



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

Arthur Masson, Benoit Combès, Romain Casey, François Cotton, Alice Dufey, David-axel Laplaud, Eric Thouvenot, Emmanuelle Leray, Jonathan Epstein, Francis Guillemin, et al.

► To cite this version:

Arthur Masson, Benoit 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. Masson^{*1}, B. Combès¹, R. Casey², F. Cotton^{3,4}, A. Dufey¹, D. Laplaud^{5,6}, E. Thouvenot^{7,8}, E. Leray^{9,10}, J. Epstein¹¹, F. Guillemin¹¹, S. Vukusic¹²,
¹³, A. Kerbrat^{1,14}

1 Inria centre at Rennes University, Rennes, France, 2 L'OFSEP, Bron, France, 3 Lyon Sud Hospital Center, Pierre-Bénite, France, 4 CREATIS, Villeurbanne, France, 5 Centre hospitalier universitaire de Nantes, Nantes, France, 6 Centre de Recherche en Transplantation et Immunologie ; Université de Nantes, Faculté de Médecine, INSERM UMR1064, Nantes, France, 7 University Hospital of Nîmes, Department of Neurology, Nîmes, France, 8 Institute Genomics Functional, Univ. Montpellier, UMR5203, INSERM 1191, Montpellier, France, 9 EhESP, Rennes, France, 10 Laboratoire Arènes CNRS UMR 6051, Rennes, France, 11 Hospital Center Regional And University De Nancy Hospital Central, Inserm CIC 1433 ClinicaEpidemiology, Nancy, France, 12 Pierre Wertheimer Hospital, Service de Neurologie, sclérose en plaques, pathologies de la myéline et neuro-inflammation, and Fondation Eugène Devic EDMUS contre la Sclérose en Plaques, Bron, France, 13 Claude Bernard University Lyon 1, Centre des Neurosciences de Lyon, INSERM 1028 et CNRS UMR5292, Villeurbanne, France, 14 CHU Rennes - Pontchaillou Hospital, Neurology department, Rennes, France

