Can we accurately assess disease activity using automated methods in large real-life MRI databases?


To cite this version:

Arthur Masson, Benoît Combès, Romain Casey, François Cotton, Alice Dufey, et al.. Can we accurately assess disease activity using automated methods in large real-life MRI databases?: Insights from the OFSEP HD database.. MSMilan2023, Oct 2023, Milano (Italy), Italy. 2023. hal-04250265

HAL Id: hal-04250265
https://hal.science/hal-04250265
Submitted on 19 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Can we accurately assess MS disease activity using automated methods in large real-life MRI databases? Insights from the OFSEP HD database

A. Masson1, B. Combes1, R. Casey2, F. Cotton3, A. Dufey4, D. Laplaud5, E. Thouvenot6, E. Leray7,9, J. Epstein1, F. Guillemin7, S. Vukusic18

Background: The systematic collection of MRI scans associated with large real-life clinical databases (DB) is still rare but offers the opportunity to directly extract a wide range of imaging metrics through the use of AI-based tools.

Methods: We included 1412 pairs of brain MRI scans with both FLAIR images available in the French OFSEP HD multicentric imaging DB at two time points, and the radiological comparison captured in the OFSEP HD clinical DB. The data was heterogeneous (FLAIR images acquired on 35 different types of scanners from 5 manufacturers). An automated tool was used to detect new lesions. A ground truth for 160 pairs of MRIs was built from two experts and differences in sensitivity, specificity and accuracy between the two methods were assessed.

Results: Figure 1: Overall, 222 out of 1412 (16%) intervals were considered active from the clinical DB, compared to 467 (35%) from the automated method.
Figure 2: The clinical DB was more specific, but the automated method was more sensitive and accurate (p<0.001) to classify MS patients.
Figure 3: Under simplified assumptions, we extrapolate a sensitivity, specificity and accuracy of about 0.69, 0.99 and 0.92 for the clinical DB and 0.99, 0.95 and 0.96 for the automated method.

Conclusion: The automated analysis of images collected in large real-life databases increases the accuracy of MS patient classification as active or inactive compared to the clinical database and offers the possibility to extract other metrics such as lesion number or volume. These results should encourage us to develop the systematic collection and re-analysis of real-life imaging data linked to large clinical databases, in order, for example, to evaluate the effectiveness of treatments in real life.

The authors report no competing interests in relation to the work described.

Funding/support: Data collection has been supported by a grant provided by the French State and handled by the "Agence Nationale de la Recherche," within the framework of the "France 2030" programme, under the reference ANR-10-COHO-002 OFSEP; We thank the Eugène Devic EDMUS Foundation against multiple sclerosis for support.