and white matter (WM) atlas (Mori et al., 2008). Quantification of the amount of tissue damage within each ROI of the atlas was obtained as described earlier (Haramati et al., 2008). Registration accuracy (Solomon et al., 2007) of the scans to the MNI template across all subjects ranged from 89.1 to 95.9% (LHD: 94.1 ± 1.3%; RHD 94.1 ± 1.2%). Overlay lesion maps (stroke lesion distribution) of LHD and RHD patients are shown in Figure 2. Individual patients' lesions are displayed on a set of standard templates in Figure S5 (supplementary materials).

The extent of tissue damage within 90 original ROIs of the AAL and WM atlases was determined for the 107 participants. This number of regions could not be handled reliably by MSA given the cohort size at hand (49 RHD; 58 LHD). Thus, we have created a new parcellation (Table S2; supplementary materials) of the entire brain by grouping adjacent regions, to obtain a set of 26 ROIs and computed the region-wise extent of damage (0–100%) for each ROI. ROIs known to be relevant to motor control of the UL, according to past studies in human and nonhuman primates (Buneo & Andersen, 2006; Byblow et al., 2015; Frenkel-Toledo et al., 2019; Groenewegen, 2003; Hoffman & Strick, 1995; Nachev et al., 2008; Rodríguez-Herreros et al., 2015; Sommer, 2003; Vingerhoets, 2014; Wenzelburger et al., 2005) were included in the new parcellation at a higher resolution than regions without such known relevance. A comparison of lesion patterns between the LHD and RHD patient groups was done in a ROI-wise manner, by calculating for each ROI the proportion of patients in whom damage to the ROI was noted (any damage, yes/no). Since lesions involved usually only a small subset of ROIs in each patient, we calculated the ROI-wise median extent of damage for lesioned patients. Differences in lesion-pattern (LHD cohort vs. RHD cohort) were assessed using chi-square test (for proportions of patients with any damage) and Mann–Whitney test (for median damage). Correction for multiple comparisons was done here using permutations methods (100,000 permutations), implemented by "coin" package in R language (Hothorn et al., 2008). All data analysis was done using custom scripting in MATLAB R2020a (The MathWorks, Natick, MA) and RStudio (RStudio Team, 2015).

2.2.3 | Application of MSA in this case

As explained earlier, while MSA requires one to know the performance of dichotomous brain working states (where each region is either intact or inactive), patient lesion-loads are graded, theoretically allowing for a graded activity. Thus, we developed the ML predictor (related to K-nearest neighbors method) described above to predict the behavioral scores of the required dichotomous working states, based upon the graded data set at hand. As explained above, due to large differences in lesion-load across ROIs, scaling was required to avoid prediction bias in favor of ROIs with high lesion loads. Hence,

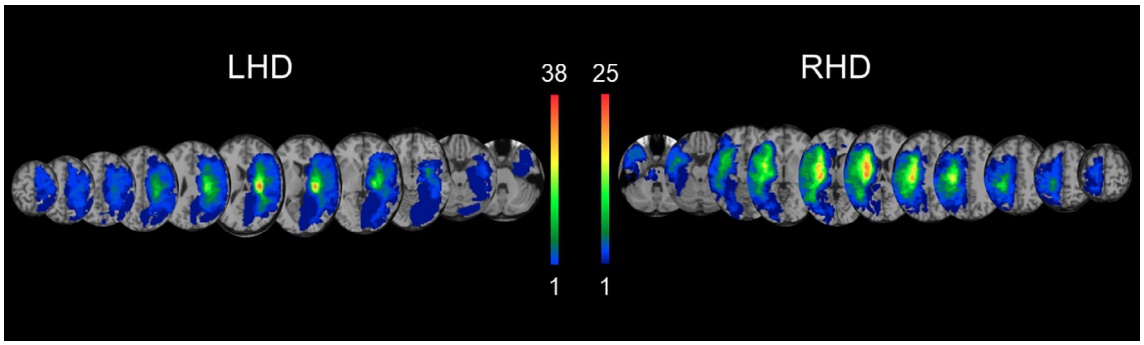

FIGURE 2 Lesion overlay maps of LHD and RHD patients ($n = 58$, $n = 49$; respectively). Representative normalized slices (out of 90 normalized slices employed) are displayed in radiological convention (right hemisphere on left side and vice versa), with warmer colors indicating greater lesion overlap (units: number of patients with lesion in this region). LHD, left-hemispheric damage; RHD, right-hemispheric damage

each lesion load was divided by the maximal lesion load seen for that region.

Following previous MSA reports (Kaufman et al., 2009; Malherbe et al., 2018; Toba et al., 2017, 2020; Zavaglia et al., 2015), we assessed the predictor's accuracy using leave-one-out cross-validation. The predictor could explain 7.2 and 13.1% of the observed variance in the LHD and RHD cohorts, respectively. The prediction mean squared error (MSE) was significantly smaller than chance-level prediction, computed by applying the predictor to 100,000 random permutations of the data (LHD: MSE = 415.1; $p = .03$; RHD: MSE = 417.9; $p = .011$).

Since the number of analyzed brain regions ($K = 26$) was too large for full-information MSA computation (as an impossible number of 2^{26} predictions are required) we used bound-perturbation MSA (Keinan et al., 2006), which involves restricting the analysis to those coalitions in which no more than J regions are inactive (i.e., bounding perturbation level to some J, as explained before). Although bound-perturbation gets closer to the full-information Shapley value as J increases (Keinan et al., 2006), preliminary analysis showed that increasing J also increases prediction error (as expressed by bootstrap-derived standard error; see Figures S6 and S7; supplementary materials). Moreover, computation time rises exponentially as J increases: For example, when considering 26 regions, $J = 4$ requires the prediction of behavioral score of 17,902 coalitions while $J = 5$ requires the prediction of 83,682 coalitions. Hence choosing the optimal perturbation depth is crucial for successful application of MSA. We thus assessed the performance of all perturbation depths from 1 to 6 using ground-truth simulation (see the next section) and used the best-performing depth in the subsequent analysis of real-world data.

Statistical inference was done using Bootstrap analysis (Efron, 1981; Toba et al., 2020), a commonly used nonparametric resampling technique, with a sample size of 1000. Subsequently, a region-wise t test was computed using the derived standard errors, to determine whether the Shapley value of a given region was significantly greater than the average Shapley value. The family-wise error rate was corrected using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995) with $\alpha = .05$.

2.3 | Ground-truth simulation

Ground-truth simulation was performed to assess the accuracy of bound-perturbation MSA using different depths ($J = 1$–6). In order to make the simulation more intelligible for clinical inference, it was based on the FMA-UE data of stroke patients from our cohort. As the critical brain regions for a given human behavior are known largely from LSM studies, simulation seems the most feasible way to validate LSM methods (Ivanova et al., 2021; Xu et al., 2018). However, different authors used different ground-truth models (Ivanova et al., 2021; Mah et al., 2014; Pustina et al., 2018; Y. Zhang et al., 2014). One of the postulated advantages of multivariate LSM methods like MSA is their ability to treat more effectively brain networks (Ivanova et al., 2021; Mah et al., 2014; Pustina et al., 2018; Zhang et al., 2014). We thus chose a relatively simplistic lesion-behavior model, as previously done by Zhang et al. (2014), in which activity of a brain region (*activity* = 1−*damage*) is linearly related to its lesion load and the total behavior score is the sum of its network members' activity:

$$y = R_1 + R_2 + R_3$$

where y is the simulated behavior and R_1, R_2, and R_3 are the activity levels of three brain regions, selected from the 26 ROIs in the parcellation (see Section 2.2.2). To avoid spurious results, we used 120 random combinations of the three regions (without replacement) out of 15,600 possible options. Each of these 120 different ground-truth brain-models was applied on the lesion-data of the two clinical subcohorts (LHD, RHD) to produce a total 240 data sets, each consisting of a lesion pattern and a simulated behavioral score. Each data set was submitted to bound-perturbation MSA with six depths ($J = 1$–6). Standard deviations for each Shapley value were calculated using 1000 bootstrap resamples of the patients' lesion data, as described before and used in previous studies (Malherbe et al., 2021; Toba et al., 2020). Each region-wise SV was then transformed to the corresponding Z-score as follows:

$$Z_i = \frac{\mathrm{SV}_i - \mathrm{mean}(\mathrm{SV})}{\mathrm{SD}}$$

2.3.1 | Area under receiver operator characteristic curves

The area under the curve (AUC; Metz, 1978; Zhang et al., 2014) quantifies overall performance. It was computed for each perturbation depth, for each MSA run. Each Z-score vector was thresholded 26 times, using each individual Z-score as a threshold. True positive and false positive rates were calculated for each threshold to yield the receiver operator characteristic (ROC) curve. ROC curve computation was followed by evaluation of AUC for each such curve. AUC of 1 represents a perfect test while AUC of 0.5 represents a test with chance-level performance. Since the AUC did not conform to a Gaussian distribution, we applied wild-bootstrap repeated-measures two-way ANOVA using "RM.MANOVA" R package (Friedrich et al., 2019) for analysis of the effect of perturbation depth ($J = 1$–6) and subcohort (LHD, RHD) on each AUC (two within-subjects factors, ground-truth model served as indexing factor). When significant effects were detected, post hoc analysis between different depths and different cohorts was done using paired bootstrap method with bias-correction and acceleration with "wBoot" R Package (Weiss, 2016) to test for any differences between the perturbation depth maximal value and the rest. Both analyses were based on 100,000 bootstraps resamples, and the FDR method (Benjamini & Hochberg, 1995) was used to correct for multiple comparisons.

2.3.2 | Sensitivity and specificity

Average sensitivity and specificity measures were computed for the best-performing perturbation depth. A region-wise t test was computed using bootstrap-derived standard errors ($\alpha = .05$) to determine which regions were significantly higher than the average SV. Subsequently, the obtained regions were compared to the brain-model critical regions to obtain sensitivity and specificity measures. This procedure was repeated for each brain model and cohort, with and without FDR correction for multiple comparisons. Confidence intervals for sensitivity and specificity were also computed using wild bootstrap method, with 100,000 resamples, implemented with "RM.MANOVA" R package (Friedrich et al., 2019).

Following the ground-truth simulation, bound-perturbation MSA with best-performing depth was applied to the two cohorts of sub-acute stroke patients (RHD, LHD), using patients' FMA-UE scores, to unravel the relative contribution of each brain ROI to overall functioning of the hemiparetic UL. As with simulations, brain regions with SVs significantly higher than the average (FDR-corrected p-value <.05, 26 comparisons) were considered as highly contributing. Confidence intervals were constructed using 1000 bootstrap resamples.

2.3.3 | Localization index

One useful feature of MSA is the ability to quantify the extent of function localization. This measure, called "localization Index" L, was described by Aharonov et al. (2003), and it is calculated as follows:

$$L = \frac{\mathrm{std}(c)}{\sqrt{(N-1)/N^2}}$$

where c is the raw SV (normalized so as to add up to 1), N is the number of regions, and L is in the range of [0, 1] (where $L = 0$ denotes a fully distributed function and $L = 1$ denotes single-region localization). The localization index was computed for each Shapley vector.

3 | RESULTS

3.1 | Comparative analysis of demographic, clinical, and anatomical characteristics of LHD and RHD patient groups

The demographic and clinical characteristics of RHD and LHD patient groups were essentially similar (Table 2), with no significant differences between the groups in either sex, dominance, time after stroke onset, lesion volume or FMA-UE scores. Overlay lesion maps (stroke lesion distribution) of LHD and RHD patients are shown in Figure 2. The proportion of patients having a lesion (yes/no) was similar for all ROIs except for the posterior limb of the internal capsule (PLIC) and the CR, which were more frequently damaged in the LHD group (chi-square independence-test, LHD vs. RHD; PLIC: 91.4 vs. 67.3%, $\chi^2 = 8.26$, permutations-corrected p-value = .003; CR: 94.8 vs. 81.6%, $\chi^2 = 3.41$, permutations-corrected p-value = .037). The median region-wise extent of damage was similar in LHD and RHD patients, except for the medial temporal lobe (MTL) where RHD patients had a more extensive damage than LHD patients (median damage for lesioned patients IQR, LHD vs. RHD: 0.9% [0.3–2.1%] vs. 6.7% [1.9–11.0%], $U = 81$, permutations-corrected p-value = .007). Descriptive statistics of damage within each ROI in the two cohorts are shown in Figure 3.