Inria

Inserm
La science pour la santé
From science to health

IRISA

Université
de Rennes

cnrs

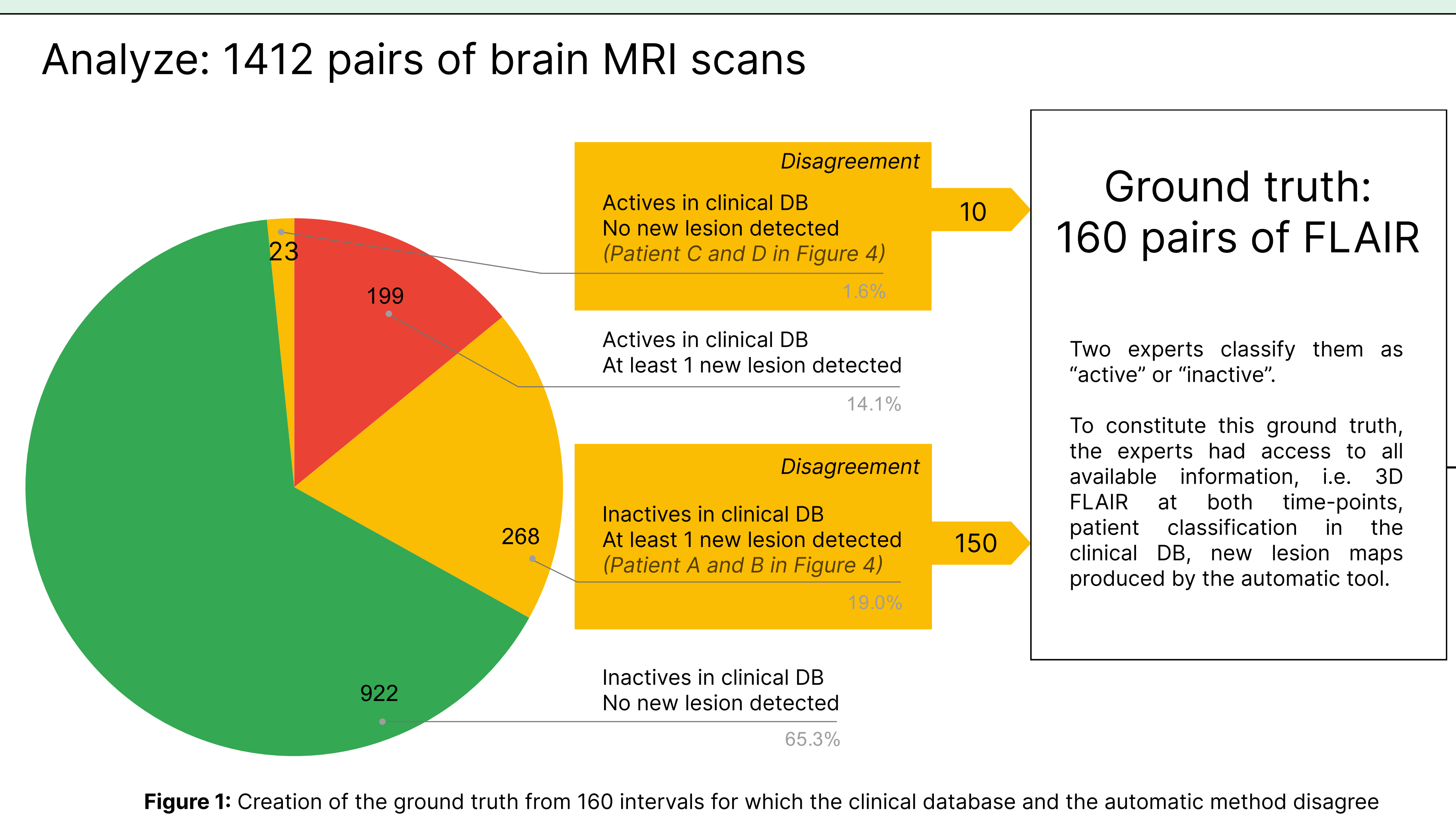
EMPENN

CHU
Rennes

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.

Objective: To compare the performance of an automatic tool that detects new lesions on FLAIR images vs radiological data entered in the clinical DB to classify MS patients as “active” or “inactive” in a large multicenter real-life DB.

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.

