Multiple Sclerosis and Related Disorders 4 (2015) 387–394

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

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

The analysis of semantic networks in multiple sclerosis identifies preferential damage of long-range connectivity

Elena Abad [a], Jorge Sepulcre [b], Elena H. Martinez-Lapiscina [c], Irati Zubizarreta [c], Jordi Garcia-Ojalvo [a], Pablo Villoslada [c,d,*]

[a] Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain
[b] Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Harvard Medical School, Boston, USA
[c] Center of Neuroimmunology, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) – Hospital Clinic of Barcelona, Barcelona, Spain
[d] Department of Neurology, University of California, San Francisco, USA

ARTICLE INFO

Article history:
Received 12 February 2015
Received in revised form
29 May 2015
Accepted 3 July 2015

Keywords:
Semantic networks
Multiple sclerosis
Brain networks
Fluency tests
Network analysis
Neuropsychology
Cognition

ABSTRACT

Objective: To analyze the characteristics of semantic networks derived from fluency tests in patients with multiple sclerosis (MS).
Methods: We built semantic networks by applying co-occurrence statistics to the data from verbal fluency tests performed on patients with MS ($n=36$) and matched controls ($n=200$), assessing the differences in network topology.
Results: As expected, the semantic networks from both patients and controls showed 'small-world' properties. Topological analysis of these semantic networks indicated that there were fewer nodes (words) and links (defined by significant co-occurrence of words) in those derived from MS patients. The average connectivity was not significantly affected, while the local connectivity (clustering coefficient) is preserved. Quantifiers of the cohesiveness of the network, which reflect long distance connectivity, such as assortativity and maximum centrality coefficients, differed significantly between MS patients and controls.
Conclusions: The analysis of semantic networks reveals quantitative differences in MS patients and identifies preferential damage of long-range connectivity. The analysis of semantic networks may be useful in clinical practice for the assessment of cognitive impairment or recovery after damage.

1. Introduction

Brain networks are damaged in neurological diseases, disturbing cognitive performance such as semantic memory (Sepulcre et al., 2012, 2009, 2008). Verbal fluency tests are used routinely to quantify global semantic memory output, and extracting semantic networks from these tests might provide additional information about the way the human brain stores and retrieves information when cognition is impaired (Lau et al., 2008; Steyvers and Tenenbaum, 2005). We previously developed a model of human semantic networks based on statistical co-occurrence in verbal fluency tasks using graph theory (Goni et al., 2011). In that model, the degree of co-occurrence between consecutive words (up to second neighbors) was measured in data from categorical fluency tests and when co-occurrence surpassed a given threshold, the two words (network nodes) were considered to be associated by a link (network edge). Using 200 fluency tests in the animal category from healthy individuals we obtained the semantic networks, which showed high modularity, displayed small-world architecture (Villoslada et al., 2009), and nicely reproduced classical definitions and measures of clustering and switching transitions (Troyer et al., 1997).

Previous analysis of verbal fluency tests in MS patients showed only a slight decrease in the total number of words, yet a consistent decrease in the switching score between subcategories was observed, as well as an increase in cluster size (Sepulcre et al., 2011; Joly et al., 2014). These observations suggest that a failure occurs in the retrieval of lexical information rather than a consistent reduction of the lexical pool. In addition to counting the number of words or their semantic similarity, one can also develop a representation of the connections between words using a statistical threshold (Goni et al., 2011), which counts how often words appear consecutively or close along the fluency tasks, representing semantic relations between words. This encouraged us to apply network analysis to semantic networks arising from this model in

* Correspondence to: IDIBAPS – Centre CELLEX, Casanova, 143 Planta 3A, 08036 Barcelona, Spain.
E-mail address: pvilloslada@clinic.ub.es (P. Villoslada).

http://dx.doi.org/10.1016/j.msard.2015.07.002
2211-0348/

order to explore the mechanisms underlying the alterations in semantic memory associated with MS.

Network analysis provides a tool for analyzing complex datasets with quantitative measurements that can be applied to the case of semantic output. In graph theory, a network is represented by a set of elements (vertices or nodes) that are connected in terms of specific relationships (links or edges). Studying the network properties allows for the identification of highly connected elements (so-called *hubs*), hierarchical organizations or critical nodes in the network cohesion. In addition, quantifying its size by counting the number of nodes and edges helps to further characterize the network (Barabasi et al., 2011). In order to analyze semantic networks, three different levels of analysis can be used: (i) the word level, focusing on specific nodes in the network, such as the most connected nodes (the largest nodes or *hubs*) and nodes of centrality that are related to the vulnerability of the network (cohesive nodes); (ii) the word group level, examining highly connected groups of nodes (neighborhoods in the network); and (iii) the word network level, corresponding to topological properties of the network as a whole (e.g. small-world phenomenon (Watts and Strogatz, 1998)). Moreover, as opposed to the classical definitions of lexical access strategies considered in verbal fluency tasks, we analyzed concepts like clustering and switching from the semantic network perspective.

