


## RESEARCH LETTER

# Optimising the therapeutic response of statins using real-world evidence and machine learning: Personalised precision dosing recommends lower statin doses for some patients

Andrew Krentz MD<sup>1,2</sup>  | Lisa Fournier MSc<sup>1,2,3</sup> | Thomas Castiglione MSc<sup>1,3</sup> |  
Vasa Curcin PhD<sup>2</sup> | Camil Hamdane MSc<sup>1,2,3</sup> | Tianyi Liu MSc<sup>2</sup> | André Jaun PhD<sup>1,2</sup>

<sup>1</sup>Metadvice, Lausanne, Switzerland

<sup>2</sup>School of Life Course & Population Sciences, King's College London, London, UK

<sup>3</sup>Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

## Correspondence

Andrew Krentz, Metadvice, Route Cantonale 109, 1025 St-Sulpice, Lausanne, Switzerland.

Email: [akrentz@metadvice.com](mailto:akrentz@metadvice.com)

**KEYWORDS:** lipid-lowering therapy, pharmaco-epidemiology, primary care, real-world evidence

## 1 | BACKGROUND

Evidence-based clinical guidelines for lipid modification are based on interventional clinical trials conducted in selected cohorts of patients according to predefined and restricted eligibility criteria.<sup>1</sup> It follows that guidelines that are applicable to most patients, that is, those with similar characteristics to trial participants, might not be ideal for all patients. In the quest to achieve patient-oriented outcomes that are as good as possible, prescribers may elect to intentionally deviate from published guidance.<sup>2</sup> A recent study using machine learning applied to real-world retrospective data from a Northern California health system reported that moderate- or low-intensity statin therapy achieved better surrogate outcomes for a substantial minority of patients compared with high-intensity statins.<sup>3</sup> We tested the hypothesis that patients can be identified from UK primary care electronic health records for whom personalised cholesterol-lowering therapy might be more appropriate than guideline-based prescribing. We also confirmed the portability of our machine learning technology in a separate clinical data set.

## 2 | METHODS

First, we developed a neural network model to reproduce prevailing UK national guidelines for cholesterol lowering, that is, National

Institute for Health and Care Excellence (NICE) CG67,<sup>4</sup> with a pre-specified level of accuracy. A simple feedforward neural network was optimised to minimise the binary cross-entropy with an equal probability over all possible recommendations. Monte Carlo testing against the rule-based outcomes finally achieved 99.7% accuracy in predicting the right therapy and 98.1% accuracy to both predict the right therapy and none of the alternatives, leaving a neural network that evaluates adherence to guidelines with high accuracy. We then applied a transfer learning procedure to refine the clinical knowledge with real-world evidence outcomes recorded in the UK Clinical Practice Research Datalink (CPRD),<sup>5</sup> associating every therapeutic intervention with a non-high-density lipoprotein (non-HDL) cholesterol reduction target. Data were split into 65% for training, 35% for testing/validation. Using artificial intelligence (AI) that combined knowledge from guidelines and real-world evidence, we identified minority 'digital twin' cohorts likely to benefit from individualisation of cholesterol-lowering therapy. A game theory concept known as Shapley values<sup>6</sup> and the kernel SHapley Additive exPlanations approximation<sup>7</sup> provided a measure of similarity to quantify the potential benefit of departing from the NICE guidelines by rejecting the no-benefit hypothesis with a proportion test at  $p = 0.05$  or 95% confidence level. Having established the neural network capabilities using the CPRD data set, an additional validation test studied the portability of the neural network into a clinical setting from South London, comprising 949 therapy decisions.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

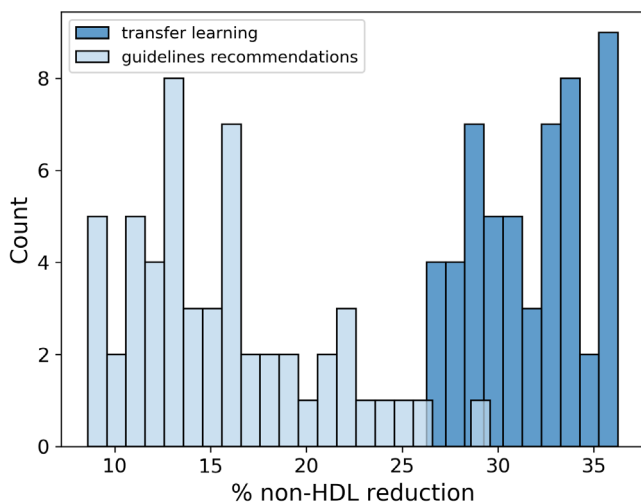
© 2024 The Author(s). *Diabetes, Obesity and Metabolism* published by John Wiley & Sons Ltd.

### 3 | RESULTS

The CPRD sample with complete records who were receiving statin therapy comprised 9675 adult patients (mean  $\pm$  SD age 74  $\pm$  11 years; M 54% vs. F 46%; 86% White or not stated ethnicity with 4% South Indian, 3.3% Black, 2.9% Asian and 1.6% classified as other ethnicities; primary prevention vs. secondary prevention, 65% vs. 35%). Major comorbidities, that is, hypertension (71%) and type 2 diabetes (21%), were similar in prevalence between the primary and secondary prevention cohorts (data not shown).

