frontiers

Model-free predictions in personalized medicine with quantified uncertainty and personalized decisions: A case study on the conversion from MCI to AD

P.G.L. Porta Mana [1,2,*], **Ingrid Rye** [3], **Alexandra Vik** [1,2], **Marek Kocinski** [2,4], **Arvid Lundervold** [2,4], **Astri J. Lundervold** [3], **and Alexander S. Lundervold** [1,2]

[1] *Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Bergen, Norway*
[2] *Mohn Medical Imaging and Visualization Centre (MMIV), Department of Radiology, Haukeland University Hospital, Bergen, Norway*
[3] *Department of Biological and Medical Psychology, University of Bergen, Norway*
[4] *Department of Biomedicine, University of Bergen, Norway*

Correspondence*:
Corresponding Author
pgl@portamana.org

ABSTRACT

[Luca & Astri, draft]

Patients with Mild Cognitive Impairment have an increased risk of Alzheimer's disease. Early identification of underlying neurodegenerative processes is essential to provide treatment before the disease is well established in the brain. It is therefore of extreme interest to study how well several kinds of predictors – from neuropsychological examinations to advanced brain-imaging techniques – allow us to prognose the onset of the disease.

But more is needed for a personalized approach to prognosis, prevention, treatment, than just the obvious requirement that prognoses be as best as they can be for each patient. Several situational elements that can be different from patient to patient must be accounted for:

- the *kinds* of clinical data and evidence available for prognosis;
- the *outcomes* of the same kind of clinical data and evidence;
- the kinds of treatment or prevention strategies available, owing to different attitudes toward life, different family networks and possibilities of familial support, different additional medical factors such as physical disabilities, different economic means;
- the advantages and disadvantages, gains and costs of the same kinds of treatment or prevention strategies; the patient has a major role in the quantification of such gains and costs;
- finally, the initial evaluation by the clinician – which often relies on too subtle clues (family history, regional history, previous case experience) to be considered as measurable data.

Statistical decision theory is the normative quantification framework that takes into account these fundamental differences. Medicine has the distinction of having been one of the first fields to adopt this framework, exemplified in brilliant old and new textbooks on clinical decision making.

Clinical decision making makes allowance for these differences among patients through two requirements. First, the quantification of prognostic evidence on one side, and of gains and costs of treatments and prevention strategies on the other, must be clearly separated and handled in a modular way. Two patients can have the same prognostic evidence, and yet very different prevention options. Second, the quantification of independent prognostic evidence ought to be in the form of *likelihoods about the health condition* (or equivalently of likelihood ratios, in a binary case), that is, of the probabilities of the observed test outcomes given the hypothesized health conditions. Likelihoods from independent clinical tests and predictors can then be combined with a simple multiplication; for one patient we could have three kinds of predictor available; for another, we could have five. The clinician's pre-test assessment is included in the form of a probability. These patient-dependent probabilities are combined with the patient-dependent costs and gains of treatment or prevention, to arrive at the best course of action for that patient. The main result underlying statistical decision theory is that decision making *must* take this particular mathematical form in order to be optimal and logically consistent.

The present work investigates the prognostic power of a set of neuropsychological and Magnetic Resonance Imaging examinations, demographic data, and genetic information about APOE status, for the prediction of the onset of Alzheimer's disease in patients defined as mildly cognitively impaired at a baseline examination. The longitudinal data used come from the ADNI database.

The prognostic power of these predictors is quantified in the form of a combined likelihood for the onset of Alzheimer's disease. As a hypothetical example application of personalized clinical decision making, three patient cases are considered where a clinician starts with prognostic uncertainties, possibly coming from other tests, of 50%/50%, 25%/75%, 75%/25%. It is shown how these pre-test probabilities are changed by the predictors.

This quantification also allows us to rank the relative prognostic power of the predictors. It is found that several neuropsychological examinations have highest prognostic power, much higher than the genetic and an imaging-derived predictors included in the present set.

Several additional advantages of this quantification framework are also exemplified and discussed in the present work:

- missing data are automatically handled, and results having partial data are not discarded; this quantification therefore also accounts for patient-dependent availability of *non-independent* predictors;
- no modelling assumptions (e.g. linearity, gaussianity, functional dependence) are made;
- the prognostic power obtained is intrinsic to the predictors, that is, it is a bound for *any* prognostic algorithm;
- variability ranges of the results owing to the finite size of the sample data is automatically quantified.
- the values obtained, being probabilities, are more easily interpretable than scores of various kinds.

