frontiers

Personalized prognosis & decision: An example study on the conversion from Mild Cognitive Impairment to Alzheimer's Disease

P.G.L. Porta Mana [1,2,*], **I. Rye** [3], **A. Vik** [1,2], **M. Kociński** [2,4], **A. Lundervold** [2,4], **A. J. Lundervold** [3], **A. 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*:
P.G.L Porta Mana, HVL, Inndalsveien 28, 5063 Bergen
pgl@portamana.org

ABSTRACT

Patients with Mild Cognitive Impairment have an increased risk of a trajectory toward Alzheimer's Disease. Early identification of patients with a high risk of Alzheimer's Disease is essential to provide treatment before the disease is well-established in the brain. great importance to study how well different kinds of predictors allow us to prognose a trajectory from Mild Cognitive Impairment towards Alzheimer's Disease in an individual patient.

But more is needed for a personalized approach to prognosis, prevention, and 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 additional medical factors such as physical disabilities, different attitudes toward life, different family networks and possibilities of familial support, different economic means;
- the advantages and disadvantages, benefits and costs of the same kinds of treatment or prevention strategies; the patient has a major role in the quantification of such benefits 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 benefits 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 benefits 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 Apolipoprotein-E4The present work investigates the prognostic power of a set of neuropsychological and Magnetic Resonance Imaging examinations, demographic data, and genetic information about Apolipoprotein-E4 (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.

(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. update this

rewrite following This quantification also allows us to rank the relative prognostic power of the predictors. It is found that several neuropsychological examinations have the highest prognostic power, much higher than the genetic and 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 are automatically quantified.
- the values obtained, being probabilities, are more easily interpretable than scores of various kinds.

Keywords: Clinical decision making, Utility theory, Probability theory, Artificial Intelligence, Machine Learning, Base-rate fallacy

1 EACH PATIENT IS UNIQUE

Meet Olivia, Ariel, Bianca, Curtis.[1] These four persons don't know each other, but they have something in common: they all suffer from a mild form of cognitive impairment, and are afraid that their impairment will turn into Alzheimer's Disease within a couple of years. In fact, this is why they recently underwent some clinical analyses and cognitive tests. Today they received the results of their analyses. From these results, available clinical statistical data, and other relevant information, their clinician will assess their risk of developing Alzheimer. The clinician and each patient will then decide among a set of possible preventive treatments.

Besides this shared condition and worry, these patients have other things in common – but also some differences. Let's take Olivia as reference and list the similarities and difference between her and the other three:

- Olivia and Ariel turn out to have exactly identical clinical results and age. They would also get similar benefits from the available preventive-treatment options. Ariel, however, comes from a different geographical region from Olivia, where the conversion rate from Mild Cognitive Impairment to Alzheimer is lower: about 30%, whereas in Olivia's it is around 45% (Petersen et al., 1999) Moreover there is no history of Alzheimer's Disease in Ariel's family.
- Olivia and Bianca also have exactly the same clinical results and age. They come from the same geographical region and have very similar family histories. In fact we shall see that they have the same probability of developing Alzheimer's disease. Bianca, however, suffers from several allergies and additional clinical conditions that would render some of the preventive options less beneficial to her.
- Olivia and Curtis have different clinical results – in particular, Olivia has the risky Apolipoprotein-E4 (APOE4) allele (Liu et al., 2013) whereas Curtis hasn't – and age, Olivia being more than 20 years older than Curtis. But they otherwise come from the same geographical region, have very similar family histories, and would get similar benefits from the preventive options.

We can categorize these differences as "difference in auxiliary information" (Olivia and Ariel), "difference in preventive benefits" (Olivia and Bianca), "difference in clinical factors" (Olivia and Curtis). Figure 1 summarizes the similarity and differences between Olivia and the other three patients. Table 1 reports the clinical results and demographic data common to Olivia, Ariel, Bianca, as well as those of Curtis need to explain the variates and refer to (Rye et al., 2022).