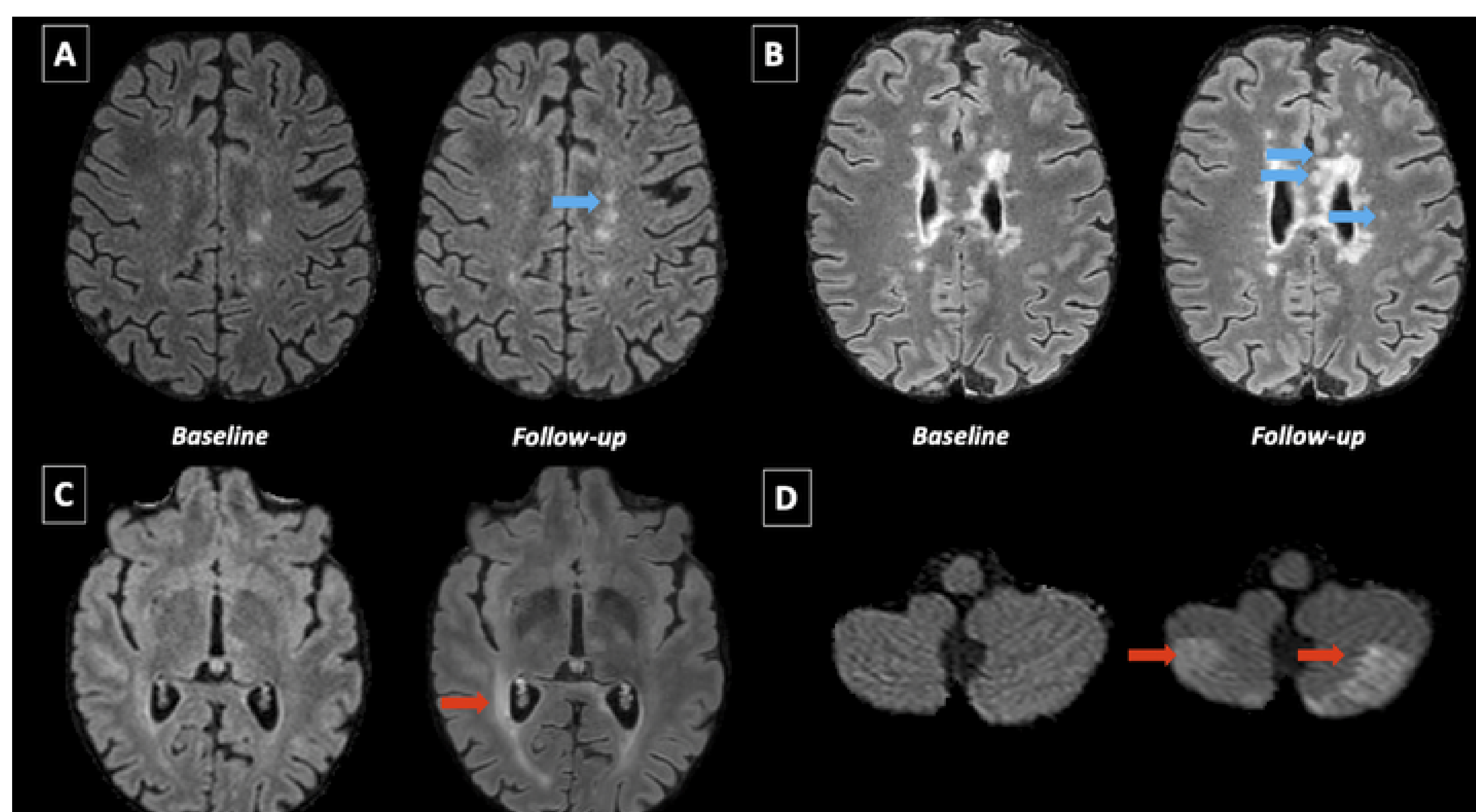
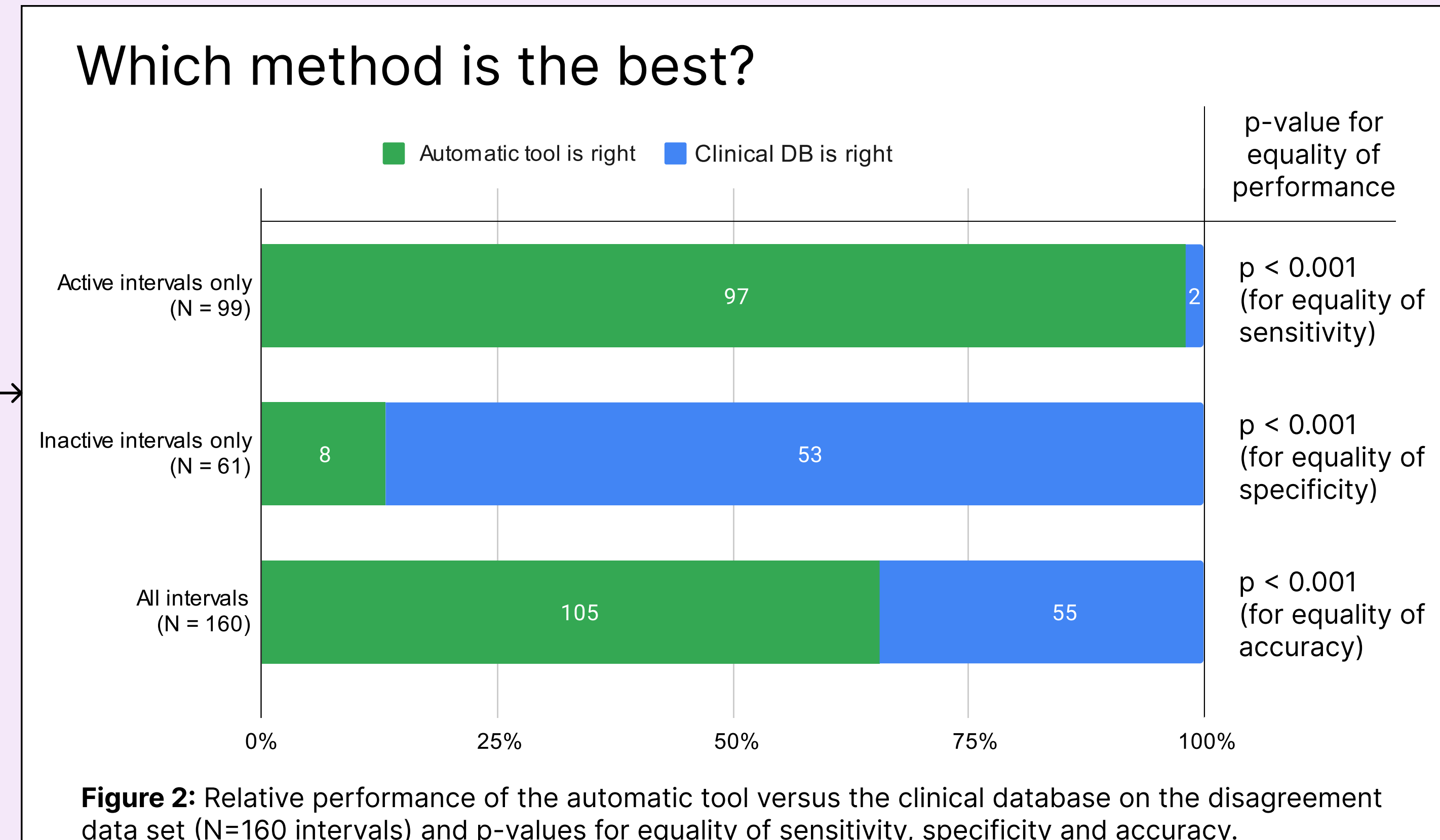


Figure 4: Examples of false negatives generated by the clinical database (A, B) and false positives generated by the automatic tool (C, D). A: patient with 1 new lesion detected by the automatic tool and classified as inactive in the database; B: patient with 21 new lesions detected by the automatic tool and classified as inactive in the database; C: patient correctly classified as inactive in the database but where changes in image characteristics between two acquisitions leads to false detection of a new lesion by the automatic tool; D: patient correctly classified as inactive in the database but where image artifacts leads to false detection of two new lesions by the automatic tool.

