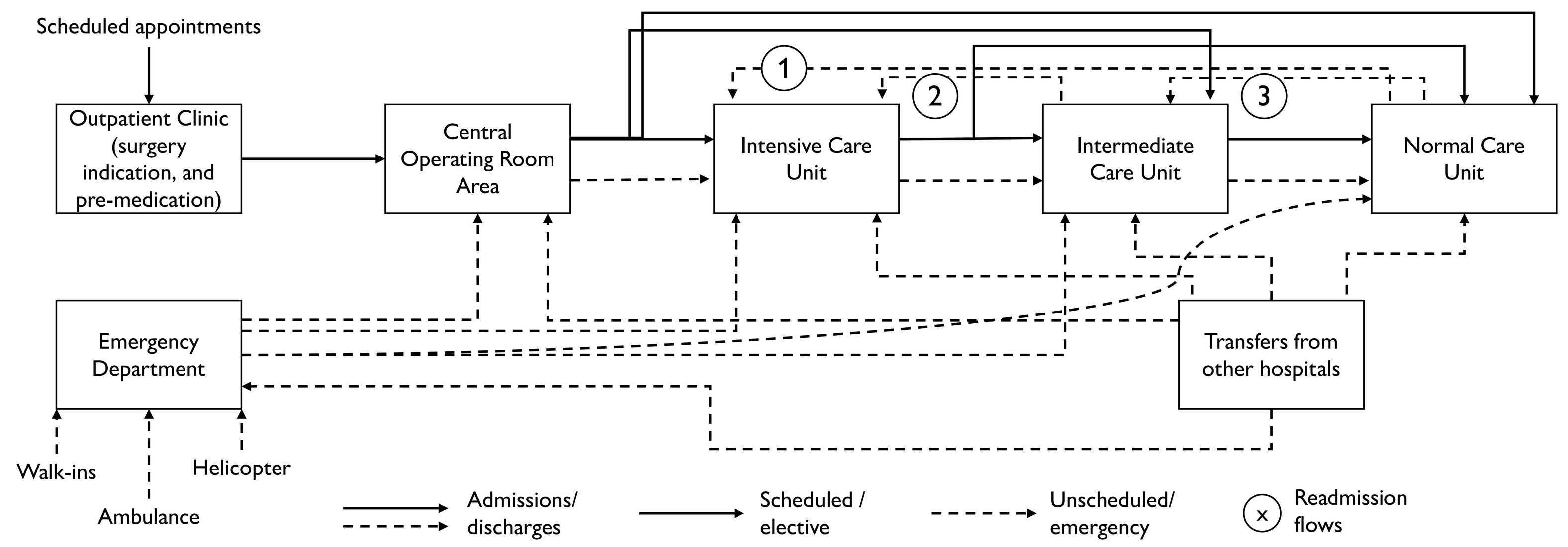
Causal Effects and Policy Learning for Intensive Care Unit Discharge Decisions to Solve Hospital Process Bottlenecks

Justus Vogel[1,c], Johannes Cordier[1], Miodrag Filipovic[2]

(1) Chair of Health Economics, Policy and Management, School of Medicine, University of St.Gallen, St. Gallen, Switzerland; (2) Cantonal Hospital of St. Gallen, St.Gallen, Switzerland; (c) Corresponding/ presenting author | Preliminary results, August 2025

BACKGROUND, RESEARCH QUESTION, OBJECTIVE



Notes: Patients can enter a hospital as scheduled, plannable cases (full arrows) through the outpatient clinic and then days to weeks later, on the day of surgery, through the central operating room area, or as unscheduled/ emergency cases (dashed arrows) through the emergency department. From the emergency department, patients are pushed onto the process area with free capacity and/ or where they need to receive care. Besides, patients enter a hospital as unscheduled transfers from other hospitals, commonly through the emergency department. Patients are pushed through the different process areas in a scheduled or oftentimes unscheduled manner. We are interested in how to minimize readmission flows (1), (2), and (3).

Intensive Care Unit (ICU) has fixed capacities

Number of arriving patients and length of stay in the ICU, i.e., demand, are uncertain

The ICU is a classic process bottleneck

Arriving patients are in immediate need for intensive care

Less critical patients then are discharged (too) early [1]

Research Question: Who should be selected for discharge?

