

Improving Health Care with Clinical Prediction Models

From Idea to Impact

Improving Health Care with Clinical Prediction Models

From Idea to Impact

Luc J.M. Smits

Sander M.J. van Kuijk

Laure Wynants

Improving Health Care with Clinical Prediction Models

From Idea to Impact

© 2026 The Authors

Authors:

Luc J.M. Smits | Maastricht University

ORCID: 0000-0003-0785-1345

Sander M.J. van Kuijk | Maastricht University Medical Centre

ORCID: 0000-0003-2796-729X

Laure Wynants | Maastricht University | KU Leuven

ORCID: 0000-0002-3037-122X

Published by: Maastricht University Press | Maastricht University, The Netherlands

Cover Design: Bas Verhage

Publication date: March 2026

ISBN hardcover: 9789403865706

ISBN softcover: 9789403870304

DOI: 10.26481/mup.2603



This work is published open access under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).

© 2026 The Authors. Published by Maastricht University Press on behalf of the authors. You are free to share and adapt the material, provided appropriate credit is given. To learn more, visit creativecommons.org/licenses/by/4.0.

For correspondence or general inquiries, please contact:

epid-cpm@maastrichtuniversity.nl

Please cite this book in the following way (Vancouver style):

Smits LJM, van Kuijk SMJ, Wynants L. *Improving Health Care with Clinical Prediction Models: From Idea to Impact*. Maastricht: Maastricht University Press; 2026. doi: 10.26481/mup.2603.

PREFACE

Over the past decades, we've been excited, and at times amazed, by the potential of clinical prediction models. These tools, which use statistical methods or artificial intelligence to estimate the likelihood of clinical outcomes based on patient data, are increasingly being relied on to support earlier diagnosis, guide treatment decisions, and tailor care to individual patients. Their applications now span a wide range of clinical areas, including oncology, cardiology, mental health, and critical care.

However, alongside that excitement comes a growing sense of disillusionment. Thousands of models have been developed and published, yet only a small fraction of them are applied in real-world settings. Sometimes the methodological quality is part of the problem. But just as often, care has not been taken to carry out the essential yet underappreciated steps involved in model development: ensuring the model meets a genuine clinical need, embedding it within a workable and user-friendly innovation, and preparing for meaningful evaluation and sustainable adoption in practice.

These challenges have practical implications for anyone interested in clinical prediction models—whether you're a researcher, clinician, policymaker, or student. This book is our attempt to shine a light on the full process. We discuss not only how to build and validate good models, but also what it takes to make them useful, usable, and used.

We—Luc Smits, Sander van Kuijk, and Laure Wynants—have spent years working in this field. Through research, teaching, and mentoring, we've helped develop new models and validate existing ones, and we've studied how they play out in everyday care. Along the way, we've felt the excitement and experienced the pitfalls—moments of genuine progress, and moments when good intentions fell short. These experiences have shaped the perspective we bring to this book.

This book is not the first to focus on clinical prediction models. Several excellent resources already provide detailed coverage of model development, validation, and updating. Our book is not intended to replace these books. Rather, its added value

lies in its comprehensive coverage of the entire journey—from the initial spark of an idea to the creation and sustainable implementation of an impactful prediction model-based innovation.

Alongside the statistical aspects of prediction modeling, we dedicate significant attention to defining the innovation goal, identifying and evaluating existing models, developing a prediction model-based innovation, evaluating its impact by means of decision modeling or empirical research, and facilitating its effective implementation in new settings.

This book is for:

- Health care professionals and researchers seeking a comprehensive but accessible guide to prediction modeling—from idea to implementation.
- Researchers looking for a practical manual on model development, validation, and impact assessment.
- Teachers and students in search of clear, practice-oriented course material on clinical prediction modeling.

We've written this book to be academically solid but approachable. If you have a basic understanding of statistical modeling—similar to that expected of a medical student—you'll be able to follow the material without difficulty.

Our aim is to help you maximize your chance of success in creating prediction model-based innovations that advance health care, while minimizing unnecessary research waste. Whether you're just starting out, or already immersed in a project, we hope this book will provide the guidance and tools you need to turn your ideas into meaningful improvements for patients.

We would welcome hearing from readers about their ideas, feedback, or experiences with prediction models in practice. Please contact us at: epid-cpm@maastrichtuniversity.nl.

Luc Smits, Sander van Kuijk & Laure Wynants

ACKNOWLEDGEMENTS

We would like to warmly thank Dr. Bas Verhage for bringing our ideas to life through the book's figures; Dr. Jane Banfield for her careful and insightful editing; Dr. Leonard Wee and Dr. Elena Albu for their expertise in reviewing the sections on Machine Learning and AI; Dr. Sabine Grimm for her thoughtful review of Chapter 13; Prof. Dr. Esther Versluis, Dr. Mindy Duffourc, and Iwan de Jong for their valuable feedback on the sections about the Medical Device Regulation and the EU AI Act; and Ron Aardening and Michel Saive of Maastricht University Press for their support in sharing our work with a wider audience.