Maybe we can also add two examples of different clinical decisions, coming from different gain/cost evaluations?

For full guidelines regarding your manuscript please refer to Author Guidelines.

As a primary goal, the abstract should render the general significance and conceptual advance of the work clearly accessible to a broad readership. References should not be cited in the abstract. Leave the Abstract empty if your article does not require one, please see Summary Table for details according to article type.

Keywords: keyword, keyword, keyword, keyword, keyword, keyword, keyword, keyword

1 INTRODUCTION

[Luca, pieces of text]

Personalized diagnosis, prognosis, treatment, and prevention strategies must make allowance for several fundamental differences among patients:

- the *kinds* of clinical data and evidence available for diagnosis or prognosis can be different;
- the *values* of the same kind of clinical data and evidence can be different;
- the kinds of treatment or prevention options can be different;
- the advantages and disadvantages, gains and costs of the same kinds of treatment or prevention can be different;
- finally, the evaluation of the clinician – which often relies on too subtle clues (family history, regional history, case experience) to be considered as measurable data – can be different.

Is there really a methodological framework that can take all these differences into account? Yes, there is, and Medicine has the distinction of having been one of the first fields to adopt it (Ledley and Lusted, 1959): Statistical Decision Theory. Its application in Medicine is explained and exemplified in several, brilliant, old and new textbooks (Weinstein and Fineberg, 1980; Sox et al., 2013; Hunink et al., 2014). This theory has mathematical and logical foundations and its principles constitute indeed the foundations for the definition and realization of Artificial Intelligence (Russell and Norvig, 2022)

The basics of clinical decision making ..basics: each piece of evidence contributes with a likelihood or odds; they combine together and together with the clinician's pre-data evaluation. Then they are combined with the different gains/costs of treatments or prevention strategies to find the optimal one. Decision trees can be necessary (but don't change this framework). Costs & gains are evaluated by clinician & patient together.

$$\overbrace{\mathrm{p}(\text{health condition} \mid \text{results of all tests},\ \text{prior info})}^{\textit{post-test probability}} \propto$$

$$\overbrace{\mathrm{p}(\text{health condition} \mid \text{prior info})}^{\textit{pre-test probability by clinician}} \times$$

$$\textit{likelihoods of tests} \begin{cases} \mathrm{p}(\text{result of 1st test} \mid \text{health condition},\ \text{prior info}) \times \\ \mathrm{p}(\text{result of 2nd test} \mid \text{health condition},\ \text{prior info}) \times \\ \ldots \end{cases} \quad (1)$$

For Original Research Articles (Name et al., 1996), Clinical Trial Articles (LastName1 et al., 2013), and Technology Reports (Surname1, 2010), the introduction should be succinct, with no subheadings (Name, 1993). For Case Reports the Introduction should include symptoms at presentation (Surname, 2002), physical exams and lab results (LastName1 et al., 2011).

2 ARTICLE TYPES

For requirements for a specific article type please refer to the Article Types on any Frontiers journal page. Please also refer to Author Guidelines for further information on how to organize your manuscript in the required sections or their equivalents for your field

3 MANUSCRIPT FORMATTING

3.1 Heading Levels

3.2 Level 2

3.2.1 Level 3

3.2.1.1 Level 4

3.2.1.1.1 Level 5

3.3 Equations

Equations should be inserted in editable format from the equation editor.

$$\sum x + y = Z \quad (2)$$

3.4 Figures

Frontiers requires figures to be submitted individually, in the same order as they are referred to in the manuscript. Figures will then be automatically embedded at the bottom of the submitted manuscript. Kindly ensure that each table and figure is mentioned in the text and in numerical order. Figures must be of sufficient resolution for publication see here for examples and minimum requirements. Figures which are

not according to the guidelines will cause substantial delay during the production process. Please see here for full figure guidelines. Cite figures with subfigures as figure 2a and 2b.

3.4.1 Permission to Reuse and Copyright

Figures, tables, and images will be published under a Creative Commons CC-BY licence and permission must be obtained for use of copyrighted material from other sources (including re-published/adapted/modified/partial figures and images from the internet). It is the responsibility of the authors to acquire the licenses, to follow any citation instructions requested by third-party rights holders, and cover any supplementary charges.