Research Objective: Utilize individualized average treatment effects (IATEs) for learning a policy that reduces readmission risk across all discharge decisions / discharges

RELATED LITERATURE

Predictive Machine Learning and causal inference are common in (applied) Operations Research (OR) [2, 3]

Discharge decision problem has been research in the OR community [4]

Causal machine learning applications are rare – especially for *causal* decision support [5-7]

EMPIRICAL MODEL AND SETTING

Decision problem: We define a bed capacity constraint $B_a \geq 0$, considering incoming and outgoing patient flows on a given day

$$B_a(x, a', d') = B - \left(\sum_{i=1}^{I} x_i + \sum_{e=1}^{E} a_e - \sum_{c=1}^{C} a_c + \sum_{r=1}^{R} a_r + \sum_{u=1}^{U} a_u^{ant}\right) + \left(\sum_{n=1}^{N} d_n + \sum_{p=1}^{P} d_p^{early}\right) + \left(\sum_{q=1}^{Q} d_q^{early} - \sum_{v=1}^{V} a_v^{add}\right)$$

Decision policy: Select those patients for discharge that have the lowest change in readmission risk due to the early discharge, minimizing overall readmission risk

Minimize

$$R(d') = \sum_{i \in C}\left(d_i' \times (RR_i^t - RR_i^{t+1})\right)$$

Causal forests (generalized random forest [8], modified causal forest [9]: IATEs are estimated from treated and untreated examples in leaves of all trees part of a forest for test point x, under unconfoundedness and with common support

$$IATE(W; x) = E[Y(1) - Y(0) | X = x]$$

$$\hat{\tau}(x) = \frac{1}{N_\ell^1} \sum_{i: X_i \epsilon \ell(x;T)} Y_i^1 - \frac{1}{N_\ell^0} \sum_{i: X_i \epsilon \ell(x;T)} Y_i^0$$

Data source and observation period: ICU stays Department of Surgical Intensive Care Medicine of the Cantonal Hospital of St. Gallen admitted between January 01, 2016, and December 31, 2023

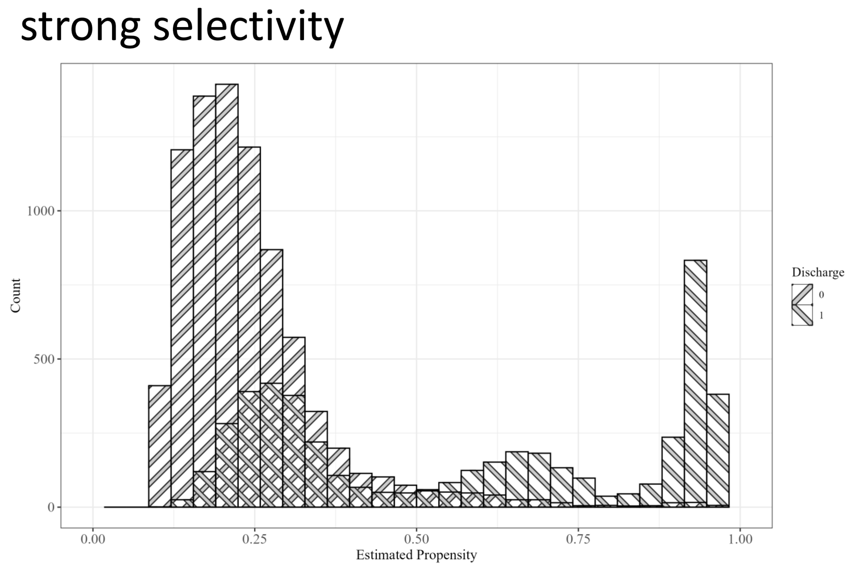
Sample: 12,950 ICU stays, 11,873 unique cases after application of exclusion criteria

Features: We leverage close to 4,700 variables as learning features

Identifying assumptions [10]: We may claim unconfoundedness because we observe all confounders that might influence treatment assignment by the physician and patients' readmission risk – unobserved patient factors such as therapy adherence cannot influence treatment assignment!