ABOUT THE AUTHORS

Luc Smits (in the middle) works as a professor of clinical epidemiology at Maastricht University, the Netherlands. Sander van Kuijk (on the left) is associate professor of clinical epidemiology at the Maastricht University Medical Centre. Laure Wynants (on the right) works as an associate professor of epidemiology at Maastricht University, and as a visiting lecturer at KU Leuven, Belgium. All authors teach prediction modeling, evidence-based medicine and decision modeling to university students and postgraduates. Luc Smits is the principal author of Chapters 1, 2, and 10–15, Sander van Kuijk of Chapters 4–7, and Laure Wynants of Chapters 3, 8, and 9. While individual chapters were primarily drafted by one author, the book was very much a joint effort. All major decisions regarding structure, content, and style were made collaboratively, and each chapter benefited from extensive feedback, discussion, and revision by all three authors.



SCHOLARLY ENDORSEMENT

Clinical prediction models represent a key enabling technology for the fundamental shift to preventative healthcare—a transformation essential for sustaining modern health systems facing ageing populations and rising prevalence of long-term conditions. Over the past two decades, the field has matured considerably, yet scholarly work and existing textbooks have predominantly emphasised statistical methodology rather than the practical considerations that determine whether models become genuinely useful, usable, and used in clinical practice. *Improving Health Care with Clinical Prediction Models: From Idea to Impact* addresses this critical gap.

This book serves primarily as a practical guide for prediction model developers, supporting the complete journey from initial conception through development and validation to real-world deployment. Core methodological topics, including model development, external validation, performance assessment, and handling missing data, are covered accessibly without excessive mathematical formalism. Whilst these topics have been addressed elsewhere, Smits, van Kuijk, and Wynants succeed in presenting them with admirable clarity for readers with modest statistical backgrounds.

The book's principal strength lies in its novel emphasis on "prediction model-based innovation" and the systematically overlooked challenges of impact evaluation and implementation. The authors progress beyond simply developing accurate models to address how they can be transformed into comprehensive clinical tools integrating decision guidance, user training, and clinical workflow considerations. The final chapters focus on selecting models for real-world impact, developing workable innovations, evaluating impact through both decision modelling and empirical research, and implementing innovations sustainably across healthcare settings. This emphasis on translating models into meaningful healthcare improvements has received insufficient attention in previous textbooks.

The book is most suited for researchers and developers in applied settings who are relatively new to clinical prediction modelling. The unified tone and style enhances readability throughout, with the strongest chapters reflecting the authors' expertise in statistics and epidemiology. Refreshingly, the book prioritises fundamentals over

speculative trends, providing readers with a solid foundation for developing impactful prediction model-based innovations.

Niels Peek | Professor of Data Science and Healthcare Improvement at THIS Institute, Department of Public Health and Primary Care, University of Cambridge

This book provides a valuable resource for healthcare professionals and researchers aiming to navigate the complexities of developing, validating, and assessing the impact of clinical prediction models. The text covers a wide array of topics relevant to prediction research, with case studies for concrete illustrations. It discusses basic principles and modeling approaches, including consideration of modern machine learning techniques and innovative performance measures such as Net Benefit. Particularly noteworthy is the discussion on selecting a decision threshold, a critical step for the implementation of a prediction model within the framework of shared decision-making. The authors stress the importance of prediction model-based innovation (PMBI) and its implementation through thoughtful study design. The book is thoughtfully designed, featuring accessible visuals, minimal formulas, and hyperlinks for easy navigation.

Improving Health Care with Clinical Prediction Models: From Idea to Impact stands out as a significant contribution to the field of health informatics, clinical epidemiology and clinical research. This book will support those committed to advancing patient care by responsible application of prediction models.

Ewout Steyerberg | Professor of Clinical Biostatistics, Medical Scientific Division Manager at the Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht

TABLE OF CONTENTS

PREFACE	1
ACKNOWLEDGEMENTS	3
ABOUT THE AUTHORS	5
SCHOLARLY ENDORSEMENT	7
1 INTRODUCTION	19
Aim of the Book	22
Structure of the Book.....	22
2 DEFINING YOUR GOAL	27
Applications of Prediction Models in Health Care and Research.....	28
Obtaining a Clear Picture of Care-as-Usual	30
End User Involvement	35
Decision Analysis.....	37
Clinical Decision Trees.....	37
Concluding Remarks.....	40
3 SEARCHING FOR EXISTING PREDICTION MODELS	43
Systematic Literature Search	44
How to Build a Search String	45
Search Filters.....	47
Example.....	49
How to Determine Which Models are Worth Further Exploration	53
What’s next?	54
References and Further Reading.....	56
4 ASSESSING PREDICTION MODEL PERFORMANCE	59
Performance of Models for Continuous Outcomes.....	60
Overall Performance.....	61
Calibration	62