Since high co-linearities between ROIs (Xu et al., 2018) might affect the interpretation of results obtained from MSA, we checked the extent of lesional dependence for each pair of ROIs, by computing the pairwise Pearson's correlation coefficient of normalized activity (1—damage) for all ROIs. After FDR correction, 77% of the pairs in the LHD and 63% of the pairs in RHD did not show a significant correlation. Although MSA can overcome partial dependencies, resolving the independence of two perfectly correlated regions is impossible (Toba et al., 2019). Figure 4 describes ROI pairs with *Pearson's* $R \geq .8$ (perfectly or almost-perfectly correlated).

3.2 | Ground-truth simulation results

Figure S2 (supplementary materials) depicts one representative example of simulation experiment results. In this specific example, the three critical brain regions were selected to be the IPL, BG, and anterior limb of internal capsule. Performance was equal to the sum of activities (1—damage) in these three regions. MSA for LSM (with $J = 5$) was applied on the LHD and RHD lesion data with their corresponding

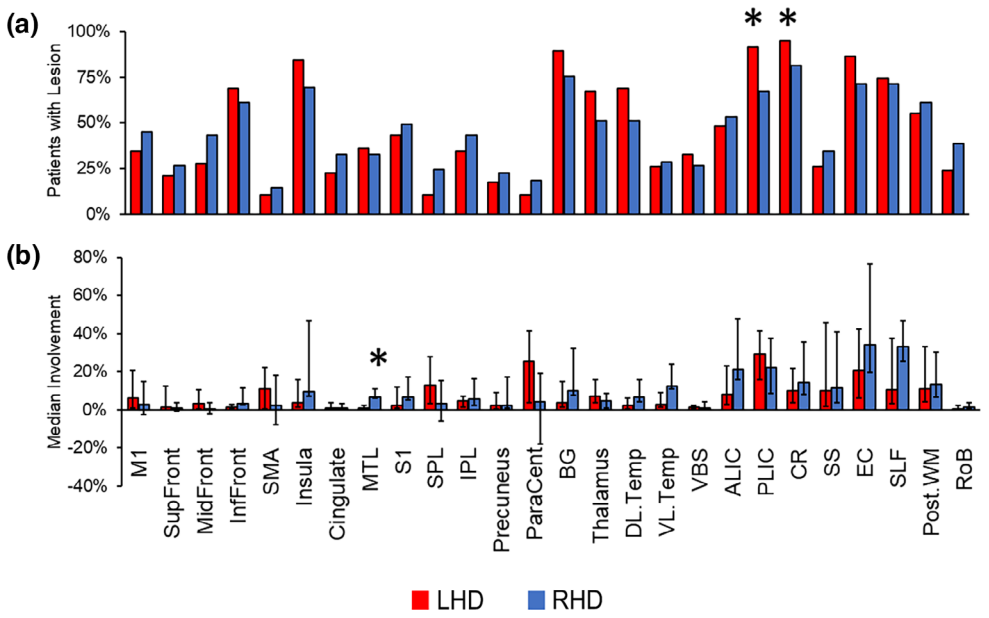

FIGURE 3 Descriptive statistics of damage within each ROI. (a) Proportion of patients (%) with any damage (yes/no) in each region. Asterisks denote significant difference between LHD and RHD (chi-square test, $p < .05$, corrected for multiple comparisons with permutations). (b) Median (IQR) ROI involvement by lesion (% of ROI's volume) for patients with any damage to the ROI. Asterisks denote significant difference between LHD and RHD (Mann–Whitney test, $p < .05$, corrected for multiple comparisons with permutations method) (see Table S2 for ROIs' correspondence to AAL and WM atlas' regions). ALIC, PLIC, anterior and posterior limbs of internal capsule; BG, basal ganglia; CR, corona radiata; DL-Temp, VL-Temp, dorsolateral and ventrolateral parts of the temporal lobe; EC, external capsule; LHD, RHD, left and right hemispheric damage; M1, primary motor cortex; MTL, medial temporal lobe; ParaCent, paracentral lobule; Post-WM, posterior white matter (superior and inferior fronto-occipital fasciculi, uncinate fasciculus); RoB, rest of the brain (other brain regions analyzed together); ROI, region of interest; S1, primary somatosensory cortex; SLF, superior longitudinal fasciculus; SMA, supplementary motor area; SPL, IPL, superior and inferior parietal lobules; SupFront, MidFront, InfFront, superior, middle and inferior frontal regions; SS, sagittal stratum; VBS, ventral brainstem

simulated behavioral scores. Both raw and cSVs are depicted. When the raw SVs are compared to the ground-truth SVs (by definition, 33% for each critical region and 0% for all of others), the raw SVs were biased upward, and their amplitude was diminished. Applying a calibration procedure (as explained in Section 2) obtained values approximating the ground-truth values. Regions which their SVs were significantly higher than the average SV (after FDR correction for multiple comparisons) were marked by asterisks. As could be seen in this example, all of the critical regions were correctly detected, in addition to some false positives (posterior white matter in LHD and RHD cohorts and external capsule in the RHD cohort). This demonstrates that MSA results are affected to some extent by lesion dimensionality, as the same ground-truth model was used for both cohorts.

Assessment of MSA performance in LSM was done using the AUC of the ROC curve method (AUROC). Figure 5 depicts the AUROC for each MSA approach, together with 95%-confidence intervals. All perturbation depths were significantly better than chance-level (AUC = 0.5) in detecting the underlying simulated ground-truth. However, performance differed between depths (main effect for depth, $F(5) = 149.32$, $p < .001$). Moreover, performance was also influenced by the cohort (LHD vs. RHD) (main effect for cohort, $F(1) = 29.12$, $p < .001$). Post hoc analysis revealed that for the LHD cohort, perturbation depth of $J = 5$ was significantly better than depths $J \leq 4$ (bootstrap paired-sample test, FDR-$p < .005$) and also than depth of $J = 6$, but with borderline significance (bootstrap paired-sample test, FDR-$p = .063$). For the RHD cohort, both depth $J = 4$ and $J = 5$ were found to be equivalent (AUC, mean, [95%-CI]; $J = 4$: 0.945 [0.927, 0.964]; $J = 5$: 0.948 [0.928, 0.967]; bootstrap paired-sample test, FDR-$p = .370$) and significantly better than bound-perturbation MSA with $J < 4$ and $J = 6$ (FDR-$p < .001$). Overall, since bound-perturbation MSA with $J = 5$ was the best performing approach for both RHD and LHD cohorts, we used it for analysis of the real-world behavioral data.

Since the AUROC is independent of the specific signal detection properties of a specific threshold, we also calculated the sensitivity and specificity for detection of critical ground-truth regions with $J = 5$. When using alpha of .05 without correction for multiple comparisons, as was done in previous MSA studies (Kaufman et al., 2009; Malherbe et al., 2018; Toba et al., 2017, 2020; Zavaglia et al., 2015), the false-positive rate [95%-CI] was 15.7% [14.4%, 17%] for the RHD cohort and 19.5% [18%, 21.1%] for the LHD cohort, while the false-negative rates were 13.9% [10.4%, 17.3%] for the RHD cohort and 16.9% [13.7%, 20.2%] for the LHD cohort. Applying FDR correction was associated with significant higher specificity (10.3% [9%, 11.5%] and 14.7% [13%, 16.3%] for the RHD and LHD cohorts, respectively; bootstrap paired-sample test, $p < .001$) and lower sensitivity (25% [20.8%, 29.2%] and 26.7% [22.4%, 30.9%] for the RHD and LHD cohorts, respectively; bootstrap paired-sample test, $p < .001$). Since FDR correction provided better specificity with reasonable sensitivity, we used it for analysis of the real-world behavioral data.

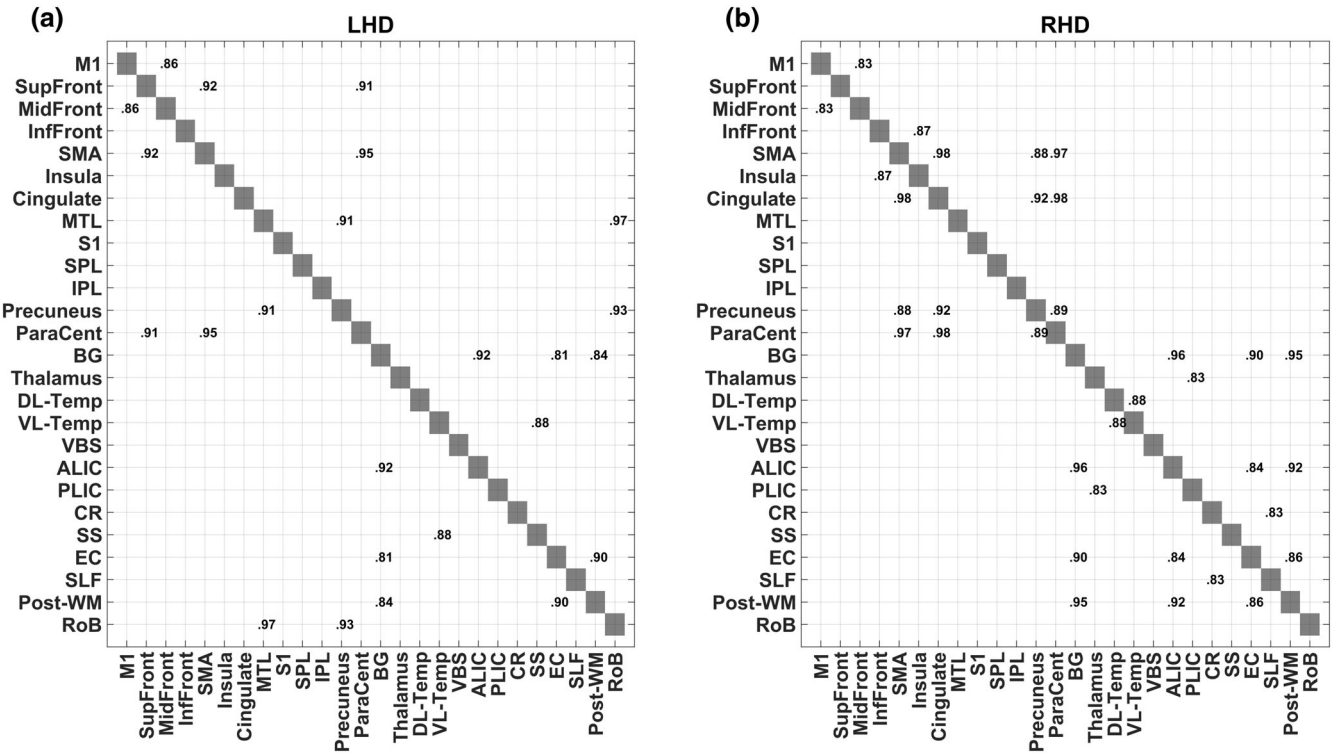

FIGURE 4 Pair-wise ROI dependencies. Pair-wise Pearson's correlation coefficients were computed between each pair of ROIs for the left-hemispheric damage and right-hemispheric damage cohorts. Correlation coefficients of 0.8 or higher were considered perfectly or almost-perfectly dependent (shown as two-decimal fraction) (see Table S2 for ROIs' correspondence to AAL and WM atlas' regions). ALIC, PLIC, anterior and posterior limbs of internal capsule; BG, basal ganglia; CR, corona radiata; DL-Temp, VL-Temp, dorsolateral and ventrolateral parts of the temporal lobe; EC, external capsule; LHD, RHD, left, right hemispheric damage; M1, primary motor cortex; MTL, medial temporal lobe; ParaCent, paracentral lobule; Post-WM, posterior white matter (superior and inferior fronto-occipital fasciculi, uncinate fasciculus); RoB, rest of the brain (other brain regions analyzed together); S1, primary somatosensory cortex; SLF, superior longitudinal fasciculus; SMA, supplementary motor area; SPL, IPL, superior and inferior parietal lobules; SupFront, MidFront, InfFront, superior, middle and inferior frontal regions; SS, sagittal stratum; VBS, ventral brainstem

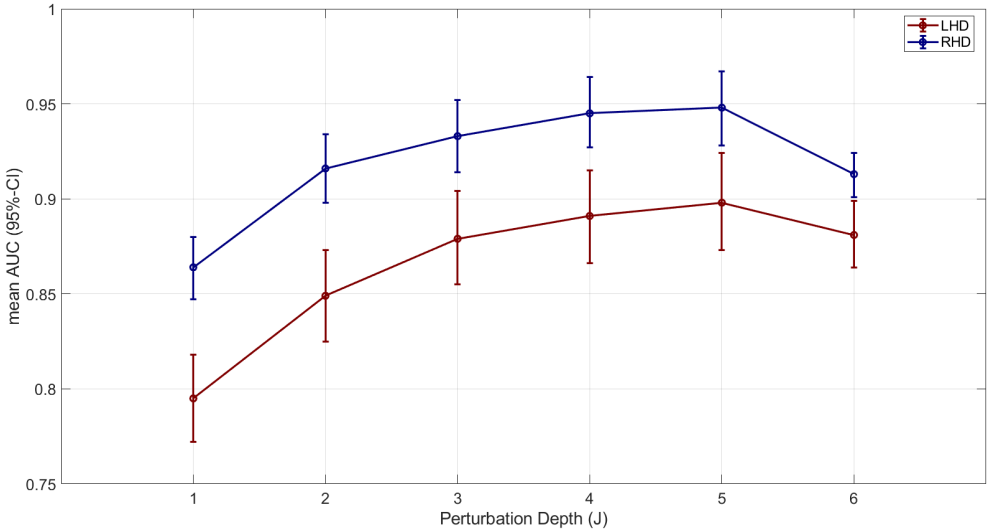

FIGURE 5 Mean area under the receiver-operator characteristic (ROC) curve. Synthetic behavioral scores were simulated 120 times for each cohort, each time assuming a different brain model (see Section 2 for details). MSA was applied on the obtained data sets using different perturbation depths (J). The area under the ROC curve was calculated by thresholding the resultant Shapley Z-scores. 95%-confidence intervals were computed using wild bootstrap method with 100,000 resamples. Each mark represents the mean and area under the ROC curve of 120 different brain models. The same brain models were used for all analyses. CI = confidence interval; LHD and RHD, left- and right-hemispheric damage; ROC, receiver-operator characteristic

3.3 | Relative contribution of different network hubs to UL motor control in the recovering brain

The relative contribution of each brain ROI to residual UL motor function, as expressed in the FMA-UE score obtained in the subacute phase of the disease, was evaluated separately for the LHD and RHD patient groups, by calculating the SVs for all possible region-coalitions with 5 or fewer inactive regions (perturbation depth $J = 5$). Standard errors were obtained from the distribution of SVs, estimated using 1000 bootstrap resamples.