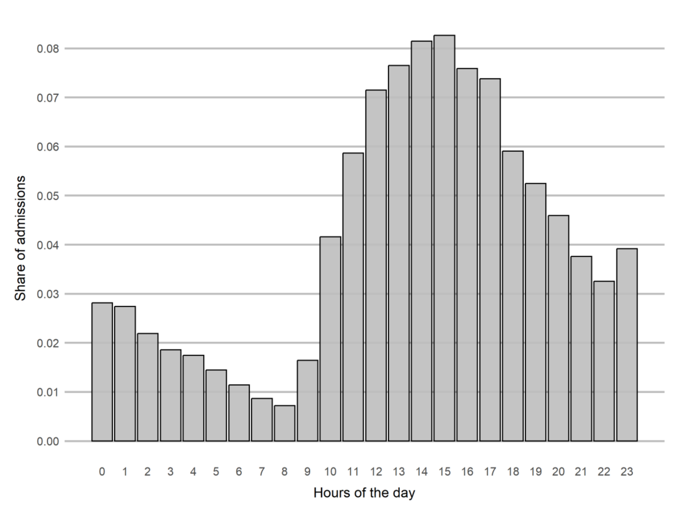
Exogeneity (outcome is observed after treatment) and Stable-Unit-Treatment-Assumption (treatment assignment of one patient cannot directly influence the outcome of another patient) are respected

Common support / overlap: Medium to strong selectivity



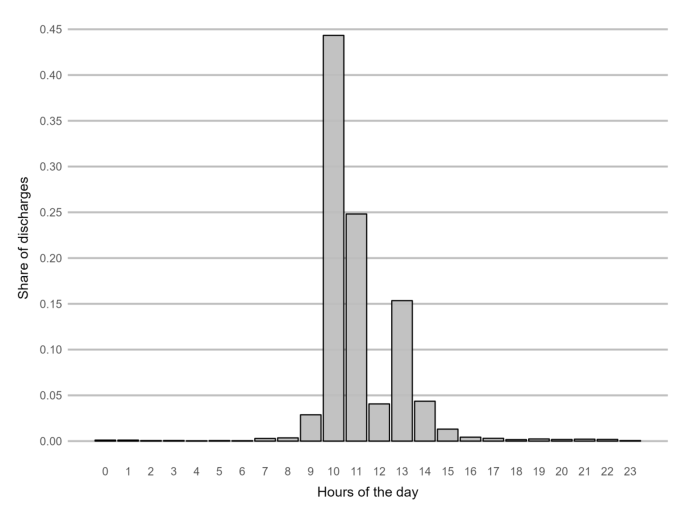
EMPIRICAL MODEL AND SETTING (continued)

Arrivals per hour of the day



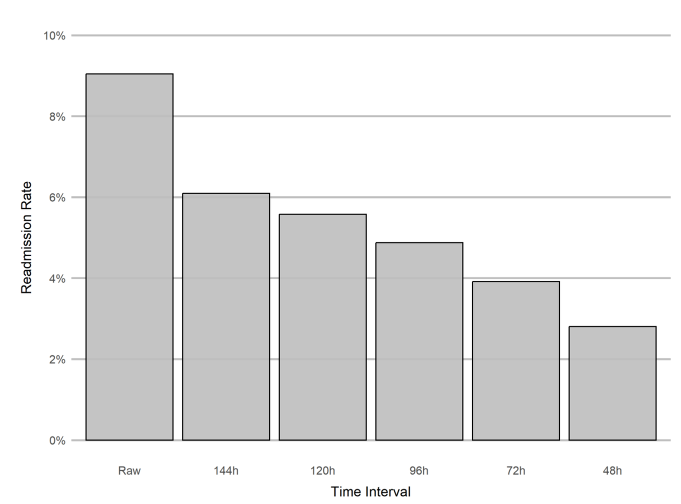
Notes: Includes 14,121 ICU stays of 12,932 patients with a positive or waived general consent admitted at our partner hospital between 2016 and 2023, no exclusion criteria applied.

