Leiden University
Medical Center

Prediction estimands

STRATOS workshop – estimands in time-to-event analysis
17th Sept 2024

Nan van Geloven

n.van_geloven@lumc.nl

DEPARTMENT OF BIOMEDICAL DATA SCIENCES
LUMC

Why use estimands in prediction?

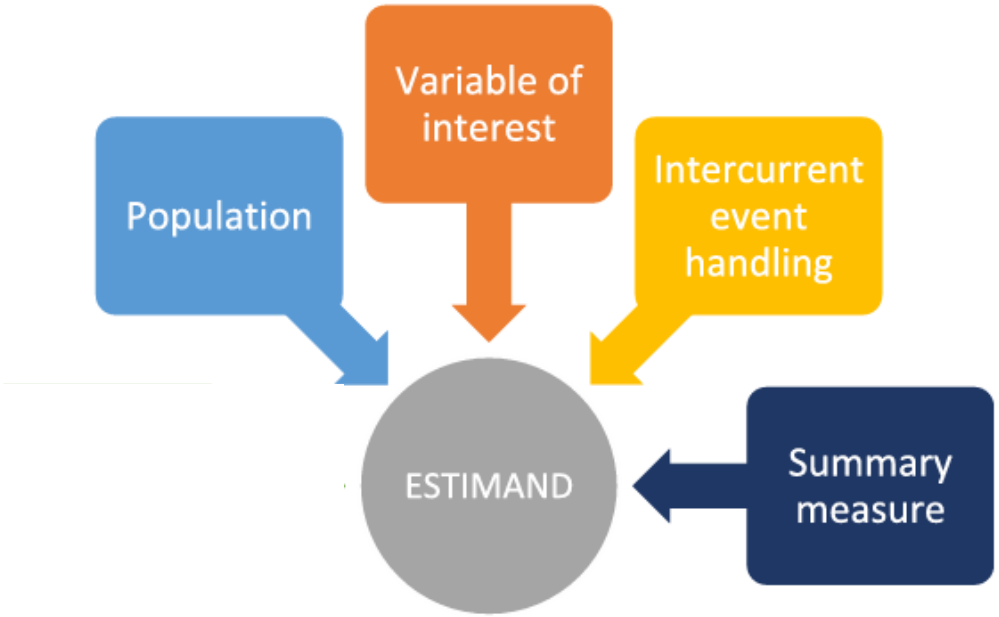


Fig. 1 The five attributes of an estimand according to the ICH E9 (R1) addendum

Lawrance et al 2020

Answer: clinicians do not just want to predict; they want to use predictions in their decision making

Reporting guidelines

PROBAST ‘risk of bias’ tool (Moons et al 2019):

> “Prognostic models can be used to aid decisions about preventive lifestyle changes, therapeutic interventions, or monitoring strategies”

TRIPOD+AI reporting guideline (Collins et al 2024) :

> “Their primary use is to support clinical decision making, such as … initiate treatment or lifestyle changes”

Examples clinical associations

American joint committee on cancer:

