

Design of an Interactive Web Application for Teaching Uncertainty Interpretations of Clinical Tests

Sebastian NEHER^{a,1}, Lorenz A. KAPSNER^{b,c}, Hans-Ulrich PROKOSCH^{a,b}, and
Dennis TODDENROTH^a

^a*Chair of Medical Informatics, Friedrich-Alexander-University Erlangen-Nürnberg
(FAU), Erlangen, Germany*

^b*Medical Center for Information and Communication Technology,
Universitätsklinikum Erlangen, Erlangen, Germany*

^c*Department of Radiology, Universitätsklinikum Erlangen, Friedrich-Alexander-
Universität Erlangen-Nürnberg (FAU), Erlangen, Germany*

Abstract. Background: Assessing the uncertainty of diagnostic findings is essential for advising patients. Previous research has demonstrated the difficulty of computing the expected correctness of positive or negative results, although clinical decision support (CDS) tools promise to facilitate adequate interpretations. Objectives: To teach the potential utility of CDS tools to medical students, we designed an interactive software module that computes and visualizes relevant probabilities from typical inputs. Methods: We reviewed the literature on recommended graphical approaches and decided to support contingency tables, plain table formats, tree diagrams, and icon arrays. Results: We implemented these functions in a single-page web application, which was configured to complement our local learning management system where students also access interpretation tasks. Conclusion: Our technical choices promoted a rapid implementation. We intend to explore the utility of the tool during some upcoming courses. Future developments could also model a more complex clinical reality where the likelihood of alternative diagnoses is estimated from sets of clinical investigations.

Keywords. Clinical Decision Support, Bayes Theorem, Medical Education

1. Introduction

The digitization of medical practice promises to facilitate various processes of clinical care with computerized inferences from patient data. Such clinical decision support (CDS) functions have been evaluated for a range of specialized applications. The development and assessment of CDS tools falls into the sphere of medical informatics, which also instructs future physicians about the benefits and limitations of current technologies [1-3].

Among the typical tasks of medical practice, the accurate appraisal and communication of diagnostic findings emerges in many subdisciplines and settings.

¹ Corresponding Author: Sebastian Neher, Chair of Medical Informatics, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany, Email: Sebastian.neher@fau.de

Physicians of course deploy diagnostic investigations across a spectrum of certitude about the present patient state. For instance, secondary prevention such as cancer screening may be scheduled irrespective of any signs or complaints, while confirmatory examinations may aim to settle some suspected diagnosis after other forms of testing. Accumulating patient information thus incrementally modifies the likelihood that the presence of a disease can be either assumed or excluded. The growing availability of quantitative information about different diagnostic methods implies that these probabilities can increasingly be calculated on an individual basis and thereby integrated into advising patients about treatment options.

The known discriminative capabilities of many clinical tests are reported as pairs of sensitivity and specificity, which denote expectations that truly affected or unaffected patients receive correct results. These essential test characteristics are typically estimated by calculating the proportions or relative frequencies among earlier patients for whom the corresponding “ground truth” is assumed to be known. From these parameters as well as the respective prevalence or pre-test estimate, the mutual probability that some positive or negative result is actually correct can be calculated via the Bayes theorem. These post-test probabilities, which are also termed positive predictive value (PPV) and negative predictive value (NPV), are fundamental to reasonable clinical interpretations as well as to informed recommendations concerning further therapeutic considerations.

Potential consequences of overestimating the probability and overdiagnosis of a present illness may include patients suffering from ‘psychological distress’ and the premature initiation of interventions with avoidable side effects [4]. Underrating the likelihood of a disease, on the other hand, might entail that physicians miss opportunities for effective treatment as progressive conditions such as infections or malignancies develop unchecked [5-8].

Previous research has demonstrated how difficult the involved calculations based on conditional probabilities can be without technical instruments [9, 10]. Psychological experiments have particularly suggested that plain numbers and formulas appeared intricate for many people, including for students and professionals [8, 11-14]. There are indications, however, that diagrams that structure and display relevant information could be helpful for an adequate appraisal [15-17].

The gradual computerization of medical education has a particular relevance for this specific topic, insofar as digital instruments may suitably constitute the medium as well as the content of teaching CDS methods to medical students. Online learning platforms, videos, mobile and browser-based applications showed great potential for instructing students about evidence-based medicine, as well as for fostering statistical knowledge and risk competence [5].