Discharges per hour of the day



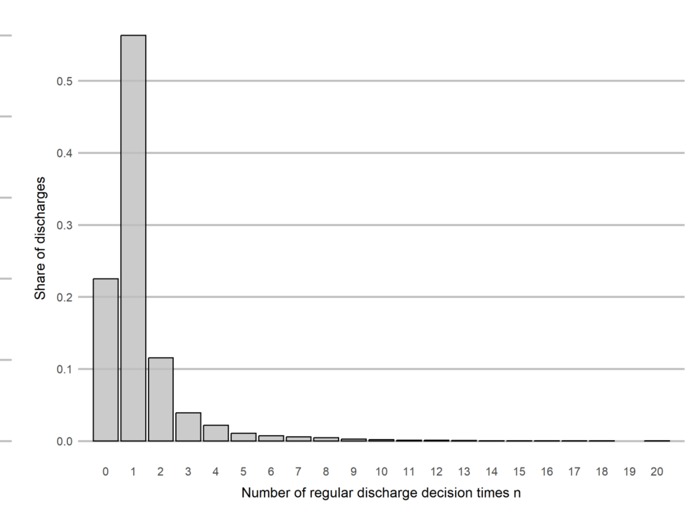
Notes: Includes 13,532 ICU stays of 12,414 patients with a positive or waived general consent admitted at our partner hospital between 2016 and 2023, excluding patients who died during their ICU stay, as the timing of these patients' "discharge" does not occur according to our formalized decision process.

Outcome: Raw readmission rate



Notes: We show readmission rates according to how much time passed between discharge and readmission. It is debated what type of readmission rate is the most relevant for ICU management. We are interested in reducing readmissions regardless of their timing. Therefore, we use readmission regardless of the time between discharge and readmission as outcome.

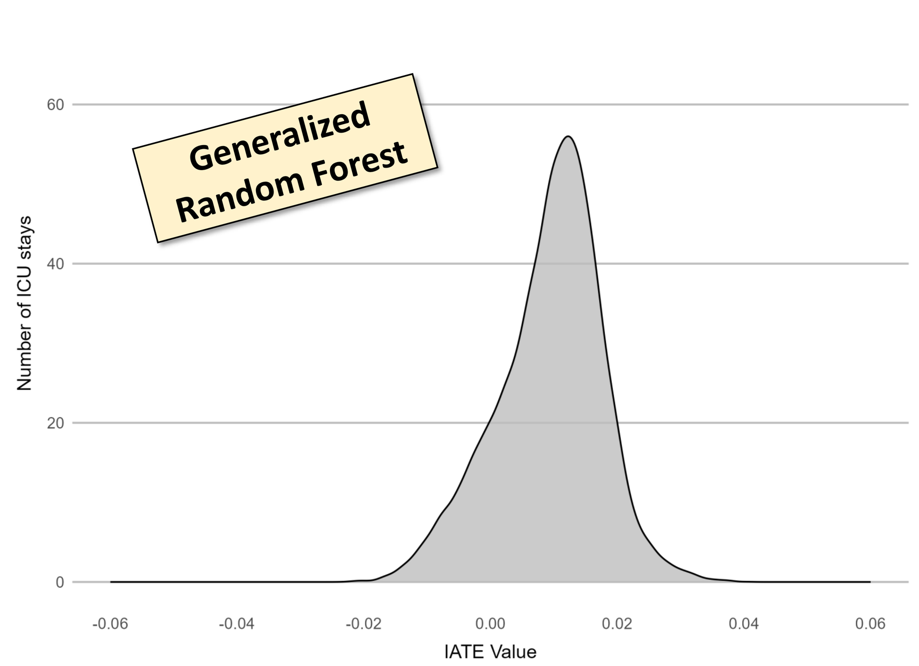
Treatment: Discharge t vs. $t+1$



Notes: To establish a binary treatment framework, we follow a four-step approach. (1) We define criteria which indicate that a patient cannot be discharged for medical reasons (mechanical ventilation, catecholamines). Once these measures stop, we identify a patient as dischargeable. (2) We define n as the number of times a patient passed the regular time (07:00 a.m.), after becoming dischargeable. (3) We define a random point in time d for each patient i with $d_i = \{max(n_i - 1{,}0), n_i\}$ with equal probability for $d_i = max(n_i - 1{,}0)$ and $d_i = n_i$. (4) we assign a positive treatment status $W_i = 1$ to all i for which $d_i = n_i$ (intervention group) and a negative treatment status $W_i = 0$ to all i for which $d_i = max(n_i - 1{,}0)$ (control group). Note that for all i, confounders measured over time (e.g., vital signs) are always considered with reference to d_i.