Performance of Models for Binary Outcomes	65
Calibration Measures for Binary Outcomes.....	65
Discrimination Measures for Binary Outcomes	70
Performance of Models for Time-to-Event Outcomes.....	73
In Summary.....	76
References and Further Reading	77
5 DESIGNING STUDIES FOR MODEL DEVELOPMENT AND VALIDATION.....	79
Epidemiological Study Designs.....	79
The Prospective Cohort Study: An Ideal Design for Prognostic Models	81
The Cross-Sectional Study: Ideal Study Design for Diagnostic Models.....	84
Alternative Study Designs	86
Randomized Controlled Trials	86
Case-Control Studies.....	87
Prospective Versus Retrospective Designs	88
Prospective Designs	88
Retrospective Designs.....	88
Choosing the Right Design for Your Study	89
Participant Selection.....	89
Selecting At-Risk Participants.....	90
Selection Bias.....	90
Measuring Candidate Predictors and Outcome.....	91
Validity and Reproducibility	91
Misclassification.....	94
Quantifying Predictors for Model Development	95
A Note on Confounding.....	96
What's next?.....	97
References and Further Reading	98
6 HANDLING MISSING DATA.....	101
Missing Data: A Threat to Precision and Validity	101
Complete Case Analysis.....	102

Missing Data Mechanisms.....	103
Missing Completely at Random.....	103
Missing at Random.....	104
Missing Not at Random.....	104
Some Remarks on Missing Data Mechanisms.....	105
Imputation Methods.....	105
Imputation With the Mean.....	105
Regression Imputation.....	107
Stochastic Regression Imputation.....	107
Multiple Imputation.....	109
Multiple Imputation in Practice.....	110
The Number of Imputations.....	110
Building the Imputation Model.....	110
Selecting the Model Type.....	111
After Imputation.....	112
References and Further Reading.....	113
7 EXTERNALLY VALIDATING A MODEL.....	115
Why External Validation is Necessary.....	115
Types of External Validation.....	116
External Validation of Prediction Models Based on Logistic Regression.....	117
External Validation of Prediction Models Based on Cox Regression.....	118
Computing Performance Parameters.....	118
Deciding whether model updating is necessary.....	119
Selecting a sample size for external validation studies.....	119
After validation.....	120
References and Further Reading.....	121
8 DEVELOPING YOUR OWN MODEL.....	123
Sample Size.....	124
Regression Models.....	127
Linear regression.....	127

Logistic Regression.....	130
Cox Regression	133
Assumptions	136
Linearity.....	136
Independent Observations.....	139
Additivity	141
Proportional Hazards	142
Normally Distributed Errors with Homogeneous Variance.....	143
Variable Selection	143
Univariable Prefiltering	144
Stepwise Selection	145
Disadvantages of Statistical Variable Selection	146
Combatting Overfitting	148
Using Pre-Existing Knowledge.....	148
Uniform Shrinkage.....	149
Model Updating	150
Estimating Performance, Revisited: Internal Validation.....	152
Apparent Performance	153
Internal Validation	153
Leave-Center-Out Cross-Validation	156
An Example of Model Development: The ADNEX Model.....	156
Beyond Basic Regression Model Development.....	158
References and Further Reading	159
9 ADVANCED REGRESSION AND ARTIFICIAL INTELLIGENCE.....	161
Advanced Regression Techniques	162
Regression for Multi-Category Outcomes.....	162
Regression for Multiple Correlated Outcomes.....	163
Time-to-Event Analysis in the Presence of Competing Risks.....	164
Regression to Predict Outcomes at Multiple Time Points.....	165
Predicting the Effect of an Intervention.....	168

Regression with Simultaneous Shrinkage	170
Artificial Intelligence and Machine Learning.....	171
Supervised Machine Learning.....	173
Is Machine Learning Better than Statistics?.....	174
Tuning Parameters.....	176
Popular Machine Learning Methods.....	177
How to Choose a Modeling Approach	187
References and Further Reading.....	189
10 SELECTING A PREDICTION MODEL WITH POTENTIAL IMPACT	191
Selecting a Prediction Model with Potential Impact	192
Shortlist	192
Predictive Performance	193
Face Validity	193
Ease of Predictor Measurement and Alignment With Point of Care	194
Net Benefit Across a Range of Thresholds of Predicted Risk: Decision Curve Analysis	194
Your Next Step.....	199
11 DEVELOPING YOUR PREDICTION MODEL-BASED INNOVATION	201
Assistive or Directive Use.....	201
Choosing Cut-Off Points.....	202
Semi-Quantitative Delphi Consultation.....	204
Minimizing Expected Total Costs.....	206
Decision Model-Based Impact Evaluation (for Cut-Off Determination)	207
Management Options: Some Additional Considerations.....	208
Further Adjusting the Target Population	209
Organization of Model Input and Output.....	209
Input	209
Processing	210
Output	211
Feasibility	212