The described relevance and difficulty of appraising clinical tests motivated our design and implementation of a CDS tool for the statistical calculation and visualization of post-test probabilities. Below we describe functional requirements and technical features of an interactive web application that we developed for teaching uncertainty interpretations of diagnostic findings to medical students.

2. Methods

To define the functionality of our CDS tool for digitally assisted uncertainty assessments of clinical findings, we conducted a non-systematic literature review to identify relevant visualization approaches. We thus pragmatically queried PubMed with the search terms

“(“Bayes Theorem”[Mesh]) AND (natural frequencies) AND teaching”, ((natural frequencies) AND (“Bayes Theorem”[Mesh])) AND (“Data Interpretation, Statistical”[Mesh]) and (“Students, Medical”[Mesh]) AND (“Bayes Theorem”[Mesh]). Results were screened by reading their abstracts, and the search was incrementally refined by also including relevant keywords used by these initial results, articles that PubMed displays as ‘similar’, as well as articles that were cited by publications from the provisional result set. After assembling relevant literature, we evaluated the practicability of implementing the described graphical approaches in an interactive software by estimating the involved effort. We selected four graphical methods that we estimated to be realizable by a medium-experienced programmer in less than a week each, and then implemented these in a web application for teaching statistical test appraisal as well as the use of CDS functions to medical students.

Steckelberg et al. reported that the *contingency table* is the most prominent visualization of the Bayes Theorem and related calculations in medical science. Due to the general utility of the table format and its wide application for data visualization in the literature, extracting information from this format is considered indispensable for any student and future researcher [15, 17, 18]. The *tree diagram* is another type of visualization that has been recommended for illustrating calculations based on the Bayes rule, which is widely used in psychology and other decision-affiliated fields [13, 16-18]. *Bar graphs* excel at displaying quantitative information by highlighting the size relations between frequency data and are present in practically every field and discipline [16]. *Icon arrays* or *pictographs* composed of human-shaped head or body symbols are particularly deployed to emphasize medical information as they illustrate the relative dimensions of different populations [11, 15].

While graphics promote an intuitive but approximate understanding of the quantitative relations between different frequencies, explicit numbers convey these parameters with greater precision, so the application should automatically compute and display all relevant statistical quantities in numerical form as well.

To compute and visualize these post-test parameters for different clinical tests, users of course need the possibility to enter different numerical inputs, particularly sensitivity, specificity, and prevalence.

Since we planned to provide the corresponding learning unit to students asynchronously, we required a widely applicable solution that we could distribute to common operating systems, including to mobile systems such as tablets or smartphones.

3. Results

To implement these functional requirements, we designed our educational CDS tool as an interactive single-page web application that students could access from various systems. This facilitated a complementary configuration with the Ilias-based online learning management system that is used at our local university for the creation and electronic distribution of instructive presentations, videos, surveys and quizzes for monitoring learning progress. This allowed students to navigate to basic tutorial documents, to a quiz module containing basic Bayesian interpretation tasks and some multiple-choice items, and to our developed application. Figure 1 illustrates this technical setup.

The CDS tool stores no individual data on the server, but since the Ilias platform provides logins that we could use to identify candidates, we can tie student responses to motivating feedback and scores depending on the adequacy of answers.