Figure 6 depicts the SVs after calibration together with their corresponding bootstrap-derived 95% confidence interval. In the LHD group, three ROIs had a cSV which passed the FDR-correction for multiple comparisons: the superior frontal gyrus (SFG), the supplementary motor area (SMA), and the paracentral lobule (PCL) (cSV, 95%-CI; SFG, 7.2 [4.7, 9.7], $p_{FDR} = .041$; SMA, 7.3 [5.1, 9.4], $p_{FDR} = .016$; PCL, 7.0 [5.0, 9.0], $p_{FDR} = .016$). In the RHD cohort, four ROIs had a cSV which passed the FDR-correction for multiple comparisons—the ventral brain stem (VBS), the primary somatosensory cortex (S1), the primary motor cortex (M1), and the BG ((cSV, 95%-CI; VBS, 28.9 [12.4, 45.3], $p_{FDR} = .016$; S1, 22.2 [16.0, 28.5], $p_{FDR} < .001$; M1, 21.0 [10.4, 31.6], $p_{FDR} = .014$), BG, 13.1 [6.4, 19.9], $p_{FDR} = .026$).

These findings point to marked differences between the left and right hemispheres in the impact of lesion topography on the level of impairment in the hemiparetic UL (see Figure 7 for summary of main results). In addition, the interregional variability was consistently lower for LHD compared to RHD, pointing to a more homogenous role of most network hubs in the dominant left hemisphere (only few patients in the cohort were non-right-handers, and repeating the analysis without their data yielded essentially the same results [see Figure S8 among the supplementary materials]). The difference between the hemispheres in this respect was quantified using a "localization index" *L* (Aharonov et al., 2003), as done previously (Kaufman

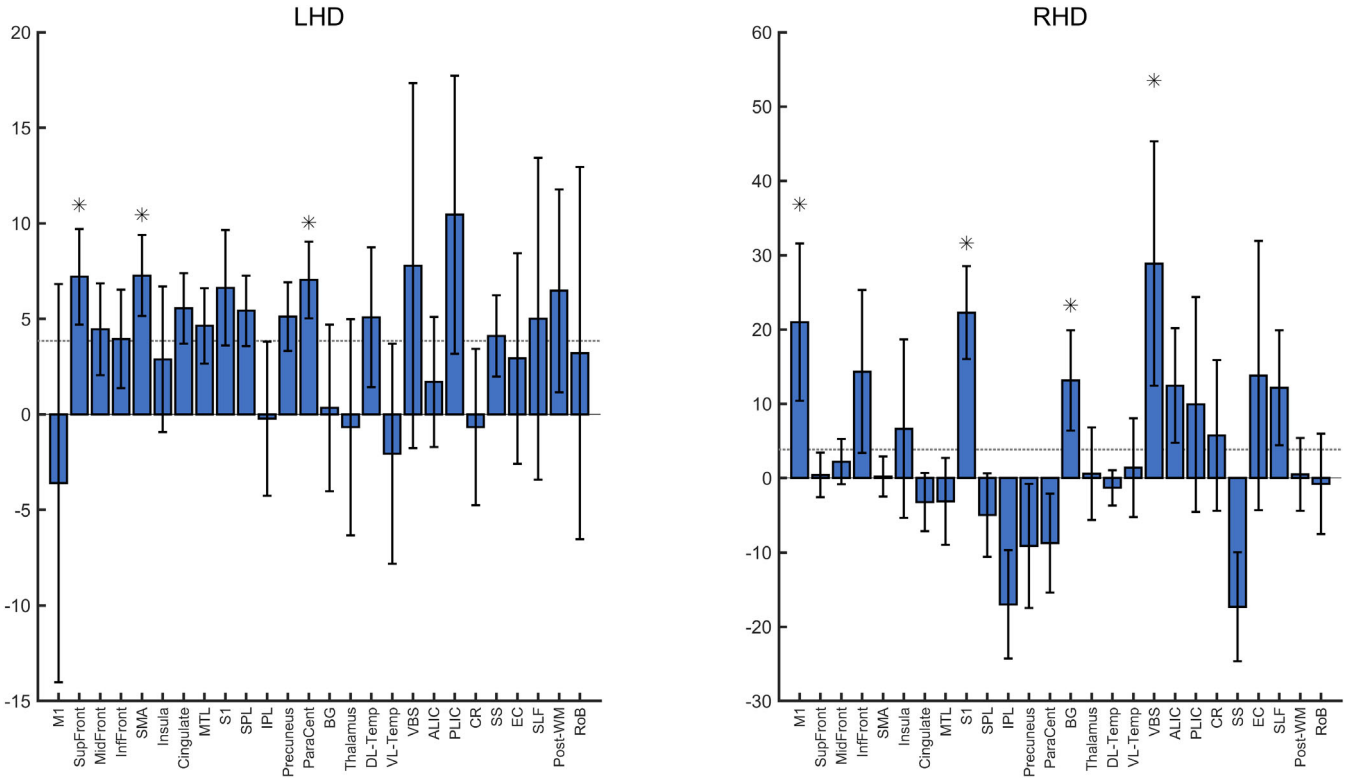

FIGURE 6 Calibrated Shapley values of brain ROIs for FMA-UE scores at the subacute phase poststroke. Calibrated Shapely values for UL motor control in the LHD and RHD cohorts, calculated using MSA for 26 brain ROIs, at the late subacute phase poststroke onset, with FMA-UE scores normalized to 100 points indicating the level of motor functioning in the hemiparetic upper limb. Whiskers denote 95%-confidence interval (derived through 1000 bootstrap resamples). The average Shapley value is marked with a dashed line (100/26 ≈ 3.85 in all figures). Asterisks denote regions with Shapley value significantly higher than this baseline (false discovery rate [FDR]-corrected $p < .05$) (see Table S2 for ROIs' correspondence to AAL and WM atlas' regions). ALIC, PLIC, anterior and posterior limbs of internal capsule; BG, basal ganglia, DL-temp, VL-temp, dorsolateral and ventrolateral parts of the temporal lobe; CR, corona radiata; EC, external capsule; FMA-UE, Fugl-Meyer Assessment Scale for the upper extremity; LHD, RHD, left, right-hemispheric damage; M1, primary motor cortex; MTL, medial temporal lobe; ParaCent, paracentral lobule; Post-WM, posterior white matter (superior and inferior fronto-occipital fasciculi, uncinate fasciculus); RoB, rest of the brain (other brain regions, analyzed together); ROI, region of interest; S1, primary somatosensory cortex; SLF, superior longitudinal fasciculus; SMA, supplementary motor area; SPL, IPL, superior and inferior parietal lobules; SupFront, MidFront, InfFront, superior, middle and inferior frontal regions; SS, sagittal stratum; VBS, ventral brainstem

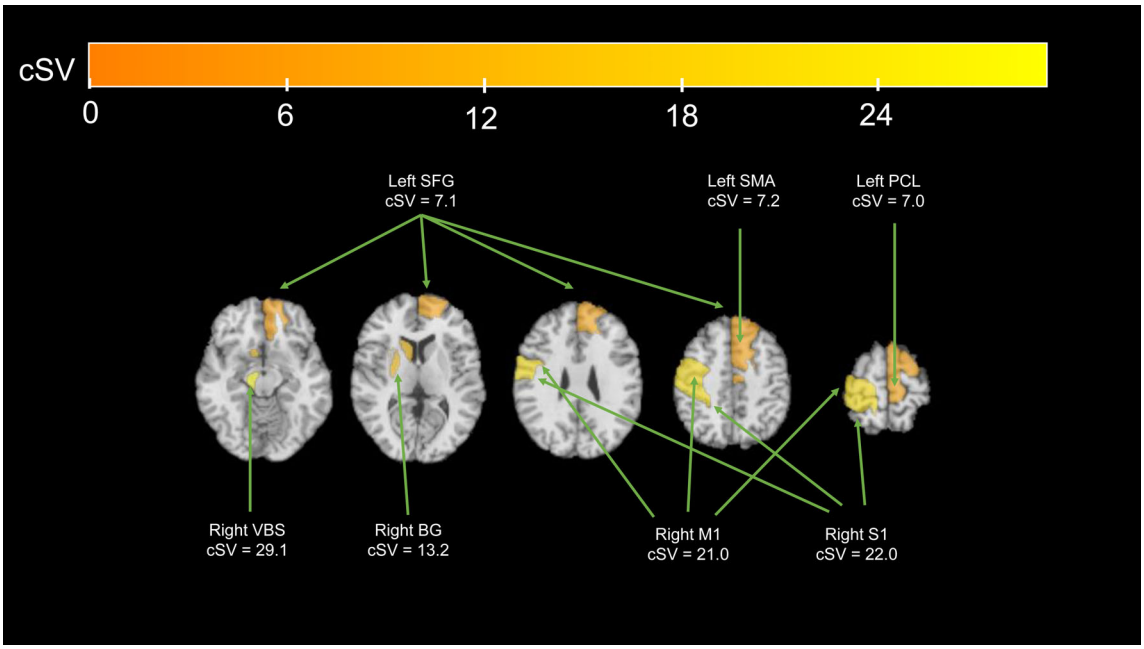

FIGURE 7 Brain regions-of-interest (ROIs) with significant calibrated Shapley values for FMA-UE scores at the subacute phase poststroke. Brain ROIs with calibrated Shapley values for FMA-UE score (normalized to 100 points) that were significantly larger than the expected baseline (false discovery rate [FDR]-corrected $p < .05$). The significant ROIs from two cohorts (LHD and RHD) are overlaid on a single MRI template (MNI single-subject, Tzourio-Mazoyer et al., 2002). Brighter colors denote larger calibrated Shapley values. BG, basal ganglia; cSV, calibrated Shapley value; FMA-UE, Fugl-Meyer Assessment for upper extremity; M1, primary motor cortex; PCL, paracentral lobule; S1, primary somatosensory cortex; SFG, superior frontal gyrus; SMA, supplementary motor area; VBS, ventral brainstem

et al., 2009). The localization index was significantly higher for the RHD cohort (L_{RHD} vs. L_{LHD}, 0.13 ± 0.01 vs. 0.09 ± 0.01; bootstrap-based Welch t test, 1000 resamples, $t(1996.6) = 59.14$, $p < .001$), suggesting that motor control in the left hemisphere is more distributed across neural hubs, relative to the right hemisphere. In order to determine if this finding reflects a real interhemispheric difference in functional neuroanatomy or stems from differences in lesion patterns between the LHD and RHD cohorts, we computed the localization index for each brain model in the ground-truth simulation. There was no effect for the cohort on the localization index ($F(1) = 0.093$, $p = .761$), supporting functional interhemispheric differences as the origin of the observed hemispheric effects.