Training Plan For the End User.....	212
Acceptability	214
Getting Ready to Test Your Innovation.....	217
12 FORMULATING YOUR RESEARCH QUESTION	219
Research Question.....	220
Patients.....	220
Intervention.....	220
Comparison.....	221
Outcome	222
Time.....	223
Example of a Research Question for an Impact Study	224
Superiority vs. non-inferiority	224
General Design Options for Impact Studies	226
13 PERFORMING A DECISION MODEL-BASED IMPACT EVALUATION	229
Choosing the Modeling Method	230
Structuring the Model.....	231
Collecting Input Data.....	233
Base-Case Analysis and Dealing with Uncertainty	234
Advantages and Disadvantages of Decision Model-Based Impact Evaluation	236
Examples of Decision Model-Based Impact Studies.....	237
References and Further Reading	247
14 PERFORMING AN EMPIRICAL IMPACT EVALUATION	249
Study Designs	249
Randomized Studies	249
Non-Randomized Studies	252
A Note on Cross-Sectional Studies.....	253
Getting the Most Out of Your Study: Baseline Variables, Impact Variables, and Process Variables.....	254
Baseline Variables.....	254
Impact Variables.....	254

Process Variables.....	255
Optimizing the Internal Validity of Your Empirical Impact Study.....	256
The Study Population Should be Representative of the Target Population.....	257
Exclude the Influence of Time Trends.....	259
Minimize Between-Group Differences in Natural Disease Course.....	259
Beware of the Unintended Influences of Study Setting on Participants’ Behavior and Experience.....	261
Exclude Differential Error in Outcome Measurement.....	262
Sample Size Considerations.....	264
Comparing Outcomes Between PMBIs and Usual Care: Key Considerations for Data Analysis.....	265
Superiority vs. Non-Inferiority: Different Goals, Different Interpretations.....	266
Intention-to-Treat vs. Per-Protocol: Two Windows on the Same Trial.....	267
Clustering and Complex Designs: Adjusting for Study Structure.....	268
Adjusting for Covariates and Confounders.....	268
Looking Beyond the Main Outcome.....	269
Examples of Empirical Impact Studies.....	269
References and Further Reading.....	278
15 IMPLEMENTING A PREDICTION MODEL-BASED INNOVATION.....	281
Dissemination.....	284
Dissemination Strategies.....	285
Audience Segmentation and Message Tailoring.....	285
Deploying Opinion Leaders and Knowledge Brokers.....	286
Your Dissemination Campaign.....	287
Implementation.....	288
Phase One: Initial Considerations Concerning Host Setting.....	290
Phase Two: Creating a Structure for Implementation.....	305
Phase Three: Ongoing Structure Once Implementation Begins — The Active Implementation Phase.....	308
Phase Four: Improving Future Applications.....	309
Closing Reflections.....	314

References and Further Reading	316
GLOSSARY.....	317
INDEX.....	331

KEYPOINTS OF THIS CHAPTER

Why Prediction Models Matter

Clinical decisions always involve uncertainty; prediction models help by turning patient data into personalized risk estimates.

Why Good Models Often Fail in Practice

Many models are never put into practice due to poor methods, failure to consider user needs, or lack of real-world testing and implementation.

Purpose of This Book

To guide you in building and applying prediction models effectively—from defining the problem to creating, validating, evaluating, and implementing your model in health care.

1

INTRODUCTION

“Medicine is a Science of Uncertainty and an Art of Probability”
Sir William Osler

Health care practice is full of uncertainty. A diagnosis may be right or wrong, complications may or may not develop, a therapeutic intervention may or may not work, or could lead to undesirable side effects. Despite this uncertainty, physicians and other health care workers must make clinical decisions every day. They do so not by ignoring uncertainty, but by factoring it in when weighing up different options.

A challenge for medical decision makers is that probabilities—such as the probability of a certain disease being present or a complication developing—vary between patients and may even change over time within a patient. Older patients, for example, face a higher risk of dying during surgery compared to younger patients, and the number of days survived after surgery strongly predicts future survival.

While clinical protocols and guidelines attempt to accommodate patient differences (in terms of severity of disease, age, sex, and other factors), they can address only a limited set of combinations. Moreover, even the most experienced clinicians cannot simultaneously process all patient-specific factors influencing risk.

That’s where *prediction models* come in. Clinical prediction models are algorithms or computational frameworks that effortlessly combine and assign weights to multiple patient characteristics. They produce an evidence-based and patient-specific estimation of the probability that a certain health outcome is present (a diagnostic model), or will occur in the future (a prognostic model). They hold the promise of enabling better decision-making based on the particularities of individual patients.

Since their introduction at the end of the last century, clinical prediction models have steadily gained popularity—a trend that shows no signs of slowing down, as

illustrated in Figure 1. Machine learning and artificial intelligence have further boosted interest in prediction modeling and its role in health care.