Considering the similarities and differences among these patients, which treatments are optimal and should prescribed to them?

[1] Fictive characters; any reference to real persons is purely coincidental

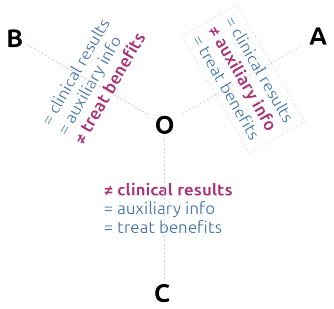

Figure 1. draft, needs better font sizes

Patient	Age	Sex	HC $\cdot 10^{-3}$	APOE4	ANART	CFT	GDS	RAVLT-im	RAVLT-del	RAVLT-rec	TMTA	TMTB
Olivia, Ariel, Bianca	83.5	F	2.77	Y	14	18	0	29	12	3	126.5	117.0
Curtis	61.5	M	4.17	N	6	13	1	16	2	0	24.8	83.4

Table 1. Clinical results & demographic data

Our main purpose in the present work is to illustrate, using the four fictitious patients above as example, how this clinical decision-making problem can today be solved methodically, exactly, and at low computational cost, when the available prognostic clinical information involves one-dimensional or categorical variates such as those listed in table 1. The solution method takes into account and integrates available clinical statistical data and each patient's unique combination of clinical results, auxiliary information, and preventive benefits. It is therefore the staple method for personalized prognosis, diagnosis, treatment.

In our example we shall find that despite the many factors in common among our four patients – even despite the identical clinical results for Olivia, Ariel, Bianca, and despite the identical probability of conversion for Olivia and Bianca – *the optimal treatment option for each patient is different from those for the other three*. This result exemplifies the importance of differences among patients with regard to clinical results, auxiliary information, or preventive benefits.

The method used is none other than decision theory, the combination of probability theory and utility theory (von Neumann and Morgenstern, 1955; Raiffa and Schlaifer, 2000; Raiffa, 1970; Berger, 1985; Jaynes, 2003). Medicine has the distinction of having been one of the first fields to adopt it (Ledley and Lusted, 1959), with old and new brilliant textbooks (Weinstein and Fineberg, 1980; Sox et al., 2013; Hunink et al., 2014) that explain and exemplify its application.

Decision theory is also the foundation for the construction of an Artificial Intelligence agent capable of rational inference and decision making (Russell and Norvig, 2022, ch. IV; Jaynes, 2003, chs 1–2). The present method can therefore be seen as the application of an *ideal machine-learning algorithm*. "Ideal" in the sense of being free from approximations, special modelling assumptions, and limitations in its informational output; not in the sense of being impracticable. Another important point of the present work, indeed, is to show that *for some kinds of dataset* such ideal machine-learning algorithm is a reality. It is preferable to popular algorithms such as neural networks, random forests, support-vector machines, which are unsuited to clinical decision-making problems owing to their output limitations. We discuss this matter further in § ***.

Add other goals, results, and some synopsis

The inferential and decision-making steps are summarized in table 2.

Table 2. Inferential and decision-making steps. Steps in **boldface** represent patient-dependent, personalized steps that cannot be obtained from the learning dataset

1. Infer the full-population frequencies of predictors and predictand, using available datasets.
2. **Assess in respect of which variates the present patient can be considered as belonging to the same population underlying the learning dataset.**
3. **Assess the prior probability of the predictand for the present patient.** This step allows us (a) to consider additional clinical information outside of the dataset's variates, and available for the present patient only; (b) to correct for mismatches between the dataset's underlying population and the patient's one.
4. Calculate the *likelihood* of the predictand for a specific patient, given the patient's predictor values. Combine this likelihood with the prior from step 3, to obtain the final probability for the predictand, for the present patient.
5. **Assess the clinical courses of action available for the present patient, together their benefits and costs.** This step is fundamentally patient-dependent and is the one open to most variability from patient to patient.
6. Choose the course of action having maximal expected benefit for the present patient, given the benefits assessed in step 5 and the final probability assessed in step 4.