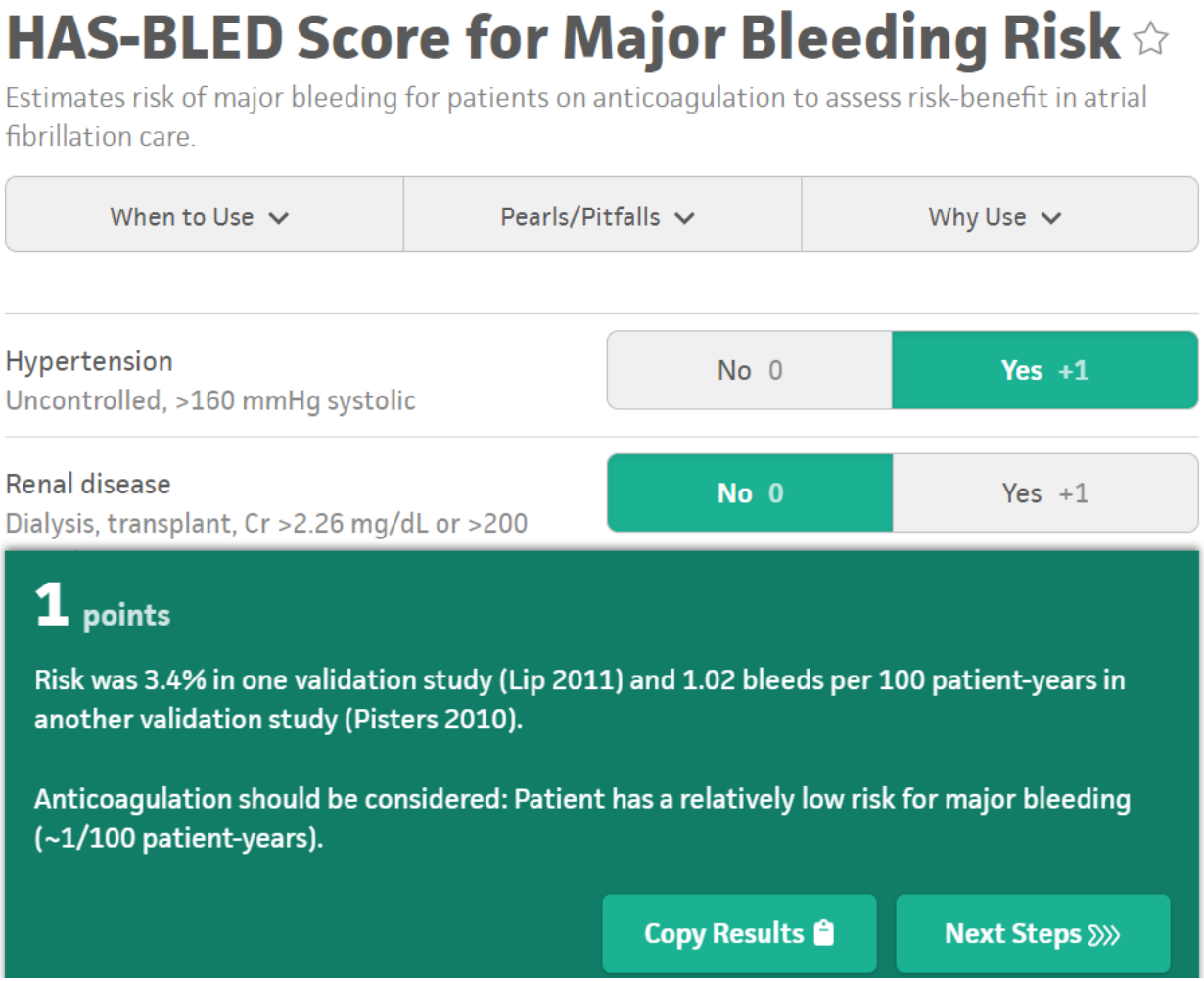
> "Describe how … this prediction might change clinical practice (e.g., patients always ask about this outcome prior to choosing a specific treatment. Or better, something actionable, such as if the prediction is < X, I would not recommend this treatment.)"

European Society of Cardiology on patients with atrium fibrillation:

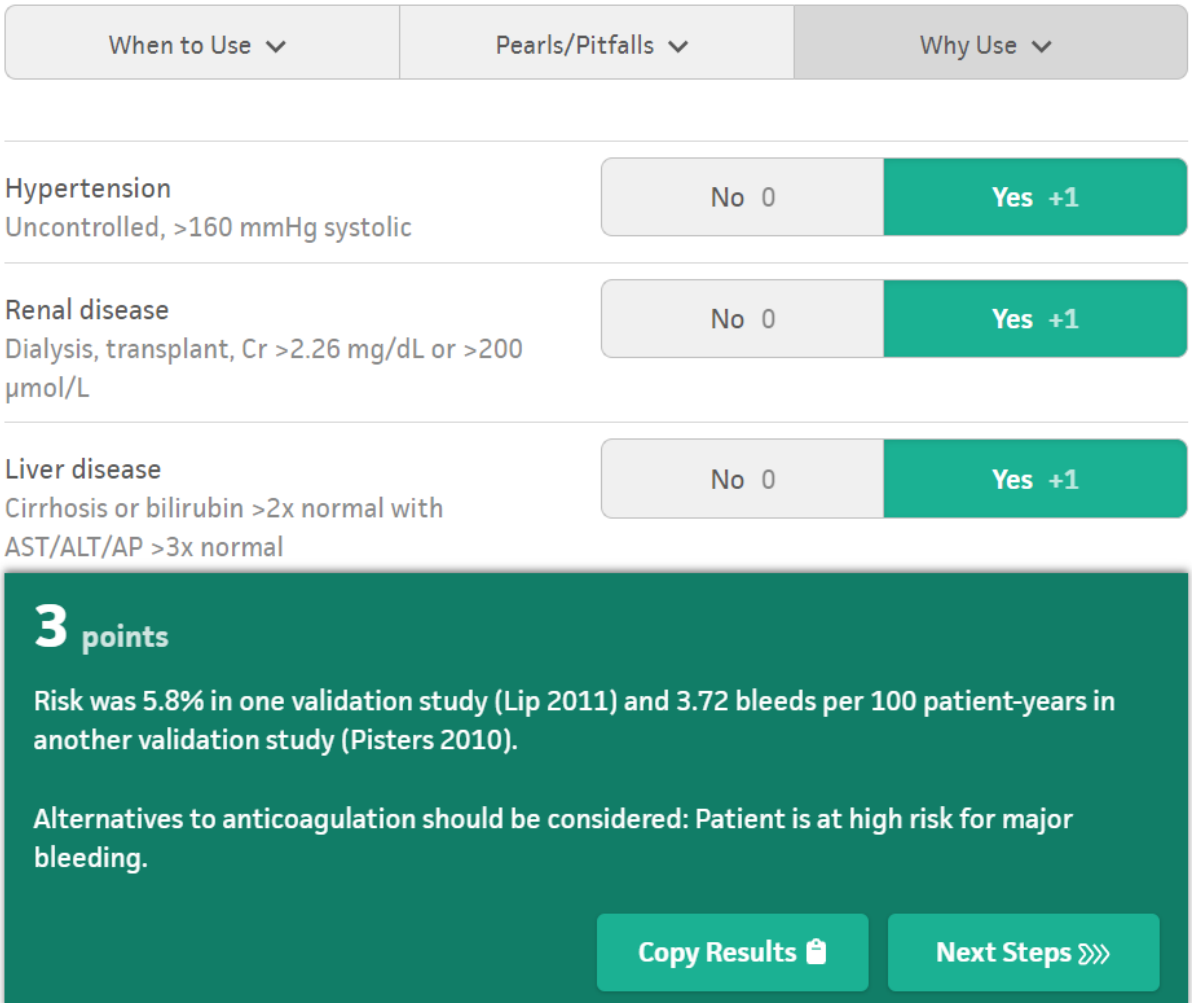
> "Long-term oral anti-coagulant therapy should be based on the stroke risk profile (using the CHA2DS2-VASc score), balanced against bleeding risk (e.g. HAS-BLED score)."

Kattan et al 2016
ESC 2020 guidelines

Risk of major bleeding risk in patients with atrial fibrillation



Risk of major bleeding risk in patients with atrial fibrillation



Systematic review covid-19 prediction models: 64% of papers recommend their model for treatment decision-making

Background. We aimed to clarify the high-risk factors with multivariate analysis and establish a prediction of disease progression, so as to help clinicians to better choose therapeutic strategy.