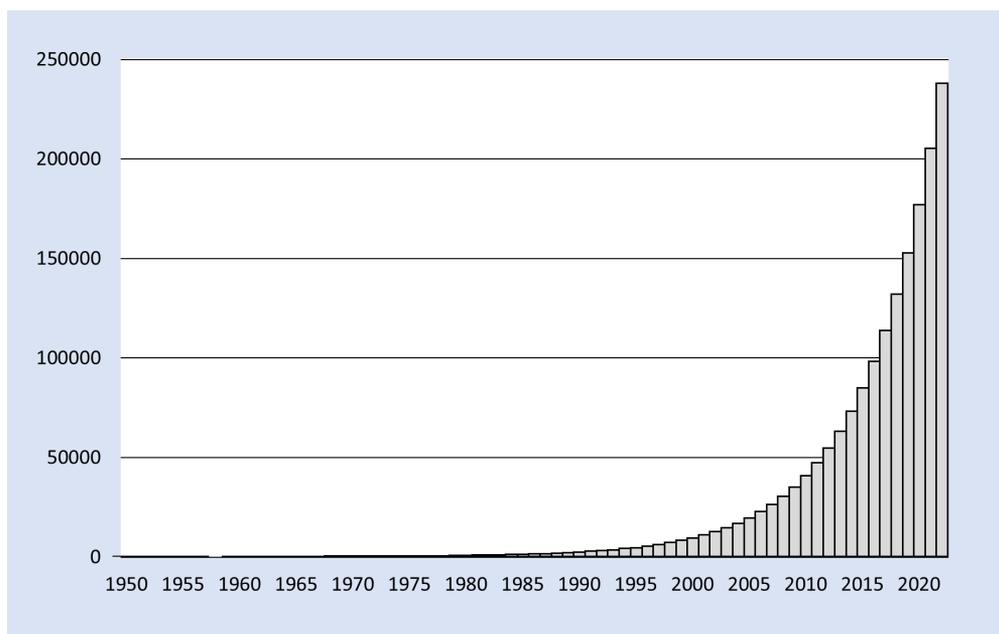


Figure 1. Cumulative Number of Articles on Prediction Model Development (Arshi et al., 2025¹).

Four decades of clinical prediction modeling have led to several important improvements in patient care. Primary care physicians and cardiologists are now better able to identify patients at increased risk of cardiovascular disease who are likely to benefit from cholesterol-lowering treatment. Gynecologists can more accurately detect women at elevated risk of developing preeclampsia or gestational diabetes during pregnancy, enabling preventive measures to be taken or intensified monitoring to be put in place when needed. Furthermore, oncologists now have tools to compare the likely effectiveness of different treatment options for women who have undergone surgery for breast cancer, supporting more individualized care.

Yet, despite these successes, a large proportion of prediction models are never put into clinical practice. A recent study asked a random sample of developers of prediction models whether their model was, or had been, used in medical practice,

¹Arshi B, Wynants L, Rijnhart E, et al. Number of publications on new clinical prediction models: a bibliometric review. *JMIR Med Inform.* 2025 Jul 4;13:e62710.

and no more than one in six reported that this was the case (Arshi et al., 2025²). In specific medical fields, the percentage of models used in patients may be under 1% (Kleinrouweler et al., 2016³, Salazar de Pablo et al., 2021⁴). From the perspective of the potential end-user, the situation isn't any rosier: while hundreds of clinical prediction models are available for general practitioners (and dozens of them are recommended in GP guidelines), 65–98% of European doctors in the target group do not use any of them (Brown et al., 2016⁵).

If a prediction model doesn't make it to practice, this doesn't necessarily mean that something has gone wrong. The process of prediction modeling can be compared to the development of new medicines: a new pharmaceutical must survive each stage of development in order to make it to the market. It may fail to work in a lab setting, prove ineffective in clinical practice, or have limited use due to severe side effects. Analogously, a prediction model may show limited predictive potential in the early stages of development. Or a model that initially appears promising may have disappointing performance when it is applied to different patient populations. And even models that predict well can turn out to have low predictive utility when their impact is evaluated in practice. These are all good reasons for ruling out a particular prediction model for use in clinical practice. Discarding some models at an early stage of development is a natural part of the scientific process.

Nevertheless, it is clear that the field of prediction modeling also produces much *research waste*. A variety of issues can contribute to this:

- New models are developed even though a search of the literature would reveal that good models, designed to achieve the same goal, already exist.

² Arshi B, Cowley LE, Rijnhart E, et al. External validation, impact assessment and clinical utilization of clinical prediction models: a prospective cohort study. *J Clin Epidemiol.* 2025 Jul;186:111902.

³ Kleinrouweler CE, Cheong-See FM, Collins GS, et al. Prognostic models in obstetrics: available, but far from applicable. *Am J Obstet Gynecol.* 2016 Jan;214(1):79-90.

⁴ Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, et al. Implementing precision psychiatry: a systematic review of individualized prediction models for clinical practice. *Schizophr Bull.* 2021 Mar 16;47(2):284-297.

⁵ Brown B, Cheraghi-Sohi S, Jaki T, et al. Understanding clinical prediction models as 'innovations': a mixed methods study in UK family practice. *BMC Med Inform Decis Mak.* 2016 Aug 9;16:106.

Rather than building new models, these existing models could be externally validated.

- Models are developed or validated using studies with methodological flaws, resulting in a high risk of bias. These models are unlikely to appeal to researchers or health care professionals.
- Models are developed that do not meet the needs of end users or take the interests of end users into account.
- Models are developed that are not easy to apply in practice: predictors are expensive, hard to come by, burdensome for the patient, or even impossible to measure at the time of intended use.
- Models are developed—and even externally validated—but no effort is made to turn them into workable decision aids, incorporate them into clinical routine in a way that is convenient for users, or evaluate their impact in comparison to usual care.