RESULTS

First modelling step: IATE estimation with Generalized Random Forest (GRF) and Modified Causal Forest (MCF)

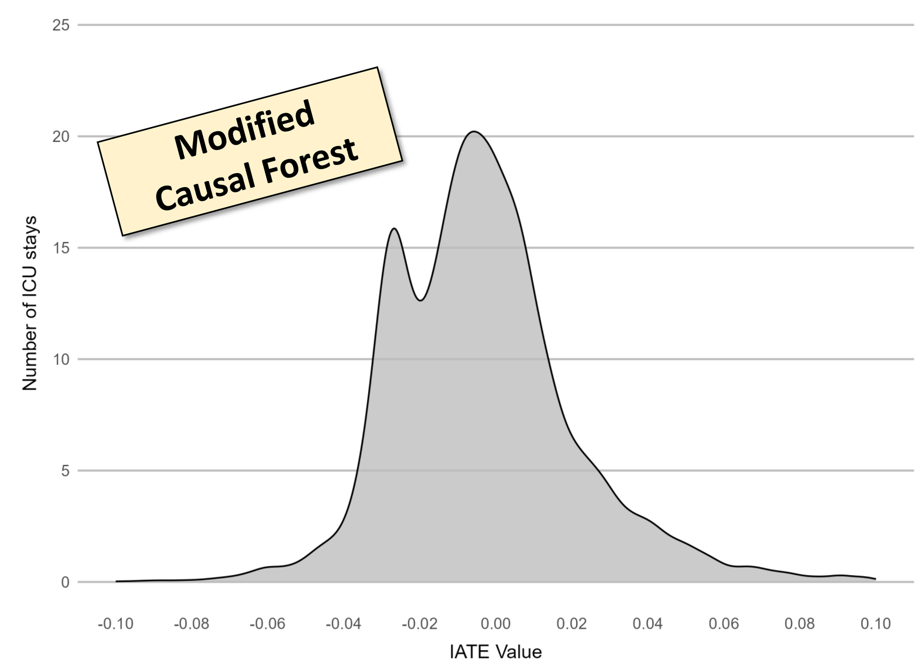


An IATE of, e.g., 0.01 means that a discharge today as opposed to tomorrow increases a patient's readmission risk by 1%-point (+11% increase as compared to average readmission rate)

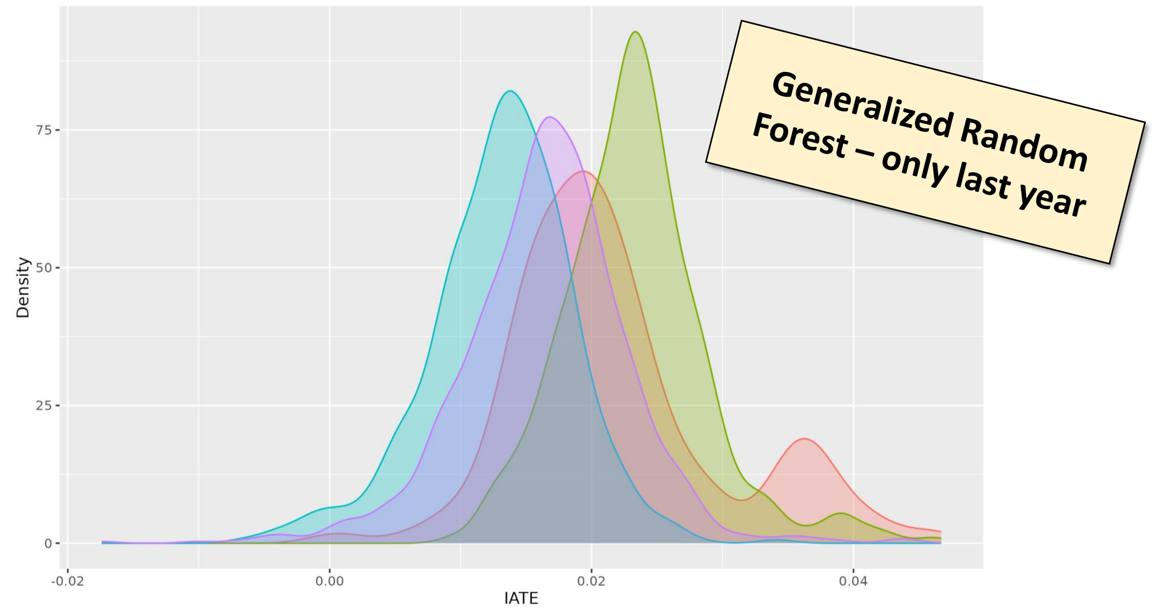
MCF seems to be able to "pick-up" more heterogeneity in IATEs