---

system to impact patient care after further validation with externally collected clinical data. Clinical decision support tools for COVID-19 have strong potential to empower healthcare providers to save lives by prioritizing critical care in patients at high risk for adverse outcomes.

---

the external validation set. We also developed a web tool to implement our predictive model. Clinicians can use this web tool to predict the mortality risk of COVID-19 patients early. For those patients with a relatively higher probability of death (e.g. >40%), more interventions could be adopted at an earlier stage by clinicians.

Prosepe et al, Front Epid 2022 (based on review by Wynants et al 2020)

Can predictions improve treatment decisions?

Yes: Risk/benefit profile of treatments may vary according to underlying risk:

low risk patients may not need treatment

high risk patients should be treated

But we need to be clear about what we mean by ‘underlying risk’?

- risk without ever being treated?
- the risk under current treatment strategies?
- often unspecified

Why is this a problem?

Hilden and Habbema (1987):

> **“Prognosis cannot be divorced from contemplated medical action”**

Development data typically observational

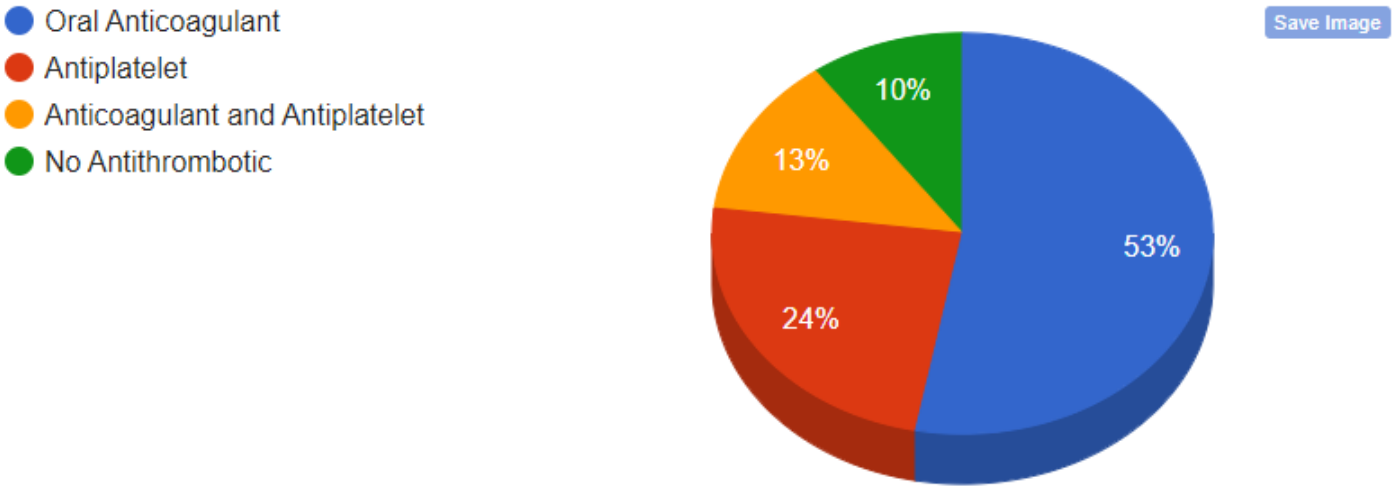
> includes some patients receiving the treatment that the prediction model wants to inform

Without clear description of the role of those treatments, how can these predictions support decisions in new patients?

Development cohort HAS-BLED

Supporting Evidence

The HAS-BLED manuscript originally included a variety of ambulatory and hospitalized patients with atrial fibrillation.[1] This patient population included those taking oral anticoagulants, antiplatelet agents, and no antithrombotic therapy (see below).



Clinician: “You have a low risk of bleeding, so you can use anticoagulants”
Patient: “Is that low risk with or without treatment”
Doctor: “Uhh…”

Pisters et al 2010

How use estimands in prediction?