Aim of the Book

The overall aim of this book is to help you, the reader, increase the odds that your prediction modeling endeavor will be a success—one that leads to tangible improvements in health care. In line with this, we want to help you minimize the production of research waste in the field of prediction modeling.

Structure of the Book

The structure of this book mirrors the authors' views on the ideal development process, from the conception of a prediction-based innovation to its integration into health care practice. This doesn't mean you must follow every step exactly, nor does it imply that deviating from this structure will hinder good research. In fact, making thoughtful exceptions can often be more beneficial than rigidly adhering to a predefined format. However, our proposed structure is designed to help you avoid common pitfalls, reduce the risk of research waste, and increase the likelihood that your work results in meaningful improvements in health care.

Figure 2 illustrates the structure we propose and highlights how each chapter aligns within it.

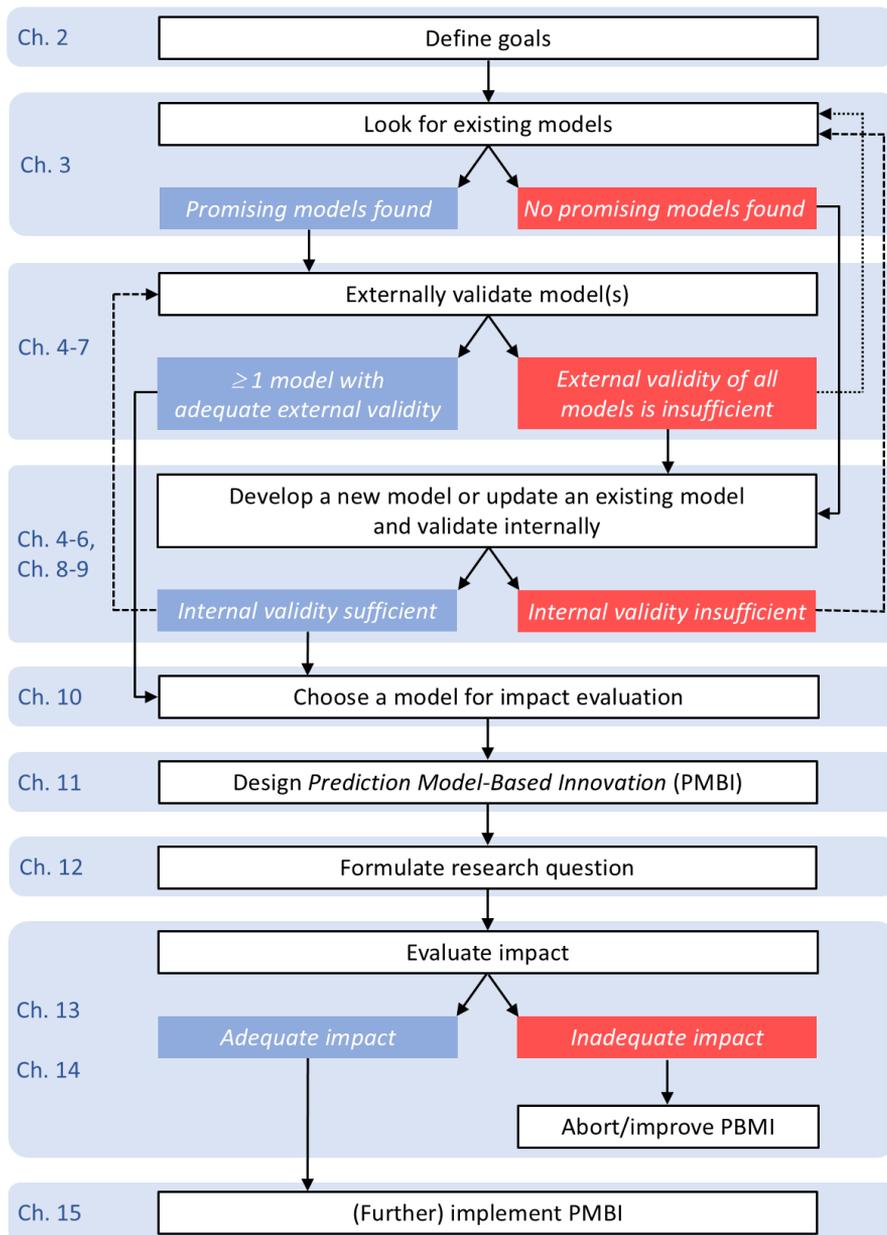


Figure 2. Improving Health Care with Prediction Models: From Defining the Problem to Implementing a Prediction Model-Based Innovation.