2 PRACTICAL EXAMPLE

It may be optimal to present the steps in reverse order: the last one explains the goal, and makes clear why the preceding steps are necessary.

2.1 Learning

The learning stage is essential to arrive at the probability that a patient might convert from Mild Cognitive Impairment to Alzheimer's Disease.

The basic idea is to consider the patient one member in a large population of similar patients. If we knew the joint frequencies of all possible combinations of predictor and predictand values in such population – and knew nothing else – then we would judge the probability for our patient to have particular values to be equal to the population frequency. Pure symmetry considerations lead to this intuitive result (de Finetti, 1930; Dawid, 2013; Bernardo and Smith, 2000, §§ 4.2–4.3). This conclusion also holds for *conditional* frequencies and probabilities, that is, probabilities for a particular variate value given that the value of some other variate values are known.

If the full statistics of such a population were known, our task would just be to "enumerate" rather than to "learn". Learning comes into play because the full population is not known: we only have a sample from it.

The most we can do is therefore to assign a probability to each possible frequency distribution for the full population. Intuitively, the probability of a candidate frequency distribution is determined by two factors: (a) how well it fits the sample data, (b) how biologically or physically reasonable it is. Figure 2 show a fictitious sample data and various candidate frequency distributions ....

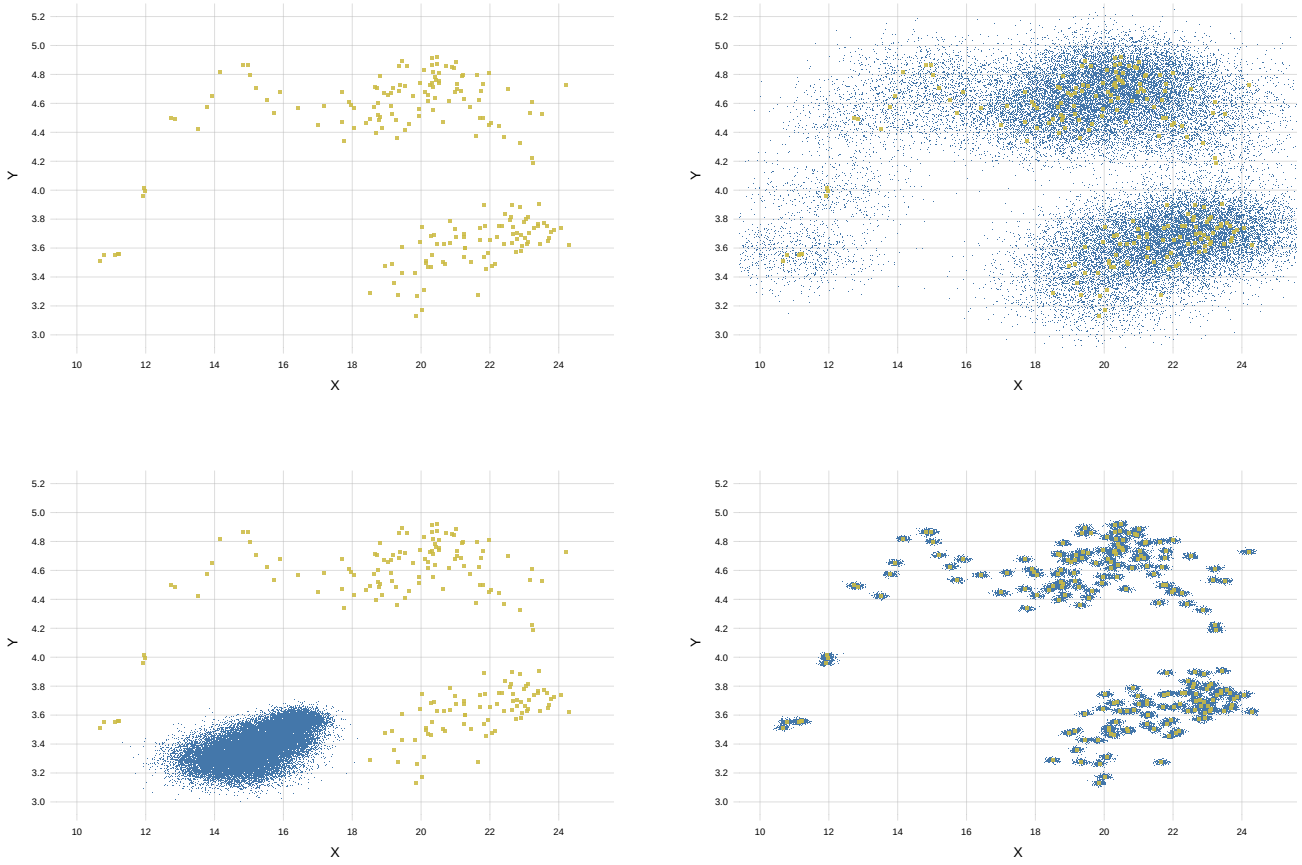

Figure 2. Upper-left: Sample data. Upper-right: candidate frequency distribution that fits the data and does not look unnatural. Lower-left: candidate distribution that might look natural but doesn't fit the sample data. Lower-right: candidate distribution that fits the data very well but looks unnatural.

Some more intuition and details about the maths, principles, and relation to machine learning (Dunson and Bhattacharya, 2011; Rossi, 2014; Rasmussen, 1999).

2.2 Population assessment

2.3 Prior probability

2.4 Likelihood and posterior probability

2.5 Benefit assessments

2.6 Maximization of expected benefit