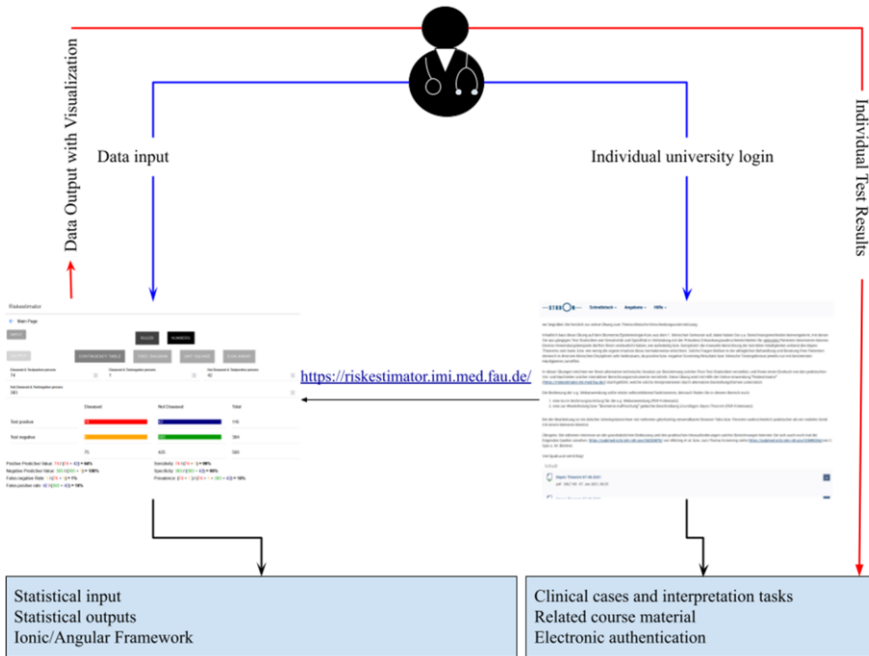


Figure 1. Architecture of developed web application vs established learning management system

With the assistance of our CDS application, students may work on Bayesian interpretation tasks and compute individual post-test probabilities for hypothetical clinical cases.

We structured the graphical user interface of our application into three horizontal visual sections, which respectively represent inputs, graphical outputs, and numerical outputs. On the top of the page users can switch between alternative data input and output options, while the actual data entry is positioned below. The middle of the page is reserved for the chosen graphical visualization, and the bottom of the page displays a tabular structure with different statistical values that are updated in real time when input parameters change.

As these Bayesian interpretations consider two hypotheses (disease or no disease) and a surrogate parameter with two values (test positive or negative), we distinguish four patient subgroups. To ensure visual consistency and easy interpretations across graphical and numerical outputs, we assigned every patient subgroup a distinct color, using red and orange for affected patients with either correct or incorrect findings. Assumed unaffected cases were rendered green and blue, depending on test results. In the entire application, this color scheme was implemented with *Cascading Style Sheets* definitions.

To facilitate data entry on conventional desktop computers as well as on handheld devices, we provided users with the alternatives of either typing numerical inputs on keyboards or accessing graphical control elements. We thus configured widgets so that relative frequencies such as sensitivity and prevalence can be entered via graphical sliders, intuitively specified as percentages between 0 and 100. The graphical slider allows quick changes of values, including on touchscreen displays, while providing an immediate visual feedback about the entered values. The alternative data input method involves entering absolute frequencies, which are automatically restricted to numbers between 0 and 1,000 to ensure intuitively comprehensible notional patient populations.

Among the alternative output formats, the *contingency table* was implemented as a typical 2x2-table in a plain layout. Changes of input parameters automatically modify the corresponding post-test estimates and are highlighted according to the described color scheme. The *tree diagram* with automatically updating frequency labels was realized by dynamically embedding *Scalable Vector Graphics* (SVG) elements in the source hypertext. Numerical outputs are likewise depicted in their respective color. The *bar chart* output was also implemented with SVG elements, with an automatically updating geometry and consistently colored subareas. The *icon array* is similarly built with a dynamic number of copies of referenced SVG images in the corresponding color. Figure 2 provides an overview of these visualization formats.