Predictions under interventions

1. Formulate a prediction estimand aligned to the targeted treatment decision

2. Assess causal assumptions and estimate accordingly

3. Evaluate predictive performance against outcomes under interventions (also assumptions needed)

Letter to TRIPOD+AI

Prognostic models for decision support need to report their targeted treatments and the expected changes in treatment decisions

Dear Editor,

TRIPOD+AI ensures that researchers report sufficient information for readers to verify whether a published prediction model is fit for purpose, doing the field a tremendous service by improving standards and reducing research waste. But when it comes to prognostic models whose primary use is to support clinical decision making, we propose researchers report on two crucial considerations related to the role of treatments.

15 May 2024

Wouter A.C. van Amsterdam
Assistant Professor
Giovanni Cinà, Vanessa Didelez, Ruth Keogh, Niels Peek, Matthew Sperrin, Andrew Vickers, Nan van Geloven, Ur Shalit
University Medical Center Utrecht
Heidelberglaan 100, Utrecht
WvanAmsterdam

1. Formulate the prediction estimand tailored to the treatment decision you want to inform

- Population: Patients ≥18 years of age, diagnosed with new-onset atrial fibrillation
- Moment(s) of intended use of prediction model: At first doctor consult after diagnosis
- Intervention option: Start using oral anti-coagulants: $a = 1$
- Outcome and prediction horizon: major bleeding within one year: Y (0 or 1)
- Predictor(s): as in HAS-BLED: X

Prediction estimand:

$E(Y^{a=1} \mid X)$, with Y^a the counterfactual outcome under treatment a

"one-year bleeding risk if patient decides to use anti-coagulants"

Sperrin et al Stat Med 2018, van Geloven et al Eur J Epi 2020, Boyer et al, ArXiv 2023, Luijken et al. ArXiv 2023

2. Estimating predictions under interventions

- Individual patient data from RCTs, e.g. PATH statement[1]
  - + confounding not a problem
  - – limited sample size
  - – absolute risks may not be representative (covariate shift)

- Combine observational data with treatment effects from RCT's (offset method)[2]
  - + confounding not a problem
  - – other assumptions needed (transportability, marginal RCT effect = conditional effect in prediction model)
  - – does not allow treatment heterogeneity

- Observational data[3]
  - + large, representative data sources
  - – challenge in addressing confounding

[1]Kent et al, Ann Intern Med 2020;
[2]Xu et al, Am J Epi, 2021, van Amsterdam and Ranganath, J Caus Inf 2023
[3]Van Geloven et al Eur J Epi 2020; Boyer et al, ArXiv 2023; Dickerman et al 2022; Keogh and van Geloven 2024

2. Estimating predictions under interventions – time varying treatment

Revisit the estimand:

- "one-year bleeding risk if patient decides to use anti-coagulants"

What does 'decides to use' mean? Starts now, but may stop tomorrow?

Most interventions are not one time only.

More informative:

- "one-year bleeding risk if patient decides to start anti-coagulants at baseline and continues this during the year (or until bleeding occurs)"

Risk estimand: $E(Y^{\underline{a}_0} \mid X)$, with $\underline{a}_0$ = (1,1,1…)

-> likely need to account for time-varying confounders (both in trials and observational data)

Lin et al 2021, Dickerman et al 2022, Boyer 2023, Keogh and van Geloven 2023

3. Evaluate counterfactual predictive performance against outcomes under interventions

Performance metrics (discrimination, calibration, prediction error)