4 | DISCUSSION

In this methodological paper, we described in detail a newly revised multivariate approach to LSM based on the game-theoretical principles of *SVMSA*. Using as a test case, a data set of 107 stroke patients, derived from a larger cohort that participated in a recent VLSM study of UL paresis (Frenkel-Toledo et al., 2019), we showed the ability of this revised approach to correctly detect ground-truth brain–behavior relationship using moderately sized cohorts (50–60 patients), with acceptable type-I and type-II error rates. The results obtained by this method can provide useful information regarding the underlying brain network supporting the behavior of interest. The following sections discuss issues related to the use of MSA for LSM in general, and findings of MSA application in the specific case of stroke-related UL paresis.

4.1 | Issues concerning MSA for LSM in general

VLSM and related univariate LSM methods are used extensively for outlining the neural correlates of impairment following stroke (Bates et al., 2003; Karnath et al., 2019; Rorden & Karnath, 2004). However, as pointed out by several authors, these methods are unable to account properly for the intricate interactions that exist between brain regions comprising a functional network, and also fail to treat properly lesion dependencies dictated by the anatomy of vascular supply (Mah et al., 2014; Xu et al., 2018; Zhang et al., 2014). Multivariate LSM methods were suggested as a possible solution for these issues (Mah et al., 2014; Pustina et al., 2018; Xu et al., 2018), as they are better equipped for treating functional dependencies. Nonetheless, other authors (Ivanova et al., 2021; Sperber & Karnath, 2017) pointed out that multivariate LSM might also suffer from similar limitations and may not outperform classical univariate LSM methods. Moreover, a recent simulation study (Ivanova et al., 2021) has shown that multivariate LSM methods require a larger sample size to achieve the accuracy of univariate LSM methods.

MSA for neurological LSM, first introduced by Kaufman et al. (2009) in a study of the neural substrate of line-bisection performance in stroke patients with spatial neglect, is a promising multivariate

approach due to its enhanced capability to detect the importance of network elements which exhibit complex interactions with each other (Keinan et al., 2004, 2006). In contrast to other multivariate methods, MSA could yield neurologically sound results when applied on relatively small cohorts (Kaufman et al., 2009; Toba et al., 2017, 2020). Also, MSA could detect the importance of brain structures which are injured only in a small proportion of stroke patients (Kaufman et al., 2009; Zavaglia et al., 2015). Nevertheless, as shown in the current study, applying MSA for LSM is more complex than applying univariate LSM, since it consists of several steps, each one critical for the final outcome.

First, computation of the original (so called, "full-information") SVs requires the knowledge of the full set of possible binary working states, while the available data are graded. Previous solutions to this problem included dichotomization of either the lesion data (Kaufman et al., 2009; Zavaglia et al., 2015) or the behavioral scores (Malherbe et al., 2018; Toba et al., 2017) before applying a ML predictor, thus potentially loosing valuable information. Here, we proposed a predictor which was specially designed to predict binary working-states scores from graded lesion data, eliminating the need for dichotomization. Also, we took into account only the less-perturbed binary working states, thus decreasing unwanted error attributed to biologically unrealistic states, as was suggested previously (Keinan et al., 2006).

Second, it seems that the prediction process introduces a bias which leads to general shrinkage of SVs toward the mean SV (i.e., to the performance in the intact state, divided by the number of brain regions, see Figure S2 in the supplementary materials for a visual example). As explained in Section 2, we addressed this bias by applying a calibration procedure and by referring only to SVs of regions which are significantly higher than the average, as indicative of critical network hubs. Since SVs distribution is unknown, we constructed confidence intervals using the bootstrap method (Efron, 1981), as was done recently (Malherbe et al., 2021; Toba et al., 2020). Moreover, since each region is tested for significance separately, spurious results were minimized by correcting for multiple comparisons (Inoue et al., 2014; Ivanova et al., 2021; Mah et al., 2014; Pustina et al., 2018; Sperber & Karnath, 2017; Y. Zhang et al., 2014).

We have tested our suggested MSA approach to LSM using ground-truth simulation. Ground-truth simulation was used recently in several studies which evaluated different LSM methodologies (Inoue et al., 2014; Ivanova et al., 2021; Mah et al., 2014; Pustina et al., 2018; Sperber & Karnath, 2017; Zhang et al., 2014), and it is considered to be the closest to "gold standard" in evaluation of LSM methods (Xu et al., 2018). The number of variables that could be assessed in simulation studies of LSM is vast and includes parameters affecting the ground-truth model (e.g., type of lesion-behavior relationship at the single region level (Amunts et al., 1996; De Gennaro et al., 2004; Guye et al., 2003; Hammond et al., 2004; Tretriluxana et al., 2009)), interregional interactions (Godefroy, 1998; Toba et al., 2019), number of regions in the functional network, factors related to lesion dimensionality (the anatomical parcellation in use, the extent of interdependency between regions, the extent of representation of different regions in the cohort), and other factors (e.g., size of the cohort, extent of behavioral impairment, or smoothness of lesion borders). Beyond all that, there are parameters related to the LSM method itself (e.g., with or without correction for total lesion volume, the thresholds set for significance and method-specific hyper-parameters). Since our aim here was to demonstrate the method's applicability and validity, assessing all of these parameters was out of the scope of the current study. We hence focused on three questions in the simulation study: (1) what is the optimal perturbation depth (J parameter); (2) what are the differences in performance between left- and right-hemisphere damaged patients; and (3) what are the theoretical false-positive and false-negative error-rates associated with the optimal perturbation depth.

Using ground-truth simulation, we showed that the perturbation-depth of $J = 5$ was associated with the best performance for both LHD and RHD cohorts. This result was also supported by analysis of the clinical data, which showed that MSA with $J > 5$ detected almost the same regions as MSA with $J = 5$, for both cohorts (Figures S6 and S7; supplementary materials). Until an automated method for determining the optimal depth is developed, we suggest trying several depths and choosing the depth in which results are stable.

Second, we showed that different lesion-behavior data sets are associated with different patterns of performance. In the current example, the LHD data set, in spite of its larger size, was associated with decreased accuracy (as expressed by lower area under the ROC curve) in comparison to the RHD data set, for all assessed J values. This result might be related to greater heterogeneity in regional lesion extent in the RHD cohort compared with the LHD cohort in our data (standard deviation of regional lesion extent, mean ± SD: 15.3 ± 9.5 vs. 12.1 ± 7.2 in RHD and LHD groups, respectively; Wilcoxon rank signed test, $z = -3.34$, $p = .0008$).

Finally, we were specifically interested in measuring the false-positive and false-negative rates of the selected perturbation depth ($J = 5$) for both data sets. Our results showed that using FDR correction yielded reasonable false-positive error rates of 10 and 14.7% for the RHD and LHD cohorts, respectively. As expected, this conservativeness came with a cost of increased false-negative rate (25 and 26.7% for the RHD and LHD cohorts, respectively). While these error rates are based on a very specific ground-truth situation which may not represent the actual brain–behavior relationship, they provide some insight into accuracy of the method, when considering the results obtained for real behavioral data.

4.2 | Case study: MSA Of UL paresis poststroke

4.2.1 | Interhemispheric differences shown in VLSM and MSA

Given the fact that FMA-UE scores were obtained 8 weeks on average after stroke onset, we should assume that neuroplasticity—the capacity of the injured brain for reorganization and adaptive remapping of structure–function relationships—affected the emerging pattern of brain ROIs involved in the tested behavior (Carmichael &

Chesselet, 2002; Cheng et al., 2014; Darling et al., 2011; Grefkes & Ward, 2014; Page et al., 2013; Puig et al., 2011; Ward et al., 2003). In a recent VLSM study by our group (Frenkel-Toledo et al., 2019) which analyzed stroke lesion effects on UL function in the subacute period, we found that damage to numerous subcortical and cortical regions—BG, CR, internal capsule, insula, rolandic operculum, and inferior frontal gyrus—affected the level of impairment in the hemiparetic UL following RHD. A different pattern was found in LHD patients, in whom lesions to only a few subcortical regions—BG, CR, internal capsule, external capsule, and superior longitudinal fasciculus—exerted a significant impact on UL function, moreover with a lower signal (Z scores passed only a lenient criterion for significance). Although some hemispheric asymmetry was expected, based on previous knowledge (Amunts et al., 1996; De Gennaro et al., 2004; Guye et al., 2003; Hammond et al., 2004; Tretriluxana et al., 2009), its extent, etas revealed in this study, was overwhelming and called for verification by other modes of analysis.

In the current study, we used MSA to analyze LSM in 107 of the 130 subjects who participated in the above VLSM research (Frenkel-Toledo et al., 2019). The data of 23 subjects from the original cohort was trimmed in order to reduce variance related to behavioral testing time after stroke onset. Multivariate analysis by MSA, like the earlier univariate VLSM analysis in (Frenkel-Toledo et al., 2019), detected largely different sets of critical regions in the RHD and LHD cohorts, suggesting a significant interhemispheric difference in the brain networks that mediate motor function of the hemiparetic UL at this stage (i.e., the later part of the early subacute period poststroke, when much of the expected recovery has already been accomplished (Duncan et al., 1992; Kwakkel et al., 2006)).

4.2.2 | Insight obtained from use of a "localization index"

Analysis of the localization index provided an important viewpoint on the hemispheric differences revealed by MSA. The "localization index" is a measure of the extent in the function of interest that can be attributed to a single brain region (Aharonov et al., 2003). The localization index was significantly lower in the LHD cohort, pointing to a more homogenous contribution of the different hubs in the left motor network compared to the right motor network. This finding is likely to indicate that the hubs of the left motor network are more versatile in their capabilities, and possess a greater capacity to compensate for the loss incurred by damage to other hubs in the network. Previous research has demonstrated increased connectivity between motor hubs in the dominant left hemisphere of healthy right-handed subjects (Guye et al., 2003). Increased redundancy of a functional network makes the contribution of specific network hubs less salient and more difficult to characterize by LSM methods, especially univariate ones (Ivanova et al., 2021). This may explain why no voxels in the LHD cohort were detected after correction for multiple comparisons in our original VLSM study (Lo et al., 2010; Meyer et al., 2016).

4.2.3 | Possible sources of hemispheric differences

The hemispheric differences revealed in the current study could be contributed by the larger variance in regional lesion extent and the greater accuracy (area under the ROC curve) in the RHD group. However, when simulated synthetic behavioral scores were used instead of the actual ones, the difference in the localization index between the RHD and LHD cohorts disappeared (L_{LHD} vs. L_{RHD}, *mean* ± SD: 0.0231 ± 0.0103 vs. 0.0227 ± 0.0093, paired $t(119) = -0.57$, $p = .57$), pointing to brain–behavior relationships rather than lesion characteristics themselves as the main source of the observed hemispheric effects. Moreover, the average lesion extent, and also the rate of lesion involvement in most of the "critical" regions (i.e., brain regions pointed by MSA to have a high contribution to UL function), did not differ significantly when the two hemispheric groups were compared (the few regions which did differ were not found to be of high importance by MSA, see Figure 3). Thus, it is unlikely that differences in accuracy yielded the two non-overlapping sets of high-impact regions (SFG, SMA, and PCL in the LHD cohort; primary motor cortex, primary somatosensory cortex, BG, and ventral brainstem in the RHD cohort). Significant contribution to the hemispheric differences by demographic and clinical factors is even less probable, as subjects in both groups were recruited using the same strict inclusion criteria and no significant difference in either demographic or behavioral variables was found when comparing the two groups (Table 2). It should be noted that various lesion studies in the past treated the left and right motor systems as parallels, applying flipping techniques in order to increase the statistical power of analyses (e.g., Lo et al., 2010; Meyer et al., 2016). Obviously, this practice ignores the fact that persons do not show the same level of dexterity when doing things with the dominant and nondominant hands (Bagesteiro & Sainburg, 2002, 2003). Our earlier VLSM study (Frenkel-Toledo et al., 2019), in which LHD and RHD cohorts were analyzed as separate groups, was one of the first LSM studies that clearly demonstrated a salient group effect with respect to the impact of lesion configuration on FMA-UE scores. We propose that the hemispheric effects demonstrated in the earlier VLSM study and in the current MSA research, stem from differences in the functional neuroanatomy of the motor network in the (dominant) left hemisphere and in the right hemisphere. Calculation of the cSVs after extraction of seven patients (three LHD, four RHD) who were not right-handers, revealed essentially the same pattern as for the entire cohort (see Figure S8; supplementary materials) with some minor differences. Past studies have demonstrated different hemispheric contributions to motor control (Amunts et al., 1996; De Gennaro et al., 2004; Guye et al., 2003; Hammond et al., 2004; Serrien et al., 2006; Tretriluxana et al., 2009; Wu et al., 2015). Thus, although lesion-dimensionality factors might have influenced the MSA results, the findings most probably reflect a profound difference between the dominant and nondominant hemispheres in the organization of the motor networks that support UL functioning at this stage poststroke onset.