— Luca, old 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, benefits 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 benefits/costs of treatments or prevention strategies to find the optimal one. Decision trees can be necessary (but don't change this framework). Costs & benefits are evaluated by clinician & patient together.

$$\overbrace{\mathrm{p}(\text{health condition} \mid \text{results of all tests},\ \text{prior info})}^{\textit{post-test probability}} \propto \overbrace{\mathrm{p}(\text{health condition} \mid \text{prior info})}^{\textit{pre-test probability by clinician}} \times \textit{likelihoods of tests} \begin{cases} \mathrm{p}(\text{result of 1st test} \mid \text{health condition},\ \text{prior info}) \times \\ \mathrm{p}(\text{result of 2nd test} \mid \text{health condition},\ \text{prior info}) \times \\ \dots \end{cases} \quad (1)$$

3 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

4 MANUSCRIPT FORMATTING

4.1 Heading Levels

4.2 Level 2

4.2.1 Level 3

4.2.1.1 Level 4

4.2.1.1.1 Level 5

4.3 Equations

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

$$\sum x + y = Z \tag{2}$$

4.4 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 **??** and **??**.

4.4.1 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.

4.5 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.

5 NOMENCLATURE

5.1 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.

5.2 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.

6 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 authors were too immersed in the development of the present work to keep a detailed record of who did what.

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

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer series in statistics (New York: Springer), 2 edn. DOI:10.1007/978-1-4757-4286-2. First publ. 1980

Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A. F. M., et al. (eds.) (2011). *Bayesian Statistics 9* (Oxford: Oxford University Press). DOI:10.1093/acprof:oso/9780199694587.001.0001

Bernardo, J.-M. and Smith, A. F. (2000). *Bayesian Theory*. Wiley series in probability and mathematical statistics (New York: Wiley), repr. edn. DOI:10.1002/9780470316870. First publ. 1994

Cifarelli, D. M. and Regazzini, E. (1979). Considerazioni generali sull'impostazione bayesiana di problemi non parametrici. Le medie associative nel contesto del processo aleatorio di Dirichlet. *Riv. mat. sci. econ. soc.* 2, 39–52, 95–111