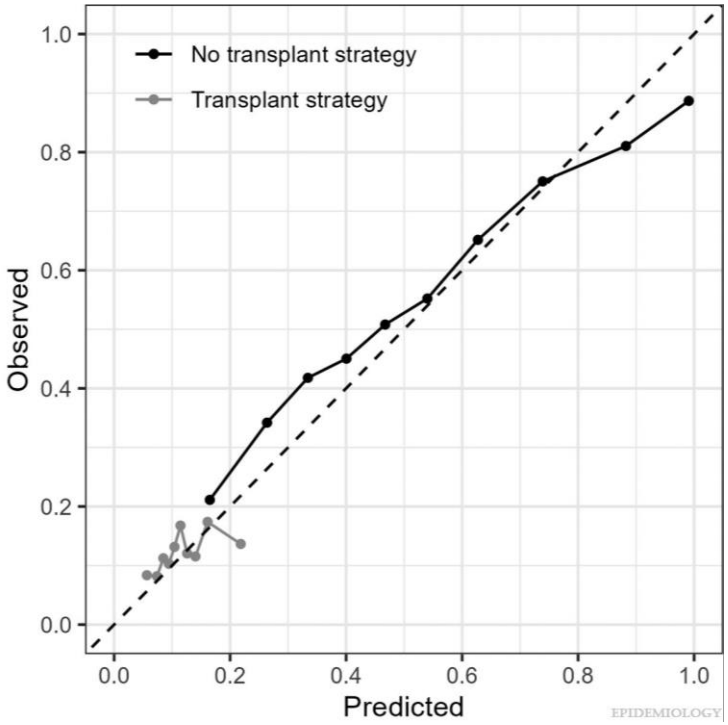
compare *estimated risks* to *observed outcomes* in a test/validation dataset

Both need to have same target -> *observed outcomes* also need to be estimated "under interventions"

This requires modified (counterfactual) performance metrics

3. Evaluate counterfactual predictive performance against outcomes under interventions – observational data

estimated by KM with time-varying reweighting to account for deviations from treatment strategy



Calibration plot: predictions against 'observed' proportion of patients with and without treatment

Similar adjustments for c-index, AUCt, Brier score, using time-varying reweighting

Boyer 2023, Keogh and Van Geloven 2024

My view on the ‘three tasks of data science’

Hernán, Hsu, Healy, Chance 2019
Carlin and Moreno-Betancur, arXiv 2024

Concluding

If a prediction model is intended to support treatment decisions:

1. Formulate prediction estimand “under interventions”
2. Estimate based on causal assumptions
3. Evaluate predictive performance “under interventions”

This ensures the right target, it does not yet assess the impact of using the prediction model on patient outcomes

(Impact depends on many other things: model accuracy, cut-points suggested for treatment decisions, do people adhere to the suggested cut-points?, does it predict better than clinicians? is stratifying treatment decisions by risk (cost)-effective?, is the treatment effective?...)

References

Van Geloven N et al. (2024) The risks of risk assessment: causal blind spots when using prediction models for treatment decisions. arXiv:2402.17366

Keogh RH, Van Geloven N. (2024) Prediction Under Interventions: Evaluation of Counterfactual Performance Using Longitudinal Observational Data. Epidemiology. 2024 May 1;35(3):329-339

Van Amsterdam W et al. (2024) When accurate prediction models yield harmful self-fulfilling prophecies. arXiv:2312.01210

Luijken K et al. (2023) Risk-based decision making: estimands for sequential prediction under interventions. arXiv:2311.17547

Boyer C et al. (2023) Assessing model performance for counterfactual predictions. arXiv:2308.13026

Dickerman BA et al. (2022) Predicting counterfactual risks under hypothetical treatment strategies: an application to HIV. Eur J Epidemiol. doi: 10.1007/s10654-022-00855-8

Prosepe I et al. (2022) The disconnect between development and intended use of clinical prediction models for Covid-19: a systematic review and real-world data illustration. Front. Epidemiol.

Van Geloven N et al. (2020) Prediction meets causal inference: the role of treatment in clinical prediction models. Eur J Epidemiol. 2020;35:619–30.

Sperrin M et al. (2018) Using marginal structural models to adjust for treatment drop-in when developing clinical prediction models. Stat Med. Dec 10;37(28):4142-4154. doi: 10.1002/sim.7913