4.2.4 | Differences between MSA and VLSM results

The current MSA study points to a set of high-impact brain regions which is different from the set of regions in which the existence of damage exerted a significant effect on the FMA-UE scores in our earlier VLSM study (Frenkel-Toledo et al., 2019). Thus, in the LHD group, damage to just a few subcortical regions passed a lenient criterion of significance in the VLSM study, whereas MSA pointed to three cortical regions as highly contributing to performance (SMA, PCL, and SFG). In the RHD group, out of the regions which were significant according to VLSM (BG, CR, internal capsule, insula, rolandic operculum, and inferior frontal gyrus—all at the core of MCA supply zone), only the BG region was pointed also by MSA as highly contributing (MSA pointed also to three other regions—primary motor and primary somatosensory cortices and the ventral brainstem). Note that all the regions pointed by MSA (and not by VLSM) to be highly important for the functioning of the hemiparetic UL, are distant from the core MCA supply zone. This is certainly true for the trajectory of the corticospinal tract in the ventral brainstem, the dorsally located hand area in primary motor and primary somatosensory cortices, the SMA, the PCL, and the SFG. All these regions are involved relatively infrequently in stroke patients. As a consequence, the importance of damage there was less likely to emerge (and survive the needed correction for multiple comparisons) in our previous VLSM study (Frenkel-Toledo et al., 2019), where analysis was employed on voxel clusters involved in at least 10–20% of the studied cohort. Unlike VLSM, studies employing MSA for LSM were shown to detect important hubs of functional networks even when these are represented in low numbers. For example, in the MSA study of Kaufman et al. (2009) the superior parietal lobule, despite being involved infrequently in stroke, emerged as an important hub in the network mediating spatial attention. This finding, which emerged from MSA of LSM in a cohort of stroke patients with spatial neglect (Kaufman et al., 2009), is in accordance with theorizing based largely on fMRI research (Corbetta & Shulman, 2011).

4.2.5 | Relevance of MSA high-impact regions to motor behavior in the recovering brain

All the regions, both in RHD and LHD groups, that were pointed by MSA to be highly contributing to UL functioning (except for the PCL), were found in past research to affect UL motor behavior. Three of the high-impact regions in the RHD cohort (primary motor cortex, primary somatosensory cortex, VBS) are related to the corticospinal tract (CST), which is the major pathway mediating cortical control of spinal motor activity, with cells of origin residing largely within the boundaries of the sensory-motor cortex. Previous studies delineated the importance of CST integrity for recovery from UL hemiparesis following stroke (Byblow et al., 2015; Feng et al., 2015; Stinear et al., 2007; J. Xu et al., 2017), with specific reference to the primary motor cortex (Kaya et al., 2015), the primary somatosensory cortex (Abela et al., 2012; Payabvash et al., 2012), and the CST itself in its passage through the CR (Lo et al., 2010), the PLIC (Schiemanck et al., 2008; Stinear et al., 2012; Wenzelburger et al., 2005; Zhu et al., 2010), and the ventral brainstem (Oh et al., 2012). The BG, though not involved directly in movement execution, play an important role in motor learning and selective activation of motor engrams appropriate for the task in question, probably by specification of a cost-function related to a learned movement gain (Turner & Desmurget, 2010). Past fMRI studies demonstrated hemispheric differences in BG activity (e.g., Molochnikov & Cohen, 2014; Zhang et al., 2017).

Several ROIs were found to contribute significantly to UL motor function in the LHD cohort. The "SMA" ROI in our parcellation included also the adjacent medial part of the SFG (FSMG). These two regions were grouped together based on their anatomical continuity and common blood supply (from the anterior cerebral artery). Since the combined ROI was found to contribute significantly to motor function in LHD cohort, the analysis was repeated with the SMA and FSMG separated to explore their individual contributions. Indeed, in both regions the SVs indicated high contribution to UL motor function in the LHD cohort (Figure S9). The SMA is known to possess vast connectivity with other motor areas (Luppino et al., 1993; Potgieser et al., 2014) and is implicated specifically in volitional, goal-directed or "top-down" movement planning (Chen et al., 2010; Tanji & Shima, 1994). The FSMG was shown in past research to be involved in skilled action anticipation (Xu et al., 2016). Thus, the high SVs of these brain regions in the LHD cohort reflect their important role in movement preparation in the dominant left hemisphere. Another high-impact region in the LHD group—the SFG—contains part of the lateral premotor cortex which, like the SMA, plays an important role in the motor planning stage preceding the execution of movement (Chouinard & Paus, 2006; Dum & Strick, 2002; Wise, 1985). The PCL—another high-impact cortical region in the LHD group—which is located more posteriorly than the former regions, controls, in the intact brain, voluntary movement of the lower limb. Yet, it is not unlikely that following damage to the brain, this region (among other perilesional cortical regions) may take part in a process of adaptive remapping and assume UL control functions. The prominence of medial and dorsolateral premotor cortices (SMA and SFG) in the Shapley analysis of the LHD cohort (but not the RHD cohort) points to a more salient involvement of these regions in motor control processes of the dominant left hemisphere. Indeed, differential effects of the left and right premotor areas were demonstrated in past studies, for example, in the control of memory-guided saccades (Gaymard et al., 1993) or the planning of object prehension (Martin et al., 2011). Increased premotor activity poststroke (Seitz et al., 1998) is thought to play a role in recovery processes, either through adaptive remapping (Plow et al., 2015) or as an alternative sources for mediation of cortical control over spinal motor activity when the CST is severely compromised (Li et al., 2022).

4.2.6 | Concluding remarks and interpretation

We propose that the differences between the hemispheric groups revealed by MSA—the prominence of network hubs related to motor execution in the RHD group and of network hubs related to motor

planning in the LHD group, and the more even contribution of network hubs in the LHD group—are connected to the different functional neuroanatomy of the dominant and nondominant hemispheres. Motor dominance is presumably based upon higher computational reserves, leading to an improved ability of intact regions of the damaged left hemisphere to take over the roles of inactive regions. The neural/computational reserve of the dominant left hemisphere seems to be more prone to adaptive remapping and activity-dependent neuroplastic processes taking place in the subacute phase after stroke (Gauthier et al., 2008) and during active rehabilitation (Sawaki et al., 2008). This may be the reason why LHD patients tend to use their contralesional hemiparetic hand more than patients with RHD, for the same level of impairment (Yadav et al., 2019) and may explain the tendency of LHD patients to show somewhat greater improvement in the hemiparetic UL compared to RHD patients (Frenkel-Toledo et al., 2019). In the case of UL paresis poststroke, MSA provided theoretically sound results which could not be obtained using univariate (VLSM) analysis, due to the seemingly improved ability of MSA to extract useful information regarding infrequently involved brain regions and its more efficient statistical framework.

4.3 | Methodological limitations

4.3.1 | Issues related to lesion dimensionality

Lesion dimensionality is a major concern for all LSM methods, including multivariate ones (Pustina et al., 2018), and MSA is not exempt. Three potential sources of error, disclosed in the current case study, should be emphasized: (1) Relatively lower variance in regional lesion extent in the LHD group compared to the RHD cohort, possibly contributing to the observed difference in accuracy between the cohorts. (2) High interregional correlation in the rate of involvement for some region pairs, based on similarity of vascular supply (e.g., SMA and PCL, both supplied by the anterior cerebral artery). In such cases, if the two regions emerge as highly contributing to the tested behavior (as found here in the LHD cohort for the two regions in the above example), the contribution attributed to one of the regions may reflect a type-I error. (3) When applying MSA on moderate-size cohorts (as in the current study), the spatial resolution (number of ROIs) of the anatomical parcellation is limited. As explained in Section 2, for the current case study, we reduced significantly the number of atlas ROIs, by pooling together regions that were not found in past research to be involved in mediation of UL motor function. Thus, different such regions were grouped in one ROI under the designation "RoB," as was done previously (Malherbe et al., 2018; Toba et al., 2017, 2020). This ROI, in both hemispheres, exhibited near-average SVs, suggesting a low probability that high-impact structures were included there. However, failed appreciation of regional importance is still a possibility in such a case, as pooling together different regions might mask the importance of one of them, as its contribution is diluted in the overall impact of the much larger aggregate region. Moreover, even without grouping of brain regions, any single ROI may contain more than one distinct functional neural hub. For example, the precentral gyrus as defined in the AAL atlas (Tzourio-Mazoyer et al., 2002) includes the primary motor representations of all body parts without discrimination. Thus, a certain lesion load might produce variable behaviors, depending on the exact locus of involvement within the precentral gyrus. This conjecture is supported by the fact that some bootstrap-based distribution of SVs was bimodal for some ROIs (see Figures S3 and S4). On the other hand, keeping spatial resolution low allowed the representation of sufficient number of cases with damage to each brain region (in the current case study, the minimal involvement rate was 10% of the LHD cohort [for the superior parietal lobule] and 14% of the RHD cohort [for the SMA]), which is similar to minimal regional involvement rate used in previous studies (Zhang et al., 2014).

Altogether, these lesion dimensionality issues might be responsible for the modest false negative (25 and 26.7%) and false positive (10.3 and 14.7%) rates (for simulated RHD and LHD, respectively). Measurement of these metrics at the voxel level by Inoue et al. (2014) yielded a false-negative rate of 48% and a false-positive rate of 5% (averages across sample sizes from 25 to 150). Another extensive simulation study (Ivanova et al., 2021), which assessed several sample sizes, reported a voxel-level false-negative rate of ~1 to ~12.5%, but this was associated with a false-positive rate of 70–95% (for a sample size of 48 patients). Although these voxel-level measures cannot be directly compared to our region-level error-rates, they provide some reference suggesting that MSA do not underperform commonly used LSM methods in this respect. A comprehensive simulation study, using the same brain-model, is warranted to determine if MSA can outperform VLSM, and in which scenarios.

4.3.2 | Issues related to functional dimensionality

The SV assigned to a brain region on the basis of LSM analysis points to its relative contribution to task performance. However, this knowledge is not enough for decoding the exact role played by the region in the examined functional network, as the overall state of the network at the time of testing deviates from the normal state, as a consequence of a poorly understood, multifactorial, recovery process. In the current case study, FMA scores were obtained in the later stage of the early subacute period (later than a month post-onset). At this time, most of the functional amelioration which is attributable to true behavioral recovery (improvement obtained in measures of basic impairments, e.g., in the capacity to execute isolated, out of synergy, movements, as examined by the FMA) already occurred (Duncan et al., 1992). Thus, patients' FMA scores at this time reflect two distinct processes—(a) spared activity in the undamaged parts of the motor network, and (b) the outcome of biologically driven and treatment-related adaptive and maladaptive plasticity, that took place from the onset of stroke until then, within and outside the original extent of the motor network (Gauthier et al., 2008). There is no direct way to clearly delineate one effect from the other (i.e., whether a high SV of a given structure reflects an important contribution to performance conveyed by the activity of intact elements in the structure, or by plastic changes within that structure).

It is of interest that MSA did not point to some regions where damage was shown in past VLSM studies, including our own study (Frenkel-Toledo et al., 2019) to have a significant impact on residual UL motor function, notably, the PLIC (Fries et al., 1993; Puig et al., 2011), CR (Lo et al., 2010; Shelton & Reding, 2001), insula, opercular regions, and association tracts (Payabvash et al., 2012; Wu et al., 2015). As in other lesion-behavior inference methods, it is impossible to draw clear conclusions regarding regions that do not pass the significance threshold (Qian et al., 2020). For example, in the LHD group, the PLIC and the ventral brainstem (both traversed by the CST, thus certainly important), were assigned high cSVs (Figure 6), which, however, (due to large variability in SVs across subjects) did not differ significantly from the computed reference value—that is, the SV a region would receive if all regions contribute equally to task performance. It is possible that the above regions might have been detected using larger cohorts, enabling use of higher spatial resolution in anatomical description.

The interpretation of SVs (whether high or low) without taking into account the network (coalitional game) context might be oversimplistic. In as much as the goal of a functional network (or of a subunit of a network) is achievable via alternative tracks, the network is more likely to resist well the impact of focal damage. This will be reflected by a lower correlation between regional damage and performance, as the functional impact of focal damage may be rapidly restituted, at least partially, by execution via alternative routes, or by the creation of new routes in other regions through adaptive plasticity. On the other hand, the hallmark of a network hub that plays a role based on a unique obligatory connectivity pattern, is noneffective restitution following its damage. Anatomical structures playing such a role will be associated with higher SVs (Keinan et al., 2006). In the current case study, the set of regions with high SVs should be viewed as neural hubs of the motor network whose role in UL function could not be compensated effectively by other regions, and not as sole determinants of motor function in the hemiparetic UL.