Damien, P., Dellaportas, P., Polson, N. G., and Stephens, D. A. (eds.) (2013). *Bayesian Theory and Applications* (Oxford: Oxford University Press). DOI:10.1093/acprof:oso/9780199695607.001.0001

Dawid, A. P. (2013). Exchangeability and its ramifications. In (Damien et al., 2013), chap. ch. 2. 19–29. DOI:10.1093/acprof:oso/9780199695607.003.0002

de Finetti, B. (1929). Funzione caratteristica di un fenomeno aleatorio. In *Atti del Congresso Internazionale dei Matematici:*, ed. S. Pincherle (Bologna: Zanichelli), vol. 6. 179–190. https://www.mathunion.org/icm/proceedings, http://www.brunodefinetti.it/Opere.htm. Transl. in (Cifarelli and Regazzini, 1979). See also (de Finetti, 1930)

de Finetti, B. (1930). Funzione caratteristica di un fenomeno aleatorio. *Atti Accad. Lincei: Sc. Fis. Mat. Nat.* IV, 86–133. http://www.brunodefinetti.it/Opere.htm. Summary in (de Finetti, 1929)

Dunson, D. B. and Bhattacharya, A. (2011). Nonparametric Bayes regression and classification through mixtures of product kernels. In (Bernardo et al., 2011). 145–158. DOI:10.1093/acprof:oso/9780199694587.003.0005, older version at https://www.researchgate.net/publication/228447342_Nonparametric_Bayes_Regression_and_Classification_Through_Mixtures_of_Product_Kernels

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

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press). Ed. by G. Larry Bretthorst. First publ. 1994. DOI:10.1017/CBO9780511790423, https://archive.org/details/XQUHIUXHIQUHIQXUIHX2, http://www-biba.inrialpes.fr/Jaynes/prob.html

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

Liu, C.-C., Kanekiyo, T., Xu, H., and Bu, G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat. Rev. Neurol.* 9, 106–118, 184. DOI:10.1038/nrneurol.2012.263, DOI:10.1038/nrneurol.2013.32

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Arch. Neurol.* 56, 303–308, 760. DOI:10.1001/archneur.56.3.303, https://jamanetwork.com/journals/jamaneurology/fullarticle/775068

Raiffa, H. (1970). *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Behavioral science: quantitative methods (Reading, USA: Addison-Wesley), 2nd pr. edn. First publ. 1968

Raiffa, H. and Schlaifer, R. (2000). *Applied Statistical Decision Theory*. Wiley Classics Library (New York: Wiley), repr. edn. First publ. 1961

Rasmussen, C. E. (1999). The infinite Gaussian mixture model. *Adv. Neural Inf. Process. Syst. (NIPS)* 12, 554–560. https://www.seas.harvard.edu/courses/cs281/papers/rasmussen-1999a.pdf

Rossi, P. E. (2014). *Bayesian Non- and Semi-parametric Methods and Applications*. The Econometric and Tinbergen Institutes lectures (Princeton: Princeton University Press). DOI:10.1515/9781400850303

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

Rye, I., Vik, A., Kocinski, M., Lundervold, A. S., and Lundervold, A. J. (2022). Predicting conversion to Alzheimer's disease in individuals with Mild Cognitive Impairment using clinically transferable features. *Sci. Rep.* 12, 15566. DOI:10.1038/s41598-022-18805-5

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

von Neumann, J. and Morgenstern, O. (1955). *Theory of Games and Economic Behavior* (Princeton: Princeton University Press), 3rd ed., 6th pr. edn. https://archive.org/details/in.ernet.dli.2015.215284. First publ. 1944

Weinstein, M. C. and Fineberg, H. V. (1980). *Clinical Decision Analysis* (Philadelphia: Saunders)