A broad distribution of responses in the primary outcome of interest, that is, non-HDL cholesterol reduction, was observed, including a majority below the 40% guidance target and even paradoxical increases in some patients (data not shown). Using the median non-HDL reduction observed in CPRD of 25% as an optimisation target, in the CPRD cohort the neural network generated two superposed histograms measuring the average non-HDL cholesterol reduction outcomes for digital twin cohorts from the test data set where the clinician either followed guidelines or happened to choose the same therapy as the neural network recommended (Figure 1). This demonstrates that the clinical outcomes are not evenly distributed in Shapley value space and that the methodology has clear forecasting power.

Learning from real-world outcomes, the model found that for up to 20% of patients, smaller statin doses achieved better lowering of non-HDL cholesterol than doses recommended by the national guidelines. In the portability validation in six South London primary care clinics, all individualised recommendations suggesting a reduction in statin dosage had *p*-values <0.05.



**FIGURE 1** Histogram showing the average non-high-density lipoprotein (non-HDL) cholesterol reduction obtained in the test set for digital twin cohorts optimised for a 25% target, in cases where the prescriber's decision was to administer guidelines (light blue) or the individualised medicine alternative (dark blue).

### 4 | CONCLUSIONS

Our proof-of-concept study, performed in patient samples that are representative of the UK primary care population, supports the contention that machine learning can identify subgroups for whom smaller statin doses deviating from clinical guidelines may be associated with greater degrees of cholesterol lowering. These results, which require further prospective validation, provide clinicians with an actionable basis for a more individualised precision approach to cholesterol-lowering pharmacotherapy. Our findings, based on independently developed and tested hypotheses, echo those of Sarraju et al.<sup>3</sup>

If sustained over time, failure to reduce non-HDL cholesterol levels to evidence-based goals may lead to avoidable cardiovascular events. Although an explanation for better cholesterol lowering using smaller statin doses cannot be determined from our analysis, a plausible mechanism is that adherence to therapy is better reflecting lower rates of statin-associated adverse effects.<sup>8</sup> This is testable in prospective cohort studies. Of potential relevance to this hypothesis, paradoxical increases in non-HDL cholesterol were observed in a proportion of patients consistent with suboptimal adherence to medication.<sup>9</sup> Of note, heterogeneity of therapeutic response is to be expected in our analysis. Cohorts identified as more optimally treated with lower intensity statin regimens may contain individuals who respond differently to a specified statin dose.

The strengths of our study include: first, the debiasing and portability that is achieved when combining two potentially biased sources of information (guidelines, real-world data). Second, direct calculation of statistical support to inform clinical decisions using retrospective data permits an individualised clinical trial to be performed independent of a black-box technology. Potential limitations of the study include well-recognised deficiencies in the completeness of electronic health record coding and the restricted generalisability of the findings to populations outside the UK.

#### ACKNOWLEDGEMENTS

The authors thank Andrew Rut for helpful comments on the manuscript. André Jaun had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. This work is based in part on CPRD data with a protocol (No.: 21\_000346) that has been submitted to CPRD for ethics approval by its Independent Scientific Advisory Committee for research using anonymised patient data. This work was carried out and financially supported through a collaboration between academic entities (KCL, EPFL) and a precision medicine technology company (Metadvice).

#### CONFLICT OF INTEREST STATEMENT

Andrew Krentz and André Jaun are shareholders in Metadvice.

#### PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/dom.16029>.

## DATA AVAILABILITY STATEMENT

CPRD data are commercially available. Specific extracts reflect the contract between King's College London and CPRD.

## ORCID

Andrew Krentz  <https://orcid.org/0000-0003-4169-0469>

## REFERENCES

1. Averitt AJ, Weng C, Ryan P, Perotte A. Translating evidence into practice: eligibility criteria fail to eliminate clinically significant differences between real-world and study populations. *NPJ Digit Med*. 2020;3:67.
2. Qureshi N, Humphries SE, Gray H. Personalised medicine in general practice: the example of raised cholesterol. *Br J Gen Pract*. 2018; 68(667):68-69.
3. Sarraju A, Ward A, Li J, et al. Personalizing cholesterol treatment recommendations for primary cardiovascular disease prevention. *Sci Rep*. 2022;12(1):23.
4. National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification. 2014 Updated 2023. Accessed 1st December 2023. <https://www.nice.org.uk/guidance/cg181>
5. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol*. 2015;44(3): 827-836.
6. Shapley L. *Notes on the End-Person Game—II. The Value of the n-Person Game*. RAND Corporation; 1951.
7. Lundberg SMLS. A unified approach to interpreting model predictions. 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA. 2017.
8. Cheeley MK, Saseen JJ, Agarwala A, et al. NLA scientific statement on statin intolerance: a new definition and key considerations for ASCVD risk reduction in the statin intolerant patient. *J Clin Lipidol*. 2022;16: 361-375. doi:10.1016/j.jacl.2022.05.068
9. Ridker PM, Mora S, Rose L, Group JTS. Percent reduction in LDL cholesterol following high-intensity statin therapy: potential implications for guidelines and for the prescription of emerging lipid-lowering agents. *Eur Heart J*. 2016;37(17):1373-1379.

**How to cite this article:** Krentz A, Fournier L, Castiglione T, et al. Optimising the therapeutic response of statins using real-world evidence and machine learning: Personalised precision dosing recommends lower statin doses for some patients. *Diabetes Obes Metab*. 2025;27(1):432-434. doi:10.1111/dom.16029