An SV not surviving the threshold set for significance may reflect several lesion-behavior relationships: a hub which is highly contributing in the intact brain, but replaceable through adaptive remapping in the recovering brain; a hub which was not damaged enough in the studied cohort to express its actual contribution; or a hub which is truly not concerned with the function of interest. One should bear in mind that SVs are assigned to brain regions that in most cases are damaged in a partial manner. For example, in the current case study, the precentral gyrus (subjected here to MSA as a single anatomical ROI) did not emerge as a significant network hub in the LHD group. This could stem from the high degree of discrete somatotopism in the primary motor cortex, making the effect of damage to the dorsally located (and relatively infrequently involved) "hand area" in the motor homunculus less salient. Most of the patients suffer from lesions to ventral parts of the precentral gyrus, which are considered irrelevant to UL motor control in the intact brain. Also, since perilesional regions usually possess connectivity patterns closer to those lost by the stroke, and as a consequence are better candidates for adaptive remapping than distant regions (Nudo & Milliken, 1996), perilesional regions could be assigned low SVs due to plasticity (blurring the effect of damage to the region). Uncovering this effect requires longitudinal assessment along the first weeks and months poststroke onset.

Finally, the interpretation of negative SVs deserves caution, as such values may reflect an inhibitory role of brain structures (Toba et al., 2017, 2019; Zavaglia & Hilgetag, 2016), as in the case of reciprocal inhibition between homologous collicular nuclei, in the midbrain system supporting bottom-up allocation of attention (Sprague & Meikle, 1965).

4.3.3 | Issues related to external validity

Beyond the above general methodological limitations of MSA for neurological LSM, one should note that LSM studies, regardless of the method of analysis employed, are prone to several kinds of sample biases. In the current study, stroke patients with severe language and/or cognitive disturbances precluding full compliance with task requirements, were not included. On the other hand, stroke patients with very mild impairment are unlikely to be represented in our cohort, which was comprised of patients referred to our hospital for in-patient rehabilitation. Thus, the sample used here to test the application of MSA for neurological LSM does not represent the very severe nor the very mild cases one sees in the general stroke population. In addition, we recruited here only subjects who were in the later stage of the early subacute period poststroke (1–3 months post onset), as the process of recovery was shown in previous studies to reach a relatively stationary level at this stage (Duncan et al., 1992; Kwakkel et al., 2006). Yet, other studies (e.g., Ward et al., 2019) showed that some dynamics in the level of impairment can be shown even at later stages, thus the picture of the relative contribution of different brain structures to the residual motor function of the hemiparetic UL is incomplete and lacks such late contributions.

4.4 | Summary

We described a revised approach to MSA, addressing various theoretical and methodological issues in application of the method for LSM. First, we demonstrated the method's ability to detect lesion-behavior relationships in ground-truth simulation. This enabled validation of the approach and determination of the optimal perturbation depth to be used in a real-life sample, comprised of stroke patients for whom we had both normalized lesion data and a quantitative measure of impairment in the hemiparetic UL. With this sample serving as a test case, MSA disclosed important differences between the motor network mediating UL function in the left hemisphere and in the right hemisphere. The left (dominant) motor network was found to possess a lower "localization index" (i.e., a more homogenous distribution of valence among the network hubs) compared to the right motor network. MSA also pointed to marked differences in the more prominent network hubs in each hemisphere. Thus, brain regions involved in motor planning had a more salient load on UL function in the LHD

group and brain regions involved in motor execution had a more salient load on UL function in the RHD group. The results of our test case demonstrate that the current approach to MSA can provide a unique insight into the relative importance of different hubs in neural networks, which is difficult to obtain using standard univariate methods.

ACKNOWLEDGMENTS

This study was part of the requirements from the author SOG for attaining Physical Medicine and Rehabilitation specialist certification. The authors wish to thank Gil Fridbeg, Justine Lowenthal-Raz, Gadi Bartur, Osnat Granot, and Shirley Handelzalts for their assistance in data collection.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available as a MATLAB data file in GitHub at https://github.com/ShayOfir/MSA.

ORCID

Shay Ofir-Geva https://orcid.org/0000-0001-7871-2778
Isaac Meilijson https://orcid.org/0000-0001-7825-9053
Silvi Frenkel-Toledo https://orcid.org/0000-0001-8991-3428
Nachum Soroker https://orcid.org/0000-0003-4909-5576

REFERENCES

Abela, E., Missimer, J., Wiest, R., Federspiel, A., Hess, C., Sturzenegger, M., & Weder, B. (2012). Lesions to primary sensory and posterior parietal cortices impair recovery from hand paresis after stroke. *PLoS One*, *7*(2), e31275. https://doi.org/10.1371/journal.pone.0031275

Aharonov, R., Segev, L., Meilijson, I., & Ruppin, E. (2003). Localization of function via lesion analysis. *Neural Computation*, *15*(4), 885–913. https://doi.org/10.1162/08997660360581949

Amunts, K., Schlaug, G., Schleicher, A., Steinmetz, H., Dabringhaus, A., Roland, P. E., & Zilles, K. (1996). Asymmetry in the human motor cortex and handedness. *NeuroImage*, *4*(3 Pt 1), 216–222. https://doi.org/10.1006/nimg.1996.0073

Bagesteiro, L. B., & Sainburg, R. L. (2002). Handedness: Dominant arm advantages in control of limb dynamics. *Journal of Neurophysiology*, *88*(5), 2408–2421. https://doi.org/10.1152/jn.00901.2001

Bagesteiro, L. B., & Sainburg, R. L. (2003). Nondominant arm advantages in load compensation during rapid elbow joint movements. *Journal of Neurophysiology*, *90*(3), 1503–1513. https://doi.org/10.1152/jn.00189.2003

Banzhaf, J. F. (1965). Weighted voting doesn't work: A mathematical analysis. *Rutgers Law Review*, *19*(2), 317–343.

Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., & Dronkers, N. F. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, *6*(5), 448–450. https://doi.org/10.1038/nn1050

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B*, *57*(1), 289–300.

Brown, P. J. (1993). *Measurement, regression, and calibration*. Clarendon Press.

Buneo, C. A., & Andersen, R. A. (2006). The posterior parietal cortex: Sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia*, *44*(13), 2594–2606. https://doi.org/10.1016/j.neuropsychologia.2005.10.011

Butnariu, D. (1980). Stability and Shapley value for an n-persons fuzzy game. *Fuzzy Sets and Systems*, *4*(1), 63–72. https://doi.org/10.1016/0165-0114(80)90064-0

Byblow, W. D., Stinear, C. M., Barber, P. A., Petoe, M. A., & Ackerley, S. J. (2015). Proportional recovery after stroke depends on corticomotor integrity: Proportional recovery after stroke. *Annals of Neurology*, *78*(6), 848–859. https://doi.org/10.1002/ana.24472

Carmichael, S. T., & Chesselet, M.-F. (2002). Synchronous neuronal activity is a signal for axonal sprouting after cortical lesions in the adult. *The Journal of Neuroscience*, *22*(14), 6062–6070. https://doi.org/10.1523/JNEUROSCI.22-14-06062.2002

Chen, X., Scangos, K. W., & Stuphorn, V. (2010). Supplementary motor area exerts proactive and reactive control of arm movements. *The Journal of Neuroscience*, *30*(44), 14657–14675. https://doi.org/10.1523/JNEUROSCI.2669-10.2010

Cheng, B., Forkert, N. D., Zavaglia, M., Hilgetag, C. C., Golsari, A., Siemonsen, S., Fiehler, J., Pedraza, S., Puig, J., Cho, T.-H., Alawneh, J., Baron, J.-C., Ostergaard, L., Gerloff, C., & Thomalla, G. (2014). Influence of stroke infarct location on functional outcome measured by the modified Rankin scale. *Stroke*, *45*(6), 1695–1702. https://doi.org/10.1161/STROKEAHA.114.005152

Chouinard, P. A., & Paus, T. (2006). The primary motor and premotor areas of the human cerebral cortex. *The Neuroscientist*, *12*(2), 143–152. https://doi.org/10.1177/1073858405284255

Corbetta, M., & Shulman, G. L. (2011). Spatial neglect and attention networks. *Annual Review of Neuroscience*, *34*, 569–599. https://doi.org/10.1146/annurev-neuro-061010-113731

Darling, W. G., Pizzimenti, M. A., & Morecraft, R. J. (2011). Functional recovery following motor cortex lesions in non-human primates: Experimental implications for human stroke patients. *Journal of Integrative Neuroscience*, *10*(3), 353–384. https://doi.org/10.1142/S0219635211002737

Dawid, A. P. (1982). The well-calibrated Bayesian. *Journal of the American Statistical Association*, *77*(379), 605–610. https://doi.org/10.1080/01621459.1982.10477856

De Gennaro, L., Cristiani, R., Bertini, M., Curcio, G., Ferrara, M., Fratello, F., Romei, V., & Rossini, P. M. (2004). Handedness is mainly associated with an asymmetry of corticospinal excitability and not of transcallosal inhibition. *Clinical Neurophysiology*, *115*(6), 1305–1312. https://doi.org/10.1016/j.clinph.2004.01.014

Dum, R., & Strick, P. (2002). Motor areas in the frontal lobe of the primate. *Physiology & Behavior*, *77*(4–5), 677–682. https://doi.org/10.1016/S0031-9384(02)00929-0

Duncan, P. W., Goldstein, L. B., Matchar, D., Divine, G. W., & Feussner, J. (1992). Measurement of motor recovery after stroke. Outcome assessment and sample size requirements. *Stroke*, *23*(8), 1084–1089.

Duncan, P. W., Propst, M., & Nelson, S. G. (1983). Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Physical Therapy*, *63*(10), 1606–1610.

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, *68*(3), 589–599. https://doi.org/10.1093/biomet/68.3.589

Efron, B., & Stein, C. (1981). The jackknife estimate of variance. *The Annals of Statistics*, *9*(3), 586–596. https://doi.org/10.1214/aos/1176345462

Efron, B., & Tibshirani, R. (1998). *An introduction to the bootstrap (Nachdr.)*. Chapman & Hall.

Feng, W., Wang, J., Chhatbar, P. Y., Doughty, C., Landsittel, D., Lioutas, V.-A., Kautz, S. A., & Schlaug, G. (2015). Corticospinal tract lesion load: An imaging biomarker for stroke motor outcomes. *Annals of Neurology*, *78*(6), 860–870. https://doi.org/10.1002/ana.24510

Forkert, N. D., Verleger, T., Cheng, B., Thomalla, G., Hilgetag, C. C., & Fiehler, J. (2015). Multiclass support vector machine-based lesion mapping predicts functional outcome in ischemic stroke patients. *PLoS One*, *10*(6), e0129569. https://doi.org/10.1371/journal.pone.0129569

Frenkel-Toledo, S., Fridberg, G., Ofir, S., Bartur, G., Lowenthal-Raz, J., Granot, O., Handelzalts, S., & Soroker, N. (2019). Lesion location impact on functional recovery of the hemiparetic upper limb. *PLoS One*, *14*(7), e0219738. https://doi.org/10.1371/journal.pone.0219738

Friedrich, S., Konietschke, F., & Pauly, M. (2019). Resampling-based analysis of multivariate data and repeated measures designs with the R package MANOVA.RM. *The R Journal*, *11*(2), 380. https://doi.org/10.32614/RJ-2019-051

Fries, W., Danek, A., Scheidtmann, K., & Hamburger, C. (1993). Motor recovery following capsular stroke. Role of descending pathways from multiple motor areas. *Brain A: Journal of Neurology*, *116*(Pt 2), 369–382.