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.

What is the sensitivity, specificity and accuracy of the two methods ?

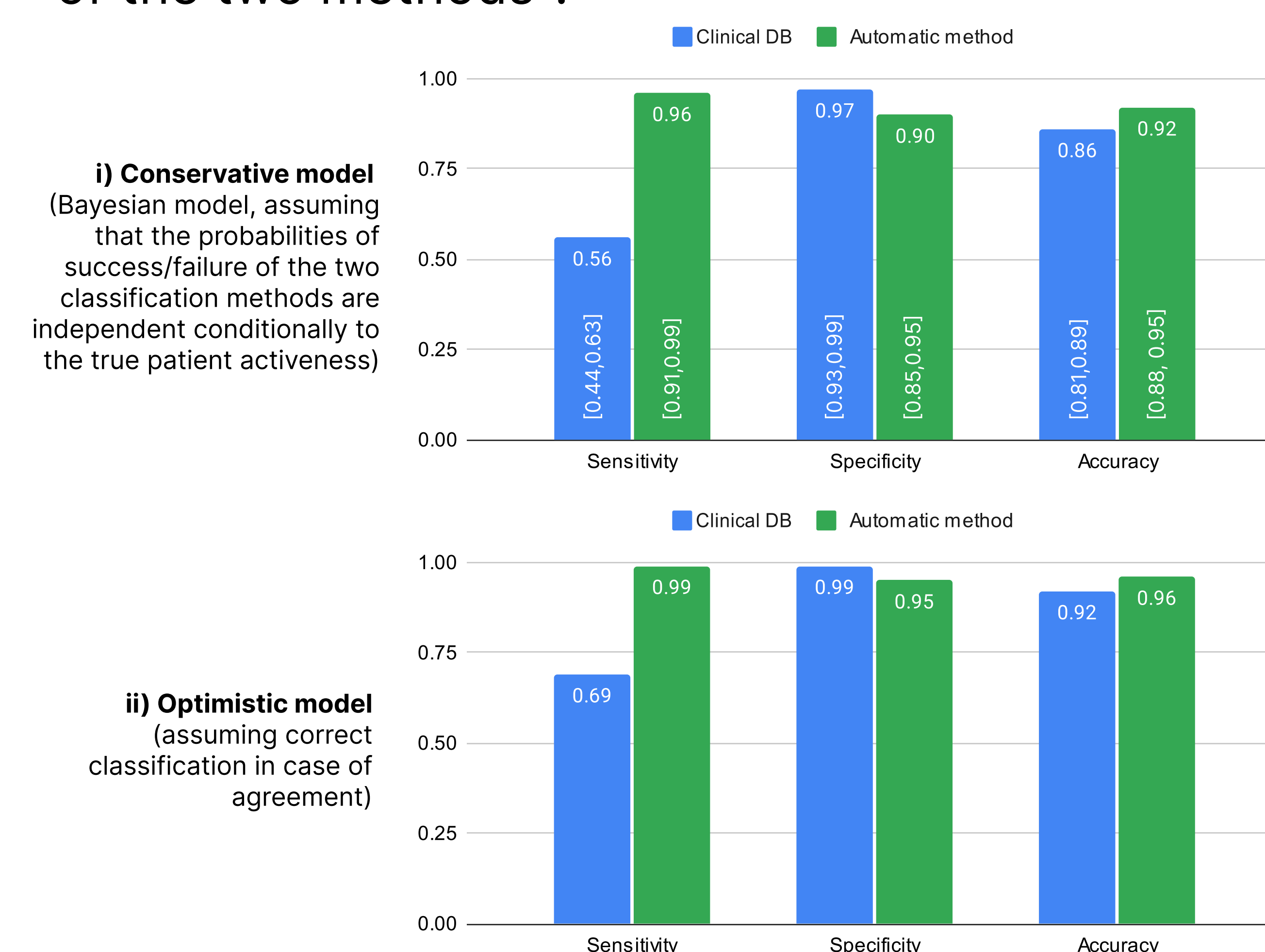


Figure 3: Specificity, sensibility and accuracy as computed i) by using the Bayesian Model assuming that the outcomes of the two classification methods (the clinical database and the automatic tool) are independent conditionally to the actual patient activeness and ii) by considering that when the two methods agreed, the patient was correctly classified.

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.