Any prediction modeling enterprise should start with asking yourself: What is the problem that I want to address? What prediction-model based innovation do I envision? *Chapter 2* helps you to reflect on these questions in more detail. If you have a clearly defined goal, the next step is to search the literature for any existing prediction models that will help you to pursue that goal. *Chapter 3* outlines the methods and strategies you can use to carry out your search. If you find a promising model—or even more than one—there may be no need to develop a model of your own! Nevertheless, it may be necessary to validate the model in a different population, possibly one from your own region. In *Chapter 5*, you will find an overview of various study designs that you can use to acquire data for such an external validation study—these designs are also useful if you want to develop your own model. *Chapter 7* explains the external validation process itself. If external validation yields disappointing results for each of the models under consideration, or if your initial search did not identify any promising models, you may decide to go ahead and develop your own. *Chapter 8* deals with the methodology of model development. If you've developed your own model, it must first undergo thorough internal validation—and ideally external (temporal) validation—before moving on to the next stage. That next stage involves evaluating the model's real-world impact when used by clinicians or other health care professionals. This kind of evaluation requires careful preparation. First and foremost, the right model must be selected—your own may not necessarily be the most suitable. *Chapter 10* discusses how to make that choice. To be effective, the model must be built into health care practice in a workable and acceptable way. *Chapter 11* helps you develop your prediction model-based innovation. The next three chapters are concerned with evaluating the impact of your innovation in comparison to usual care, starting with the formulation of your research question (*Chapter 12*), and followed by chapters covering impact evaluation using decision modeling (*Chapter 13*) and empirical research (*Chapter 14*). If the innovation is shown to be effective in practice, it should be possible to implement it in a wider range of settings, after any necessary modifications have been made. *Chapter 15* provides an overview of implementation methods and designs.

Between these chapters, there are sections that provide essential background knowledge to support your understanding of what follows. *Chapter 4* introduces key measures of model performance, while *Chapter 6* explains strategies for handling

missing data. Finally, *Chapter 9* offers an overview of advanced regression techniques and machine learning models used in clinical prediction.

KEYPOINTS OF THIS CHAPTER

Define Your Goal

Be explicit about the problem you want your model to solve and the improvement it should bring—ensure your effort has a meaningful impact.

Understand Current Care and Context

Map out how decisions are currently made, what information guides them, and where there's room for improvement—involve real users early to align your plan with practical needs.

Plan and Communicate with Clarity

Use tools such as decision trees to structure your ideas, compare strategies, and communicate clearly with stakeholders.

2

DEFINING YOUR GOAL

Before you engage in any prediction modeling activity, be it developing or validating a model, or assessing the impact of a prediction model-based innovation (PMBI), it is important that you clarify for yourself what problem you aim to solve or what improvement you wish to achieve by means of model-based prediction. This will increase the likelihood that your efforts pay off in terms of better patient outcomes, greater efficiency, or some other health care improvement.

In this chapter, you will discover the various ways in which prediction models can be, and have been, applied in health care practice. You will learn that when you start thinking about your own PMBI, it is important to have a clear picture of what's happening in current care ('care-as-usual'). You will find a checklist of aspects to consider regarding your tool, ranging from the target population to various outcomes on different levels. Furthermore, you will see how the use of a decision tree can help with structuring thoughts and communicating with important stakeholders. But first, let's specify what a clinical prediction model actually is.

In its simplest form, a clinical prediction model is an algorithm or a computational framework that produces an estimate of the probability of a health outcome for a particular patient on the basis of his or her characteristics. If it estimates the probability of a particular condition being present, the model is called a diagnostic or screening model; if it estimates the probability of an outcome occurring in the future, it is called a prognostic model. The latter type also includes models that estimate the probability of a certain outcome occurring following an intervention (e.g., in terms of success, or an unintended effect). Not all prediction models produce a probability as an output; some estimate the value of a variable on a continuous scale, for example, in relation to blood pressure, or quality of life. Although a model based on a single predictor variable could be used as a prediction model (as you can estimate probabilities with it), the vast majority of published prediction models use more than one variable (i.e., multivariable models).

Clinical prediction models can be developed using various types of regression modeling, and are increasingly being developed with machine learning or other artificial intelligence techniques. In this book, we address both ways of generating prediction models. Irrespective of the technique used to generate a clinical prediction model, the goal is always the same: to support clinical decision making by predicting states or outcomes in individual patients, either in the present or in the future. No matter how a model is generated, it will need to go through similar steps before it is ready to be implemented in clinical practice.

Applications of Prediction Models in Health Care and Research

Prediction models can serve as the foundation for a wide range of health care innovations. We intentionally use the word *foundation* because the mere existence of a prediction model is typically insufficient to qualify it as a health care innovation, no matter how accurate the model may be in predicting health outcomes. The model must also be accessible and user-friendly, for example. End users may have to be trained; they need to trust the model and believe that its use will improve clinical decision making. Cut-off points may have to be defined, and decision rules formulated for different outcomes of the model. It is therefore important and useful to distinguish between clinical prediction models and prediction model-based innovations (PMBIs). The necessary preparations and evaluations that pave the way from prediction model to effective PMBI will be discussed in later chapters of this book (Chapters 10 to 14). For now, let's take a look at some examples of PMBIs in health care.

Example 1. Avoiding Invasive Diagnostic Procedures

The ADNEX model was developed to estimate the probability of malignancy in an ovarian tumor in a minimally invasive way. The model uses a biomarker, ultrasound measurements, and a few other non-invasive measurements as predictors. It enables physicians to better adjust clinical management to the severity of the disease—to refer patients with a high risk of ovarian malignancy from gynecology to oncology, or to decide between surgery and conservative follow-up, for example.