Gauthier, L. V., Taub, E., Perkins, C., Ortmann, M., Mark, V. W., & Uswatte, G. (2008). Remodeling the brain: Plastic structural brain changes produced by different motor therapies after stroke. *Stroke*, *39*(5), 1520–1525. https://doi.org/10.1161/STROKEAHA.107.502229

Gaymard, B., Rivaud, S., & Pierrot-Deseilligny, C. (1993). Role of the left and right supplementary motor areas in memory-guided saccade sequences. *Annals of Neurology*, *34*(3), 404–406. https://doi.org/10.1002/ana.410340317

Gil, M., & Goral, M. (2004). Nonparallel recovery in bilingual aphasia: Effects of language choice, language proficiency, and treatment. *International Journal of Bilingualism*, *8*(2), 191–219. https://doi.org/10.1177/13670069040080020501

Gladstone, D. J., Danells, C. J., & Black, S. E. (2002). The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties. *Neurorehabilitation and Neural Repair*, *16*(3), 232–240. https://doi.org/10.1177/154596802401105171

Godefroy, O. (1998). Brain-behaviour relationships. Some models and related statistical procedures for the study of brain-damaged patients. *Brain*, *121*(8), 1545–1556. https://doi.org/10.1093/brain/121.8.1545

Grefkes, C., & Ward, N. S. (2014). Cortical reorganization after stroke: How much and how functional? *The Neuroscientist*, *20*(1), 56–70. https://doi.org/10.1177/1073858413491147

Groenewegen, H. J. (2003). The basal ganglia and motor control. *Neural Plasticity*, *10*(1–2), 107–120. https://doi.org/10.1155/NP.2003.107

Guye, M., Parker, G. J. M., Symms, M., Boulby, P., Wheeler-Kingshott, C. A. M., Salek-Haddadi, A., Barker, G. J., & Duncan, J. S. (2003). Combined functional MRI and tractography to demonstrate the connectivity of the human primary motor cortex in vivo. *NeuroImage*, *19*(4), 1349–1360.

Halligan, P. W., Cockburn, J., & Wilson, B. A. (1991). The behavioural assessment of visual neglect. *Neuropsychological Rehabilitation*, *1*(1), 5–32. https://doi.org/10.1080/09602019108401377

Hammond, G., Faulkner, D., Byrnes, M., Mastaglia, F., & Thickbroom, G. (2004). Transcranial magnetic stimulation reveals asymmetrical efficacy of intracortical circuits in primary motor cortex. *Experimental Brain Research*, *155*(1), 19–23. https://doi.org/10.1007/s00221-003-1696-x

Haramati, S., Soroker, N., Dudai, Y., & Levy, D. A. (2008). The posterior parietal cortex in recognition memory: A neuropsychological study. *Neuropsychologia*, *46*(7), 1756–1766. https://doi.org/10.1016/j.neuropsychologia.2007.11.015

Hart, S., & Mas-Colell, A. (1989). Potential, value, and consistency. *Econometrica*, *57*(3), 589–614.

Hastie, T., & Tibshirani, R. (1998). Classification by pairwise coupling. *The Annals of Statistics*, *26*(2), 451–471. https://doi.org/10.1214/aos/1028144844

Hoffman, D. S., & Strick, P. L. (1995). Effects of a primary motor cortex lesion on step-tracking movements of the wrist. *Journal of Neurophysiology*, *73*(2), 891–895. https://doi.org/10.1152/jn.1995.73.2.891

Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2008). Implementing a class of permutation tests: The coin package. *Journal of Statistical Software*, *28*(1), 1–23. https://doi.org/10.18637/jss.v028.i08

Inoue, K., Madhyastha, T., Rudrauf, D., Mehta, S., & Grabowski, T. (2014). What affects detectability of lesion–deficit relationships in lesion studies? *NeuroImage: Clinical*, *6*, 388–397. https://doi.org/10.1016/j.nicl.2014.10.002

Ivanova, M. V., Herron, T. J., Dronkers, N. F., & Baldo, J. V. (2021). An empirical comparison of univariate versus multivariate methods for the analysis of brain–behavior mapping. *Human Brain Mapping*, *42*(4), 1070–1101. https://doi.org/10.1002/hbm.25278

Karnath, H.-O., Sperber, C., & Rorden, C. (2019). Reprint of: Mapping human brain lesions and their functional consequences. *NeuroImage*, *190*, 4–13. https://doi.org/10.1016/j.neuroimage.2019.01.044

Kaufman, A., Keinan, A., Meilijson, I., Kupiec, M., & Ruppin, E. (2005). Quantitative analysis of genetic and neuronal multi-perturbation experiments. *PLoS Computational Biology*, *1*(6), e64. https://doi.org/10.1371/journal.pcbi.0010064

Kaufman, A., Kupiec, M., & Ruppin, E. (2004). Multi-knockout genetic network analysis: The Rad6 example. Proceedings. IEEE Computational Systems Bioinformatics Conference. 332–340.

Kaufman, A., Serfaty, C., Deouell, L. Y., Ruppin, E., & Soroker, N. (2009). Multiperturbation analysis of distributed neural networks: The case of spatial neglect. *Human Brain Mapping*, *30*(11), 3687–3695. https://doi.org/10.1002/hbm.20797

Kaya, D., Dincer, A., Arman, F., Bakirci, N., Erzen, C., & Pamir, M. N. (2015). Ischemic involvement of the primary motor cortex is a prognostic factor in acute stroke. *International Journal of Stroke*, *10*(8), 1277–1283. https://doi.org/10.1111/j.1747-4949.2011.00640.x

Keinan, A., Sandbank, B., Hilgetag, C. C., Meilijson, I., & Ruppin, E. (2004). Fair attribution of functional contribution in artificial and biological networks. *Neural Computation*, *16*(9), 1887–1915. https://doi.org/10.1162/0899766041336387

Keinan, A., Sandbank, B., Hilgetag, C. C., Meilijson, I., & Ruppin, E. (2006). Axiomatic scalable neurocontroller analysis via the Shapley value. *Artificial Life*, *12*(3), 333–352. https://doi.org/10.1162/artl.2006.12.3.333

Koch, P., Schulz, R., & Hummel, F. C. (2016). Structural connectivity analyses in motor recovery research after stroke. *Annals of Clinical and Translational Neurology*, *3*(3), 233–244. https://doi.org/10.1002/acn3.278

Kwakkel, G., Kollen, B., & Twisk, J. (2006). Impact of time on improvement of outcome after stroke. *Stroke*, *37*(9), 2348–2353. https://doi.org/10.1161/01.STR.0000238594.91938.1e

Li, Z., Hu, J., Wang, Z., You, R., & Cao, D. (2022). Basal ganglia stroke is associated with altered functional connectivity of the left inferior temporal gyrus. *Journal of Neuroimaging*, *32*, 744–751. https://doi.org/10.1111/jon.12978

Lo, R., Gitelman, D., Levy, R., Hulvershorn, J., & Parrish, T. (2010). Identification of critical areas for motor function recovery in chronic stroke subjects using voxel-based lesion symptom mapping. *NeuroImage*, *49*(1), 9–18. https://doi.org/10.1016/j.neuroimage.2009.08.044

Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *The Journal of Comparative Neurology*, *338*(1), 114–140. https://doi.org/10.1002/cne.903380109

Mah, Y.-H., Husain, M., Rees, G., & Nachev, P. (2014). Human brain lesion-deficit inference remapped. *Brain*, *137*(9), 2522–2531. https://doi.org/10.1093/brain/awu164

Malherbe, C., Cheng, B., Königsberg, A., Cho, T.-H., Ebinger, M., Endres, M., Fiebach, J. B., Fiehler, J., Galinovic, I., Puig, J., Thijs, V.,

Lemmens, R., Muir, K. W., Nighoghossian, N., Pedraza, S., Simonsen, C. Z., Wouters, A., Gerloff, C., Hilgetag, C. C., & Thomalla, G. (2021). Game-theoretical mapping of fundamental brain functions based on lesion deficits in acute stroke. *Brain Communications*, *3*(3), fcab204. https://doi.org/10.1093/braincomms/fcab204

Malherbe, C., Umarova, R. M., Zavaglia, M., Kaller, C. P., Beume, L., Thomalla, G., Weiller, C., & Hilgetag, C. C. (2018). Neural correlates of visuospatial bias in patients with left hemisphere stroke: A causal functional contribution analysis based on game theory. *Neuropsychologia*, *115*, 142–153. https://doi.org/10.1016/j.neuropsychologia.2017.10.013

Martin, K., Jacobs, S., & Frey, S. H. (2011). Handedness-dependent and -independent cerebral asymmetries in the anterior intraparietal sulcus and ventral premotor cortex during grasp planning. *NeuroImage*, *57*(2), 502–512. https://doi.org/10.1016/j.neuroimage.2011.04.036

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283–298. https://doi.org/10.1016/s0001-2998(78)80014-2

Meyer, S., Kessner, S. S., Cheng, B., Bönstrup, M., Schulz, R., Hummel, F. C., De Bruyn, N., Peeters, A., Van Pesch, V., Duprez, T., Sunaert, S., Schrooten, M., Feys, H., Gerloff, C., Thomalla, G., Thijs, V., & Verheyden, G. (2016). Voxel-based lesion-symptom mapping of stroke lesions underlying somatosensory deficits. *NeuroImage: Clinical*, *10*, 257–266. https://doi.org/10.1016/j.nicl.2015.12.005

Molochnikov, I., & Cohen, D. (2014). Hemispheric differences in the mesostriatal dopaminergic system. *Frontiers in Systems Neuroscience*, *8*, 1–14. https://doi.org/10.3389/fnsys.2014.00110

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., Woods, R., Toga, A. W., Pike, G. B., Neto, P. R., Evans, A., Zhang, J., Huang, H., Miller, M. I., van Zijl, P., & Mazziotta, J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*, *40*(2), 570–582. https://doi.org/10.1016/j.neuroimage.2007.12.035

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, *9*(11), 856–869. https://doi.org/10.1038/nrn2478

Nudo, R. J., & Milliken, G. W. (1996). Reorganization of movement representations in primary motor cortex following focal ischemic infarcts in adult squirrel monkeys. *Journal of Neurophysiology*, *75*(5), 2144–2149. https://doi.org/10.1152/jn.1996.75.5.2144

Oh, S., Bang, O. Y., Chung, C.-S., Lee, K. H., Chang, W. H., & Kim, G.-M. (2012). Topographic location of acute pontine infarction is associated with the development of progressive motor deficits. *Stroke*, *43*(3), 708–713. https://doi.org/10.1161/STROKEAHA.111.632307

Page, S. J., Gauthier, L. V., & White, S. (2013). Size doesn't matter: Cortical stroke lesion volume is not associated with upper extremity motor impairment and function in mild, chronic hemiparesis. *Archives of Physical Medicine and Rehabilitation*, *94*(5), 817–821. https://doi.org/10.1016/j.apmr.2013.01.010

Payabvash, S., Souza, L. C. S., Kamalian, S., Wang, Y., Passanese, J., Kamalian, S., Fung, S. H., Halpern, E. F., Schaefer, P. W., Gonzalez, R. G., Furie, K. L., & Lev, M. H. (2012). Location-weighted CTP analysis predicts early motor improvement in stroke: A preliminary study. *Neurology*, *78*(23), 1853–1859. https://doi.org/10.1212/WNL.0b013e318258f799

Platz, T., Pinkowski, C., van Wijck, F., Kim, I.-H., di Bella, P., & Johnson, G. (2005). Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer test, action research arm test and box and block test: A multicentre study. *Clinical Rehabilitation*, *19*(4), 404–411. https://doi.org/10.1191/0269215505cr832oa

Plow, E. B., Cunningham, D. A., Varnerin, N., & Machado, A. (2015). Rethinking stimulation of the brain in stroke rehabilitation: Why higher motor areas might be better alternatives for patients with greater impairments. *The Neuroscientist*, *21*(3), 225–240. https://doi.org/10.1177/1073858414537381

Potgieser, A. R. E., de Jong, B. M., Wagemakers, M., Hoving, E. W., & Groen, R. J. M. (2014). Insights from the supplementary motor area syndrome in balancing movement initiation and inhibition. *Frontiers in Human Neuroscience*, *8*, 1–11. https://doi.org/10.3389/fnhum.2014.00960

Puig, J., Pedraza, S., Blasco, G., Daunis-I-Estadella, J., Prados, F., Remollo, S., Prats-Galino, A., Soria, G., Boada, I., Castellanos, M., & Serena, J. (2011). Acute damage to the posterior limb of the internal capsule on diffusion tensor tractography as an early imaging predictor of motor outcome after stroke. *AJNR. American Journal of Neuroradiology*, *32*(5), 857–863. https://doi.org/10.3174/ajnr.A2400

Pustina, D., Avants, B., Faseyitan, O. K., Medaglia, J. D., & Coslett, H. B. (2018). Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. *Neuropsychologia*, *115*, 154–166. https://doi.org/10.1016/j.neuropsychologia.2017.08.027

Qian, S. S., Refsnider, J. M., Moore, J. A., Kramer, G. R., & Streby, H. M. (2020). All tests are imperfect: Accounting for false positives and false negatives using Bayesian statistics. *Heliyon*, *6*(3), e03571. https://doi.org/10.1016/j.heliyon.2020.e03571

Quenouille, M. H. (1949). Problems in plane sampling. *The Annals of Mathematical Statistics*, *20*(3), 355–375. https://doi.org/10.1214/aoms/1177729989

