Forecast of the MMSE score up to 6 years ahead, with cross-cohort replications

Etienne Maheux, Igor Koval, Damiano Archetti, Alberto Redolfi, Stanley Durrleman

Forecast of the MMSE score up to 6 years ahead, with cross-cohort replications

E. Maheux[a1], I. Koval[a], D. Archetti[b], A. Redolfi[b] and S. Durrleman[a]

[a] *Inria, Aramis project-team, Institut du Cerveau et de la Moelle épinière (ICM), Hôpital de la Pitié Salpêtrière, Inserm U 1127, CNRS UMR 7225, Sorbonne Université, F-75013, Paris, France;* [b]*Laboratory of Alzheimer's Neuroimaging and Epidemiology (LANE), IRCCS Instituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy*

Keywords: cognitive decline forecast, personalized prediction, longitudinal study, cross-cohort replication

1. Introduction

Most of the studies that focus on the Alzheimer's disease (AD) prediction forecast a dementia label: cognitively normal, mild cognitive impairments or AD. However, these labels cannot always be compared across cohorts, they are not perfectly accurate and they describe the disease thanks to a limited number of stages (only 3) contrary to what is, in practice, a continuous progression. For these reasons, we predict the Mini-Mental State Examination (MMSE), a cognitive test whose scale presents 30 discrete values. Predicting this measure, assessed in multiple AD-related cohorts, allows better staging of the disease progression at the individual level, in particular in the ADNI, AIBL and PharmaCOG cohorts.

2. Methods

2.1 Model calibration

The Bayesian mixed-effects model introduced in [1] allows to estimate the average temporal progression of biomarkers out of individual longitudinal data, i.e. repeated measurements over time. This group-average scenario is *calibrated* on the ADAS, MMSE, MOCA, RAVLT, Boston Naming Test, CDR-SoB, Geriatric Depression Score, Category Fluency (Animals) assessments of the ADNI database, letting aside the stable cognitively normal (CN) subjects. Selected individuals are thus CN converters, mild cognitive impaired (MCI) subjects with stable diagnosis, MCI converting to AD and stable AD. The group-average scenario describes the continuous conversion of each biomarker from a normal to an abnormal state in the form of a logistic curve, reconstructed from individual measurements with potentially missing values [2]. In a cross-validation setting (10 splits of 70/30% between train and test), it is possible to hide future values of test patients to forecast them (1.5, 3.0, 4.5 and 6.0 years after the last seen visit). The prediction is compared to the real value that has been hidden during the *personalization* of the model to the individual data.

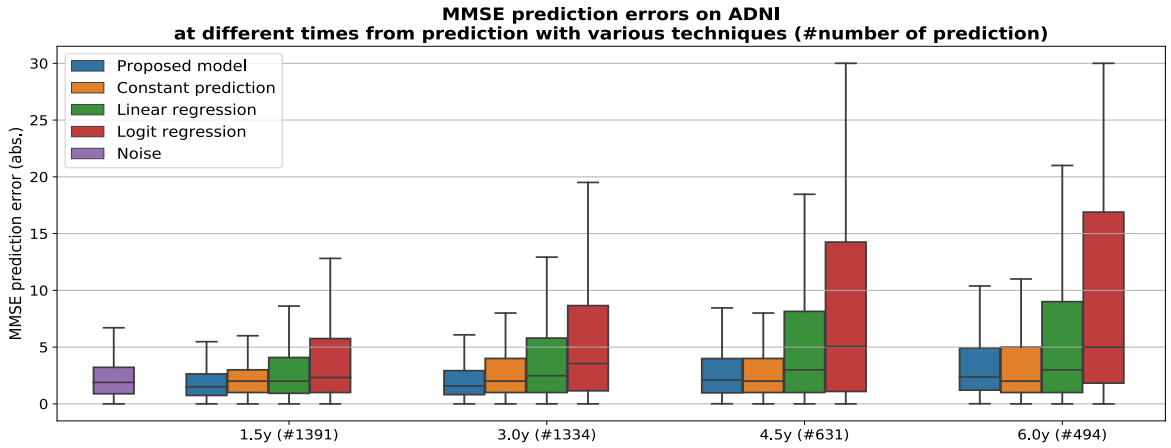


Figure 1: MMSE forecast absolute errors (AE) with respect to time to prediction (years after the last known visit). Boxplots whiskers correspond to the 5th and 95th percentile. Outliers have been removed for the sake of clarity. The *(# number)* indicates the number of predictions at each temporal horizon.