The aim of this study was, using data from verbal fluency tests, to assess the semantic networks of MS patients and compare them with those of controls. The network structures formed from the statistical co-occurrence of concepts in the tests contain, albeit indirectly, information on the amount of content that is accessed during the test (number of words/nodes) as well as the routes or paths that are followed during content retrieval in the memory task of the fluency tests (links/edges).

2. Methods

2.1. Subjects

Patients with MS were recruited by their neurologist, providing their signed informed consent. Patients with MS were diagnosed according to the McDonald 2005 criteria (Polman et al., 2005). Demographic and clinical features of the cohorts are described in Table 1. Patients in an active phase of MS relapse, taking steroids, or who had suffered a clinical relapse within the previous three months were excluded from the study. None of the patients were suffering significant dysarthria, in line with previous reports suggesting that although most cases were mild, dysarthria was not rare (Piacentini et al., 2014). Given the well-known effect of psychiatric disorders on cognitive assessment (e.g., depression and anxiety), subjects with anxiety, depression or those taking psychoactive drugs did not participate in the study. Specifically, we ruled out patients taking recreational drugs (e.g. cocaine, amphetamines, LSD, and heroin) but not drugs prescribed for symptomatic therapy (e.g. amantadine for fatigue, baclofen for spasticity or carbamazepine for neuropathic pain). Because we excluded patients with on-going depression or anxiety, drugs used for these conditions were avoided (e.g. benzodiazepines, IRSS, etc.). The Cummings' Neuropsychiatric Inventory was used to confirm the absence of psychiatric disturbances (Cummings, 1997) (exclusion if any positive item was reported), in conjunction with Hamilton's Depression Rating Scale (Hamilton, 1960) (≥ 8 points) and Anxiety Rating Scale (Hamilton, 1959) (≥ 6 points).

Patients were assessed for cognitive impairment with the BRB-N test as previously described (Sepulcre et al., 2006), and patients with severe cognitive impairment (defined as >3SD in 3 subtests, which represents 30% of patients with MS (Sepulcre et al., 2006)) were excluded from the study. We did not assess levels of fatigue, a factor known to influence cognitive performance (Weinges-Evers et al., 2010; Andreasen et al., 2010; Flachenecker and Meissner, 2008), but tests were performed in the morning, when patients typically experience less fatigue. All patients had an MRI study in the previous 3 months before testing, including T1, T2 Flair and T1 post-gadolinium. Although MRI scans were not used to exclude patients due to the relationship between inflammatory activity and cognition (Bellmann-Strobl et al., 2009), no participants were experiencing significant disease activity at the time of the study as measured by new gadolinium enhancing lesions. Finally, only right-handed ($>70\%$ Oldfield scale (Oldfield, 1971)) native Spanish-speakers took part in the study. We also analyzed a cohort of sex, age and education matched control subjects. The Institutional Review Board at our center approved this study.

2.2. Verbal fluency tests

MS patients and controls were asked to name as many animals as possible in 90 s. All tests were performed by the same neuropsychologist and at the same time of the day (between 9 am and 1 pm). In these verbal fluency tests, proper names were not allowed and all the responses were transcribed verbatim. Repetitions and rule violations were not included when calculating the total verbal fluency scores. None of the individuals had previously been subjected to this test and none refused to perform it. Also, this test is less sensitive to longitudinal changes than other cognitive domains (Duque et al., 2008). No subjects suffered clinical reactivation of the disease during the neuropsychological assessment at baseline or at the endpoint of the study.

2.3. Building semantic networks and resampling methods

Semantic networks were built using a method based on co-occurrence statistics, as described previously (Goni et al., 2011). To address the difference in dataset sizes between the patient and control groups (number of individual tests), we performed statistical resampling using a jackknifing method (tests resampling). We initially dealt with the difference in size of each dataset: 36 tests from MS patients and 200 tests from controls. From the control dataset we evaluated the different combinations of 36 tests without repetitions, improving the final statistical power. Second, for each combination of 36 control and disease tests we computed distinct subsamples of 25 tests (without repetitions) through random permutations (at least 800 replicates). We built semantic networks by using a co-occurrence statistical model to link words (by edges in the graph) and we extracted the giant network component matrix (the maximum connected component) for each subsample of tests. Finally, we computed the different topological

Table 1
Demographics and clinical data of MS patients and controls. The data are expressed as the mean $\pm$ standard deviation, or the median and range or proportions, depending on the parametric or non-parametric distribution of the variable. The *p*-value is the mean of the Mann–Whitney *U*-test.

	Controls ($n=200$)	MS ($n=36$)	*p* Value
Sex female *n* [%]	117 [58%]	27 [75%]	ns
Age (year)[1]	32 ± 12	36 ± 8	ns
Education (year)[2]	15 (5–30)	15 (7–28)	ns
Verbal fluency test[1]	30 ± 6	25 ± 5	0.00004
EDSS	–	2.5 (0–7)	–
DMD (*Y*/*N*)	–	15/21	–

ns, No statistically significant differences between groups; –, not applicable.
EDSS: Expanded disability status Scale; DMD: Disease modifying drug.

[1] Mean $\pm$ standard deviation.
[2] Median or means, range in parentheses.