GRF results are more "intuitive" from the clinical point of view (14% negative IATEs as compared to 60% with mcf)

Standard errors (GRF): 83% (53%) between 0.025 and 0.125 (0.05 and 0.10)



Second modelling step: Application of decision policy and comparison with empirical decisions by physicians



CONSENSUS Policy = Physicians
GRF: No discharge - Physician: No discharge
GRF: Discharge - Physician: Discharge

NO CONSENSUS Policy ≠ Physicians
GRF: No discharge - Physician: Discharge
GRF: Discharge – Physician: Not discharged

Decision Policy: Find IATE for all discharge candidates for each day of our observation period, discharge patients with lowest IATEs

Evaluation: Physicians and GRF agree on discharge decisions for more moderate point estimates, but especially if point estimates are negative, GRF will suggest a different discharge as decided by physicians

DISCUSSION AND NEXT STEPS

Interim milestone: We have made immense progress in the direction of causal clinical decision support! Our application could be a blueprint for similar decision problems in hospitals and beyond. **However, some issues remain**:

Common Support // Model Matrix: (1) Explore areas with little overlap, potentially exclude from second modelling step and welfare estimation, (2) Re-assess data cleaning and feature engineering strategy, potentially affecting overlap

Modelling strategy: IATE model should be learned with 6 out of 7 years of data in the first modelling step and held out year should be used for the second modelling step and welfare estimation

Standard errors // Choice of primary causal forest algorithm: (1) Explore outlier values of standard errors and implement potential improvements in estimation strategy, (2) Compare GRF standard errors with MCF and experiment with MCF propensity score penalty for growing trees.

Welfare estimation: How could the welfare of our decision support model be estimated?

REFERENCES

[1] Bai, J., Fügener, A., Schoenfelder, J., & Brunner, J. O. (2018). Operations research in intensive care unit management: a literature review. Health Care Management Science, 21(1), 1–24. https://doi.org/10.1007/s10729-016-9375-1

[2] Chou YC, Chuang HHC, Chou P, Oliva R (2023) Supervised machine learning for theory building and testing: Opportunities in operations management. Journal of Operations Management 69(4):643–675.

[3] Yılmaz Ö, Son Y, Shang G, Arslan HA (2024) Causal inference under selection on observables in operations management research: Matching methods and synthetic controls. Journal of Operations Management 70(5):831–859.

[4] Chan CW, Farias VF, Bambos N, Escobar GJ (2012) Optimizing Intensive Care Unit Discharge Decisions with Patient Readmissions. Oper Res 60(6):1323–1341. https://doi.org/10.1287/OPRE.1120.1105

[5] Wang G, Li J, Hopp WJ (2021) An Instrumental Variable Forest Approach for Detecting Heterogeneous Treatment Effects in Observational Studies. Manage Sci 68(5):3399–3418.

[6] Wang G, Li J, Hopp WJ (2023) Personalized Healthcare Outcome Analysis of Cardiovascular Surgical Procedures. Manufacturing & Service Operations Management 25(4):1567–1584.

[7] Hopp WJ, Li J, Wang G (2018) Big Data and the Precision Medicine Revolution. Prod Oper Manag 27(9):1647–1664.

[8] Athey S, Tibshirani J, Wager S (2019) Generalized random forests. Ann Stat 47(2):1148–1178.

[9] Lechner M, Mareckova J (2025) Comprehensive Causal Machine Learning. Working Paper. doi: 10.48550/arXiv.2405.10198

[10] Lechner M (2001) Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. Lechner M, Pfeiffer F, eds. Econometric Evaluation of Labour Market Policies (Physica, Heidelberg), 43–58.

University of St.Gallen
School of Medicine

hoch Kantonsspital St.Gallen

Acknowledgements
We thank Patrick Münger from the Cantonal Hospital of St. Gallen for extracting all data necessary for this study. We thank our student assistants David Klug and Andrés Gallardo for their support in data cleaning and descriptive analyses. We thank Jana Mareckova for providing conceptual feedback.

Funding
This study is supported by a Stepping Stone Grant of the University of St. Gallen (project number 2300181).

Conflict of interest
None.

Ethics approval and consent to participate
This study was approved by the Ethical Commission of Eastern Switzerland (project number 2024-00829).

Consent for publication
Not applicable.

Data availability
Anonymized personal health data with restricted access.

Code availability
Available from the corresponding author upon request.