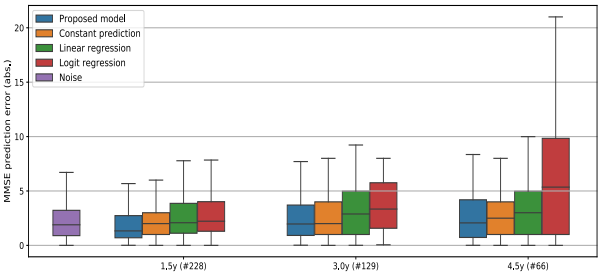
[1] Corresponding author. Email: etienne.maheux@inria.fr

2.2 Model comparisons

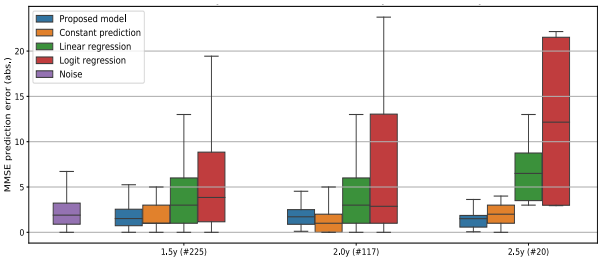
Our forecast errors are compared to those obtained with (i) constant predictions, i.e. no change in MMSE score, (ii) individual linear regressions of patients' past scores with respect to their past ages, and (iii) individual logit regressions. Besides, they are compared to the best achievable forecast which is the noise in the data. It has been measured in [3] – thanks to test-retests assessments including inter and intra-rater variability – leading to a standard deviation of 2.8 out of 30.

2.3 Cohort replications

Although the previous group-average scenario of biomarkers progression has been calibrated on 8 cognitive tests, it is possible to personalize it to patients that only present a subset of these assessments. Thanks to this, we personalized the model to patients from the AIBL and PharmaCOG cohorts so to forecast their MMSE scores. Due to cohorts variability (especially the number of follow-up visits available / the time patients were followed), we evaluate forecasts at different temporal horizons.



(a) MMSE prediction errors on the AIBL cohort 1.5, 3.0 and 4.5 years ahead.



(b) MMSE prediction errors on the PharmaCOG cohort 1.5, 2.0 and 2.5 years ahead.

Figure 2: Replication of the MMSE forecasts on patients from AIBL and PharmaCOG cohorts – the model having been fitted on ADNI subjects only.

3. Results and discussion

Figure 1 shows our forecasts on the ADNI cohort. While the linear and logit regressions present poor predictive power, the constant prediction, even if trivial, is somehow accurate (especially at short temporal horizons) due to the fact that the disease settles over long periods of time. Nevertheless, at any time horizon, our method predicts MMSE scores as well as or better than any of these techniques, and often with a lower dispersion. Besides, for some conditions, distributions of absolute errors compare with the intrinsic level of noise in the data. This comparison stresses the fact that short-term forecasts are worthless as the constant prediction already achieves noise level prediction. Furthermore, our model does not overfit and correctly estimates the long-term progression of the biomarkers as its personalization to patients of the AIBL and PharmaCOG cohorts shows prediction of the noise level or below constant prediction, outperforming the linear and logit regressions [Figure 2].

4. Conclusions

The constant prediction outperforms standard regression techniques and is of noise level for small temporal horizons (up to 2 or 3 years). Such comparison should always be undertaken in a forecasting context to evaluate a minimal temporal horizon for worthy predictions.

Additionally, our model has proven to accurately estimate the average scenario of biomarkers progression. The latter, once personalized to individual data, enables to accurately forecast the MMSE for a wide range of times to prediction, even for patients from unseen cohorts (potentially having fewer cognitive assessments). This paves the way to accurate and robust prediction in real-life applications.

5. References

[1] J.-B. Schiratti, et al. A Bayesian mixed-effects model to learn trajectories of changes from repeated manifold-valued observations. *Journal of Machine Learning Research*, 18(133):1-33. 2017.

[2] R. Couronné, et al. Learning disease progression models with longitudinal data and missing values. *ISBI*. 1033-1037. 2019.

[3] C.M. Clark, et al. Variability in annual MMSE score in patients with probable Alzheimer's disease in late onset families. *Archives of Neurology.* 56(7):857-862. 1999.