Example 2. Adjusting the Treatment Decision Based on the Patient's Natural Outlook

Couples with fertility problems presenting to a gynecologist may suffer from underlying disorders that completely block fertility but may also just be unfortunate in the chance process that will lead to a pregnancy (or not). Prediction models (e.g., the models developed by Hunault et al., 2004) can be used by gynecologists to estimate the probability that a particular couple will conceive without medical intervention within 1 year. These models can help to prevent the side effects and costs associated with overtreatment.

Example 3. Targeting Preventive Interventions at Individuals with an Increased Risk of Developing Disease

Myocardial infarction and other forms of cardiovascular disease can be prevented, in part, by individuals taking cholesterol-lowering medication (statins). However, statins can have (rare) side effects and their cost-effectiveness is low in persons with low risk of cardiovascular disease. Therefore, before deciding whether or not to prescribe statins, many GPs and cardiologists make use of prediction models, such as the QRISK3 , that estimate the risk of cardiovascular events occurring in the future.

Example 4. Counseling Patients on the Outlook of Their Disease

Conveying personalized information to a patient about his or her prognosis can be a goal in itself, for instance, when treatment has been decided upon or started with, or no treatment alternatives are obvious or available. The patient can use the information to adjust their expectations, level of care, and practicalities of their life to the expected outlook.

These are only a few examples of the possible applications of prediction models in health care. At this time, you may want to reflect on your own ambitions. What is it that you want to accomplish? Would you like to help improve patient outcomes? Make health care more efficient? Be able to inform patients more accurately about their prognosis? A combination of the above? Although these questions seem obvious, the medical literature is full of examples of prediction models developed by researchers who did not carefully consider what kind of improvement the implementation of their products should bring about. This is undoubtedly one of the

reasons why so many prediction models do not even make it past the development stage, and do not get validated and implemented.

Defining your goal implies that you gain insight into how things currently work in the care segment that you want to make an improvement in, and how prediction, or better prediction, can contribute to the improvement you envision. Basically, you should pose yourself a number of questions and answer them as thoroughly as possible. Be sure not to skip this step, however tempting it may be to jump ahead to the next stage.

Obtaining a Clear Picture of Care-as-Usual¹

To gain a clear understanding of care-as-usual, it may be helpful to answer the following questions:

- Which patients are at stake, in terms of disorder, stage, age, sex?
- What are the options to choose from? These can be treatment options, but also diagnostic procedures.
- What are the advantages and disadvantages of each option? Think about costs, patient burden, potential benefit, complications.
- Which patient-related information is the choice based on?
- In what way are the various pieces of information compiled to underpin the choice?
- Is the patient (or their representative) actively involved in the choice (i.e., is there shared decision-making)?
- Are there guidelines about decision making in this specific situation?
- Does care-as-usual differ substantially from the guidelines?
- Is there any substantial variation between physicians / institutes / countries in care-as-usual?
- If any prognostic or diagnostic assessments are made in care-as-usual, is there any substantial variation in accuracy between the health care professionals who carry out the assessments (e.g., related to profession, experience)?

¹ If we make a reference to *care*, this also includes *preventive* care. Prediction models can be designed to support this type of care too, which can be targeted at both patients and healthy people.

If you are not involved, as a care professional or otherwise, in the usual care that you are now describing, it is usually necessary to engage with clinicians or others who are involved in it, even if specific guidelines are available. The reason for this is that care-as-usual (as implied in the above questions) does not always follow guidelines closely enough to be able to consider them to be a true reflection of reality. If there is a big difference between usual care and the guidelines, there is probably a reason for this divergence—which is worth exploring since it may underscore the need for a change. Another reason for gathering information from other care professionals involved in decision making in care-as-usual is that usual care may vary between physicians, institutions, or countries. If such variation does exist, and you expect your PMBI to be widely adopted, make sure that you are aware of these differences. If there are several dominant care-as-usual strategies, it may be worth comparing each of them separately to the new model-based strategy in the impact analysis phase (see Chapters 10-14).

Next, consider what you would like—or expect—to improve by using a prediction model instead of usual care. Try to go beyond technical accuracy (e.g., “we will be better able to estimate the likelihood that disorder X is present in patients with symptoms Y”, or “our estimation of the risk of complications within a month after surgery Z will be greatly improved”). Rather, indicate, in qualitative terms, the intended or expected effects of using the model on patient outcomes, costs, efficiency, or other relevant outcomes (e.g., the facilitation of shared decision-making) in comparison with care-as-usual.

Describe your prediction model by addressing the following questions. Note that you may not know the answer to some of these questions at this stage:

- For which type of patients (or more generally, individuals) will the model be used?
- In which clinical context will the model be used (e.g., primary care, secondary care, tertiary care)?
- At what moment will the model be applied (e.g., at presentation, immediately after diagnosis, immediately after starting treatment)?
- What outcome will the model predict, and over what time period or periods?
- What kind of predictors will be used in the model, and how will these be assessed?