3.5 Tables

Tables should be inserted at the end of the manuscript. Please build your table directly in LaTeX.Tables provided as jpeg/tiff files will not be accepted. Please note that very large tables (covering several pages) cannot be included in the final PDF for reasons of space. These tables will be published as Supplementary Material on the online article page at the time of acceptance. The author will be notified during the typesetting of the final article if this is the case.

4 NOMENCLATURE

4.1 Resource Identification Initiative

To take part in the Resource Identification Initiative, please use the corresponding catalog number and RRID in your current manuscript. For more information about the project and for steps on how to search for an RRID, please click here.

4.2 Life Science Identifiers

Life Science Identifiers (LSIDs) for ZOOBANK registered names or nomenclatural acts should be listed in the manuscript before the keywords. For more information on LSIDs please see Inclusion of Zoological Nomenclature section of the guidelines.

5 ADDITIONAL REQUIREMENTS

For additional requirements for specific article types and further information please refer to Author Guidelines.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

The Author Contributions section is mandatory for all articles, including articles by sole authors. If an appropriate statement is not provided on submission, a standard one will be inserted during the production process. The Author Contributions statement must describe the contributions of individual authors referred

to by their initials and, in doing so, all authors agree to be accountable for the content of the work. Please see here for full authorship criteria.

FUNDING

Details of all funding sources should be provided, including grant numbers if applicable. Please ensure to add all necessary funding information, as after publication this is no longer possible.

ACKNOWLEDGMENTS

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

SUPPLEMENTAL DATA

Supplementary Material should be uploaded separately on submission, if there are Supplementary Figures, please include the caption in the same file as the figure. LaTeX Supplementary Material templates can be found in the Frontiers LaTeX folder.

DATA AVAILABILITY STATEMENT

The datasets [GENERATED/ANALYZED] for this study can be found in the [NAME OF REPOSITORY] [LINK].

REFERENCES

Hunink, M. G. M., Weinstein, M. C., Wittenberg, E., Drummond, M. F., Pliskin, J. S., Wong, J. B., et al. (2014). *Decision Making in Health and Medicine: Integrating Evidence and Values* (Cambridge: Cambridge University Press), 2 edn. doi:10.1017/CBO9781139506779. First publ. 2001

[Dataset] LastName1, A., LastName2, A., and LastName3, A. (2011). Data title. doi:10.000/55555

LastName1, A., LastName2, A., and LastName3, A. (2013). Article title. *Frontiers in Neuroscience* 30, 10127–10134. doi:10.3389/fnins.2013.12345

Ledley, R. S. and Lusted, L. B. (1959). Reasoning foundations of medical diagnosis: Symbolic logic, probability, and value theory aid our understanding of how physicians reason. *Science* 130, 9—21. doi:10.1126/science.130.3366.9

Name, A. (1993). *The title of the work* (The city: The name of the publisher)

Name, C., Surname, D., and LastName, F. (1996). The title of the work. In *The title of the conference proceedings*, eds. E. Name1 and E. Name2 (The name of the publisher), 41–50

Russell, S. J. and Norvig, P. (2022). *Artificial Intelligence: A Modern Approach*. Pearson series in artificial intelligence (Harlow, UK: Pearson), fourth global ed. edn. `http://aima.cs.berkeley.edu/global-index.html`, `https://archive.org/details/artificial-intelligence-a-modern-approach-4th-edition`. First publ. 1995

Sox, H. C., Higgins, M. C., and Owens, D. K. (2013). *Medical Decision Making* (New York: Wiley), 2 edn. doi:10.1002/9781118341544. First publ. 1988

Surname, B. (2002). The title of the work. In *The title of the book*, ed. E. Name (The city: The name of the publisher). 201–213

Surname1, H. (2010). *The title of the work* (Patent country: Patent number)
Weinstein, M. C. and Fineberg, H. V. (1980). *Clinical Decision Analysis* (Philadelphia: Saunders)

FIGURE CAPTIONS

Figure 1. Enter the caption for your figure here. Repeat as necessary for each of your figures

Figure 2a. This is Subfigure 1.

Figure 2b. This is Subfigure 2.

Figure 2. Enter the caption for your subfigure here. **(A)** This is the caption for Subfigure 1. **(B)** This is the caption for Subfigure 2.