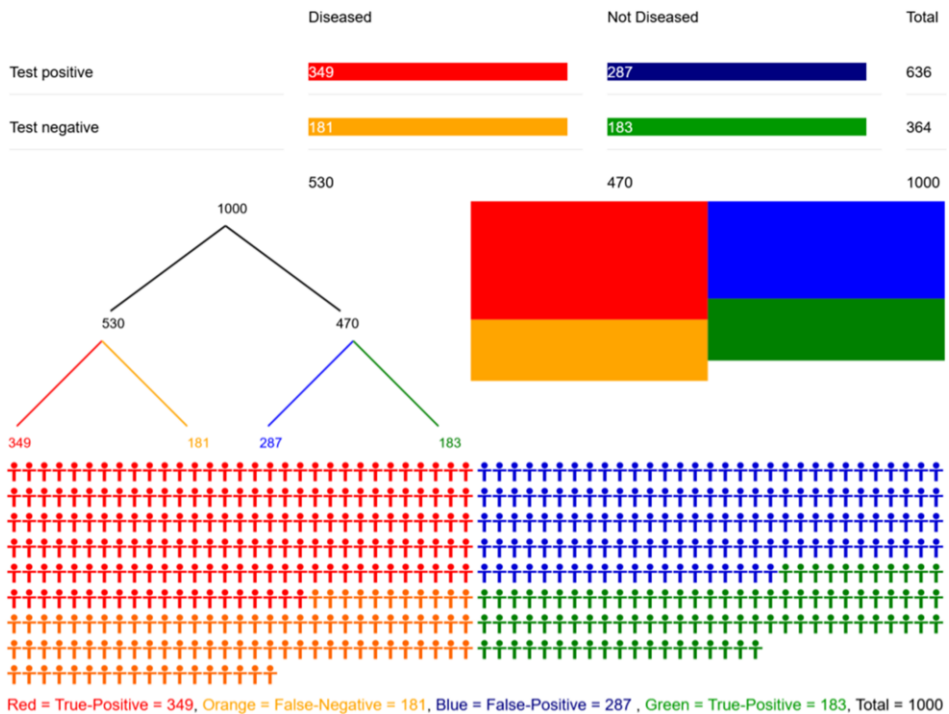


Figure 2. Alternative output formats from top to bottom: contingency table, tree diagram, bar graphs, icon array

Post-test probabilities and other statistical outputs are displayed at the bottom of the page and are automatically recalculated and updated whenever inputs change. To illustrate the computations, the respective formulas are also shown with self-updating terms.

Considering the prior described functionality by the Ilias Platform we only had to implement the computational and display logic. We created the computational and display logic with the Typescript-based Ionic/Angular framework. A German version of the application can be accessed at <https://riskestimator.imi.med.fau.de/>. The source code of the computational and display logic is available at <https://github.com/nehers/Riskestimator.git>. Ionic applications support an uncomplicated distribution to commonly used operating systems by tailoring hypertext elements to target systems. This leads to slightly varying optics of interface elements such as sliders or buttons on different operating systems. This yields two upsides. Programmers benefit from being able to support software varieties optimized for different target systems from a single code base, while users get their accustomed looks and familiar handling of interactive elements.

4. Discussion

Our described technical choices appeared to promote a rapid implementation. With respect to the realized functionality, it has to be noted that users who are completely unfamiliar with the Bayes theorem or the involved statistical calculations could be overwhelmed by the number of possibilities for input and output. But for a presumed majority of students in advanced semesters, we believe our created proof-of-concept tool seems useful for simplifying an important aspect that many medical students as well as physicians see as rather complex and inaccessible. The contingency table as well as the tree diagram are often utilized in teaching the Bayes theorem, so we assume that many students have worked with these visualizations, and when given a choice might prefer them. As a next step, we intend to explore the acceptance and utility of our tool among medical students during upcoming informatics courses that make use of it. There are several more promising visualization methods that need further investigation and evaluation in a digital environment; a notable addition could be the Fagan nomogram, which allows the calculation of post-test probabilities with the help of likelihood ratios and pre-test estimates [19], or the extended ruler proposed by Herich, Lehmacher and Hellmich, which allows the calculation with sensitivity, specificity and prevalence [20].

While our proposed CDS tool currently supports Bayesian calculations that handle statistical parameters from a single clinical test, future developments could also reflect a more complex clinical reality by computing the probabilities of alternatively considered differential diagnoses that originate from sets or sequences of conducted investigations. In addition to feasible interfaces of such CDS functions to clinically deployed electronic medical records, as well as to their potential educational value, the field thus appears to offer exciting opportunities for future studies.

Declarations

Conflict of interest: All authors declare that they have no conflict of interest.

Author Contributions: All authors participated in the design of the described web application. SN and DT reviewed the literature. SN and LK implemented the described web application. SN wrote the manuscript. DT revised the manuscript. LK and HUP revised intermediate manuscript versions. All authors read and approved the final version.

The present work was performed in (partial) fulfillment of the requirements for obtaining the degree „Dr. med.”