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika*, *43*(3/4), 353. https://doi.org/10.2307/2332914

Rodríguez-Herreros, B., Amengual, J. L., Gurtubay-Antolín, A., Richter, L., Jauer, P., Erdmann, C., Schweikard, A., López-Moliner, J., Rodríguez-Fornells, A., & Münte, T. F. (2015). Microstructure of the superior longitudinal fasciculus predicts stimulation-induced interference with online motor control. *NeuroImage*, *120*, 254–265. https://doi.org/10.1016/j.neuroimage.2015.06.070

Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Reviews. Neuroscience*, *5*(10), 813–819. https://doi.org/10.1038/nrn1521

RStudio Team. (2015). *RStudio: Integrated development environment for R*. RStudio http://www.rstudio.com/

Sawaki, L., Butler, A. J., Leng, X., Wassenaar, P. A., Mohammad, Y. M., Blanton, S., Sathian, K., Nichols-Larsen, D. S., Wolf, S. L., Good, D. C., & Wittenberg, G. F. (2008). Constraint-induced movement therapy results in increased motor map area in subjects 3 to 9 months after stroke. *Neurorehabilitation and Neural Repair*, *22*(5), 505–513. https://doi.org/10.1177/1545968308317531

Schiemanck, S. K., Kwakkel, G., Post, M. W. M., Kappelle, L. J., & Prevo, A. J. H. (2008). Impact of internal capsule lesions on outcome of motor hand function at one year post-stroke. *Journal of Rehabilitation Medicine*, *40*(2), 96–101. https://doi.org/10.2340/16501977-0130

Schmeidler, D. (1969). The nucleolus of a characteristic function game. *SIAM Journal on Applied Mathematics*, *17*(6), 1163–1170. https://doi.org/10.1137/0117107

Seitz, R. J., Höflich, P., Binkofski, F., Tellmann, L., Herzog, H., & Freund, H.-J. (1998). Role of the premotor cortex in recovery from middle cerebral artery infarction. *Archives of Neurology*, *55*(8), 1081. https://doi.org/10.1001/archneur.55.8.1081

Serrien, D. J., Ivry, R. B., & Swinnen, S. P. (2006). Dynamics of hemispheric specialization and integration in the context of motor control. *Nature Reviews. Neuroscience*, *7*(2), 160–166. https://doi.org/10.1038/nrn1849

Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, *39*(10), 1095–1100.

Shelton, F. N., & Reding, M. J. (2001). Effect of lesion location on upper limb motor recovery after stroke. *Stroke*, *32*(1), 107–112. https://doi.org/10.1161/01.str.32.1.107

Smith, D. V., Clithero, J. A., Rorden, C., & Karnath, H.-O. (2013). Decoding the anatomical network of spatial attention. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(4), 1518–1523. https://doi.org/10.1073/pnas.1210126110

Solomon, J., Raymont, V., Braun, A., Butman, J. A., & Grafman, J. (2007). User-friendly software for the analysis of brain lesions (ABLe). *Computer Methods and Programs in Biomedicine*, *86*(3), 245–254. https://doi.org/10.1016/j.cmpb.2007.02.006

Sommer, M. A. (2003). The role of the thalamus in motor control. *Current Opinion in Neurobiology*, *13*(6), 663–670. https://doi.org/10.1016/j.conb.2003.10.014

Sperber, C., & Karnath, H.-O. (2017). Impact of correction factors in human brain lesion-behavior inference: Validity of human brain lesion-behavior inference. *Human Brain Mapping*, *38*(3), 1692–1701. https://doi.org/10.1002/hbm.23490

Sperber, C., & Karnath, H.-O. (2018). On the validity of lesion-behaviour mapping methods. *Neuropsychologia*, *115*, 17–24. https://doi.org/10.1016/j.neuropsychologia.2017.07.035

Sperber, C., Wiesen, D., & Karnath, H.-O. (2019). An empirical evaluation of multivariate lesion behaviour mapping using support vector regression. *Human Brain Mapping*, *40*(5), 1381–1390. https://doi.org/10.1002/hbm.24476

Sprague, J. M., & Meikle, T. H. (1965). The role of the superior colliculus IN visually guided behavior. *Experimental Neurology*, *11*, 115–146. https://doi.org/10.1016/0014-4886(65)90026-9

Stinear, C. M., Barber, P. A., Petoe, M., Anwar, S., & Byblow, W. D. (2012). The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain*, *135*(8), 2527–2535. https://doi.org/10.1093/brain/aws146

Stinear, C. M., Barber, P. A., Smale, P. R., Coxon, J. P., Fleming, M. K., & Byblow, W. D. (2007). Functional potential in chronic stroke patients depends on corticospinal tract integrity. *Brain*, *130*(Pt 1), 170–180. https://doi.org/10.1093/brain/awl333

Tanji, J., & Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature*, *371*(6496), 413–416. https://doi.org/10.1038/371413a0

Toba, M. N., Godefroy, O., Rushmore, R. J., Zavaglia, M., Maatoug, R., Hilgetag, C. C., & Valero-Cabré, A. (2019). Revisiting 'brain modes' in a new computational era: Approaches for the characterization of brain-behavioural associations. *Brain*, *143*(4), 1088–1098. https://doi.org/10.1093/brain/awz343

Toba, M. N., Zavaglia, M., Malherbe, C., Moreau, T., Rastelli, F., Kaglik, A., Valabrègue, R., Pradat-Diehl, P., Hilgetag, C. C., & Valero-Cabré, A. (2020). Game theoretical mapping of white matter contributions to visuospatial attention in stroke patients with hemineglect. *Human Brain Mapping*, *41*(11), 2926–2950. https://doi.org/10.1002/hbm.24987

Toba, M. N., Zavaglia, M., Rastelli, F., Valabrégue, R., Pradat-Diehl, P., Valero-Cabré, A., & Hilgetag, C. C. (2017). Game theoretical mapping of causal interactions underlying visuo-spatial attention in the human brain based on stroke lesions. *Human Brain Mapping*, *38*(7), 3454–3471. https://doi.org/10.1002/hbm.23601

Tretriluxana, J., Gordon, J., Fisher, B. E., & Winstein, C. J. (2009). Hemisphere specific impairments in reach-to-grasp control after stroke: Effects of object size. *Neurorehabilitation and Neural Repair*, *23*(7), 679–691. https://doi.org/10.1177/1545968309332733

Tsurumi, M., Tanino, T., & Inuiguchi, M. (2001). A Shapley function on a class of cooperative fuzzy games. *European Journal of Operational Research*, *129*(3), 596–618. https://doi.org/10.1016/S0377-2217(99)00471-3

Tukey, J. W. (1980). We need both exploratory and confirmatory. *The American Statistician*, *34*(1), 23–25. https://doi.org/10.1080/00031305.1980.10482706

Turner, R. S., & Desmurget, M. (2010). Basal ganglia contributions to motor control: A vigorous tutor. *Current Opinion in Neurobiology*, *20*(6), 704–716. https://doi.org/10.1016/j.conb.2010.08.022

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. https://doi.org/10.1006/nimg.2001.0978

Vingerhoets, G. (2014). Contribution of the posterior parietal cortex in reaching, grasping, and using objects and tools. *Frontiers in Psychology*, *5*, 1–17. https://doi.org/10.3389/fpsyg.2014.00151

Ward, N. S., Brander, F., & Kelly, K. (2019). Intensive upper limb neurorehabilitation in chronic stroke: Outcomes from the queen square programme. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(5), 498–506. https://doi.org/10.1136/jnnp-2018-319954

Ward, N. S., Brown, M. M., Thompson, A. J., & Frackowiak, R. S. J. (2003). Neural correlates of motor recovery after stroke: A longitudinal fMRI study. *Brain*, *126*(Pt 11), 2476–2496. https://doi.org/10.1093/brain/awg245

Weiss, N. A. (2016). wBoot: Bootstrap methods. https://CRAN.R-project.org/package=wBoot

Wenzelburger, R., Kopper, F., Frenzel, A., Stolze, H., Klebe, S., Brossmann, A., Kuhtz-Buschbeck, J., Gölge, M., Illert, M., & Deuschl, G. (2005). Hand coordination following capsular stroke. *Brain*, *128*(Pt 1), 64–74. https://doi.org/10.1093/brain/awh317

Wise, S. P. (1985). The primate promoter cortex fifty years after Fulton. *Behavioural Brain Research*, *18*(2), 79–88. https://doi.org/10.1016/0166-4328(85)90064-6

Wu, O., Cloonan, L., Mocking, S. J. T., Bouts, M. J. R. J., Copen, W. A., Cougo-Pinto, P. T., Fitzpatrick, K., Kanakis, A., Schaefer, P. W., Rosand, J., Furie, K. L., & Rost, N. S. (2015). Role of acute lesion topography in initial ischemic stroke severity and long-term functional outcomes. *Stroke*, *46*(9), 2438–2444. https://doi.org/10.1161/STROKEAHA.115.009643

Xu, H., Wang, P., Ye, Z., Di, X., Xu, G., Mo, L., Lin, H., Rao, H., & Jin, H. (2016). The role of medial frontal cortex in action anticipation in professional badminton players. *Frontiers in Psychology*, *7*, 1–9. https://doi.org/10.3389/fpsyg.2016.01817

Xu, J., Ejaz, N., Hertler, B., Branscheidt, M., Widmer, M., Faria, A. V., Harran, M. D., Cortes, J. C., Kim, N., Celnik, P. A., Kitago, T., Luft, A. R., Krakauer, J. W., & Diedrichsen, J. (2017). Separable systems for recovery of finger strength and control after stroke. *Journal of Neurophysiology*, *118*(2), 1151–1163. https://doi.org/10.1152/jn.00123.2017

Xu, T., Jha, A., & Nachev, P. (2018). The dimensionalities of lesion-deficit mapping. *Neuropsychologia*, *115*, 134–141. https://doi.org/10.1016/j.neuropsychologia.2017.09.007

Yadav, G., Haaland, K. Y., & Mutha, P. K. (2019). Laterality of damage influences the relationship between impairment and arm use after stroke. *Journal of the International Neuropsychological Society*, *1–9*, 470–478. https://doi.org/10.1017/S1355617718001261

Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *14*(2), 65–72. https://doi.org/10.1007/BF01769885

Yourganov, G., Fridriksson, J., Rorden, C., Gleichgerrcht, E., & Bonilha, L. (2016). Multivariate connectome-based symptom mapping in Post-stroke patients: Networks supporting language and speech. *The Journal of Neuroscience*, *36*(25), 6668–6679. https://doi.org/10.1523/JNEUROSCI.4396-15.2016

Zavaglia, M., Forkert, N. D., Cheng, B., Gerloff, C., Thomalla, G., & Hilgetag, C. C. (2015). Mapping causal functional contributions derived from the clinical assessment of brain damage after stroke. *NeuroImage: Clinical*, *9*, 83–94. https://doi.org/10.1016/j.nicl.2015.07.009

Zavaglia, M., Forkert, N. D., Cheng, B., Gerloff, C., Thomalla, G., & Hilgetag, C. C. (2016). Technical considerations of a game-theoretical approach for lesion symptom mapping. *BMC Neuroscience*, *17*(1), 40. https://doi.org/10.1186/s12868-016-0275-6

Zavaglia, M., & Hilgetag, C. C. (2016). Causal functional contributions and interactions in the attention network of the brain: An objective multi-perturbation analysis. *Brain Structure & Function*, *221*(5), 2553–2568. https://doi.org/10.1007/s00429-015-1058-z

Zhang, S., Hu, S., Chao, H. H., & Li, C. R. (2017). Hemispheric lateralization of resting-state functional connectivity of the ventral striatum: An exploratory study. *Brain Structure and Function*, *222*(6), 2573–2583. https://doi.org/10.1007/s00429-016-1358-y

Zhang, Y., Kimberg, D. Y., Coslett, H. B., Schwartz, M. F., & Wang, Z. (2014). Multivariate lesion-symptom mapping using support vector regression. *Human Brain Mapping*, *35*(12), 5861–5876. https://doi.org/10.1002/hbm.22590

Zhu, L. L., Lindenberg, R., Alexander, M. P., & Schlaug, G. (2010). Lesion load of the corticospinal tract predicts motor impairment in chronic stroke. *Stroke*, *41*(5), 910–915. https://doi.org/10.1161/STROKEAHA.109.577023

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ofir-Geva, S., Meilijson, I., Frenkel-Toledo, S., & Soroker, N. (2023). Use of multi-perturbation Shapley analysis in lesion studies of functional networks: The case of upper limb paresis. *Human Brain Mapping*, *44*(4), 1320–1343. https://doi.org/10.1002/hbm.26105