References

- [1] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005; 330(7494):765. doi: 10.1136/bmj.38398.500764.8F.
- [2] Greenes RA, Bates DW, Kawamoto K, Middleton B, Osheroff J, Shahar Y. Clinical decision support models and frameworks: Seeking to address research issues underlying implementation successes and failures. *J Biomed Inform* 2018; 78:134–43. doi: 10.1016/j.jbi.2017.12.005.
- [3] Varghese J, Kleime M, Gessner SI, Sandmann S, Dugas M. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review. *J Am Med Inform Assoc* 2018; 25(5):593–602. doi: 10.1093/jamia/ocx100.
- [4] Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev*. 2013;(6):CD001877. doi:10.1002/14651858.CD001877.pub5 Cited in: PubMed; PMID 23737396.
- [5] Kyriakoulis K, Patelarou A, Laliotis A, Wan AC, Matalliotakis M, Tsiou C et al. Educational strategies for teaching evidence-based practice to undergraduate health students: systematic review. *J Educ Eval Health Prof* 2016; 13:34. doi: 10.3352/jeehp.2016.13.34.
- [6] Wegwarth O, Gigerenzer G. The Barrier to Informed Choice in Cancer Screening: Statistical Illiteracy in Physicians and Patients. *Recent Results Cancer Res* 2018; 210:207–21. doi: 10.1007/978-3-319-64310-6_13.
- [7] Bergus G, Vogelgesang S, Tansey J, Franklin E, Feld R. Appraising and applying evidence about a diagnostic test during a performance-based assessment. *BMC Med Educ* 2004; 4:20. doi: 10.1186/1472-6920-4-20.
- [8] Whiting PF, Davenport C, Jameson C, Burke M, Sterne JAC, Hyde C et al. How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open* 2015; 5(7):e008155. doi: 10.1136/bmjopen-2015-008155.
- [9] Weber P, Binder K, Krauss S. Why Can Only 24% Solve Bayesian Reasoning Problems in Natural Frequencies: Frequency Phobia in Spite of Probability Blindness. *Front Psychol* 2018; 9:1833. doi: 10.3389/fpsyg.2018.01833.
- [10] Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. *Acad Med* 1998; 73(5):538–40. doi: 10.1097/00001888-199805000-00024.
- [11] Brase GL. Pictorial representations in statistical reasoning. *Appl. Cognit. Psychol.* 2009; 23(3):369–81. doi: 10.1002/acp.1460.
- [12] Johnson KV: The two by two diagram; a graphical truth table. *J Clin Epidemiol* 1999, 52:1073-1082.
- [13] Binder K, Krauss S, Bruckmaier G, Marienhagen J. Visualizing the Bayesian 2-test case: The effect of tree diagrams on medical decision making. *PLoS One* 2018;
- [14] Brase GL. The power of representation and interpretation: Doubling statistical reasoning performance with icons and frequentist interpretations of ambiguous numbers. *Journal of Cognitive Psychology*. 2014;26(1):81–97. doi:10.1080/20445911.2013.861840.
- [15] Böcherer-Linder K, Eichler A. How to Improve Performance in Bayesian Inference Tasks: A Comparison of Five Visualizations. *Front Psychol* 2019; 10:267. doi: 10.3389/fpsyg.2019.00267.
- [16] Böcherer-Linder K, Eichler A. The Impact of Visualizing Nested Sets. An Empirical Study on Tree Diagrams and Unit Squares. *Front Psychol* 2016; 7:2026. doi: 10.3389/fpsyg.2016.02026.
- [17] Binder K, Krauss S, Bruckmaier G. Effects of visualizing statistical information - an empirical study on tree diagrams and 2×2 tables. *Front Psychol* 2015; 6:1186. doi: 10.3389/fpsyg.2015.01186.
- [18] Steckelberg A, Balgenorth A, Berger J, Mühlhauser I. Explaining computation of predictive values: 2×2 table versus frequency tree. A randomized controlled trial ISRCTN74278823. *BMC Med Educ* 2004; 4:13. doi: 10.1186/1472-6920-4-13.
- [19] Fagan TJ. Letter: Nomogram for Bayes theorem. *N Engl J Med*. 1975;293(5):257. doi:10.1056/NEJM197507312930513 Cited in: PubMed; PMID 1143310.
- [20] Herich L, Lehmacher W, Hellmich M. Drop the Likelihood Ratio. A Novel Non-electronic Tool for Interpreting Diagnostic Test Results. *Methods Inf Med* 2015; 54(3):283–7. doi: 10.3414/ME14-01-0091.