



# Early identification of patients likely to benefit from paroxysmal nocturnal hemoglobinuria workup using machine learning on large-scale real-world data

J. Muralidharan<sup>1</sup>, CW. Pike<sup>1</sup>, S. Gombar<sup>1,2</sup>, S. Jain<sup>1,3</sup>, J. Jones<sup>1</sup>

1. Atropos Health Inc., New York City, NY, United States

2. Stanford School of Medicine, Department of Pathology, Stanford, CA, United States

3. Warren Alpert Medical School of Brown University, Division of Hematology/Oncology, Providence, RI, United States



## INTRODUCTION

Paroxysmal Nocturnal Haemoglobinuria (PNH) is a rare (3.81 per 100K), treatable, clonal, hematopoietic stem cell (HSC) disorder characterized by intravascular hemolysis, thrombosis, and smooth muscle dystonias, with bone marrow failure occurring in some cases. Patients undergo lengthy diagnostic journeys, frequently exceeding a year. Diagnosis is often made following a high morbidity/mortality event, such as a stroke. Earlier identification and treatment may improve disease burden.

## AIM

Building on prior work by Worker *et al.* (2024)<sup>1</sup>, we sought to **replicate and extend predictive modeling for earlier PNH identification**, using a large, real-world dataset to pinpoint patients likely to benefit from PNH workup **3–12 months before diagnosis**.

## METHOD

- Data: Sourced from the Apollo Database as part of the Atropos Health Evidence Network (67 million U.S. patients; 1,208 with PNH).
- Analysis: 416K adults with hemoglobin testing (2016–2023) and no prior PNH diagnosis.
- Cases: ICD-10 D59.5 diagnosis 3–12 months post-index Hb (n = 306).
- Controls: Random sample of Hb-tested patients.
- Clinical Features: ≤24 months pre-index.
- An XGBoost model (train/test = 70/30) was optimized for precision-recall performance, clinical interpretability, and deployability, followed by manual physician review of high-discordance predictions.

## RESULTS

We identified 306 PNH patients (0.07% of sample) with an initial diagnosis of PNH 3 to 12 months after a hemoglobin test. Important features included: history of anemia (aplastic and broadly defined), overall disease burden, age, haptoglobin, and having a diagnosis for myelodysplastic syndrome (MDS) (Figure 1). AUPRC and AUROC were 0.09 and 0.77, respectively, for a full model with 375 features, and 0.08 and 0.77 for a model with 13 features (Figure 2).

Clinical review agreed that 100% of “unlikely” PNH patients did not warrant further workup. Among the “false positives” (>34% probability of PNH, but no PNH diagnosis found in 3 to 12 months), two physicians indicated 75% to 90% of these patients still warranted further workup.

The predictive model is available at <https://github.com/atroposhealth/pnh-undiagnosed>

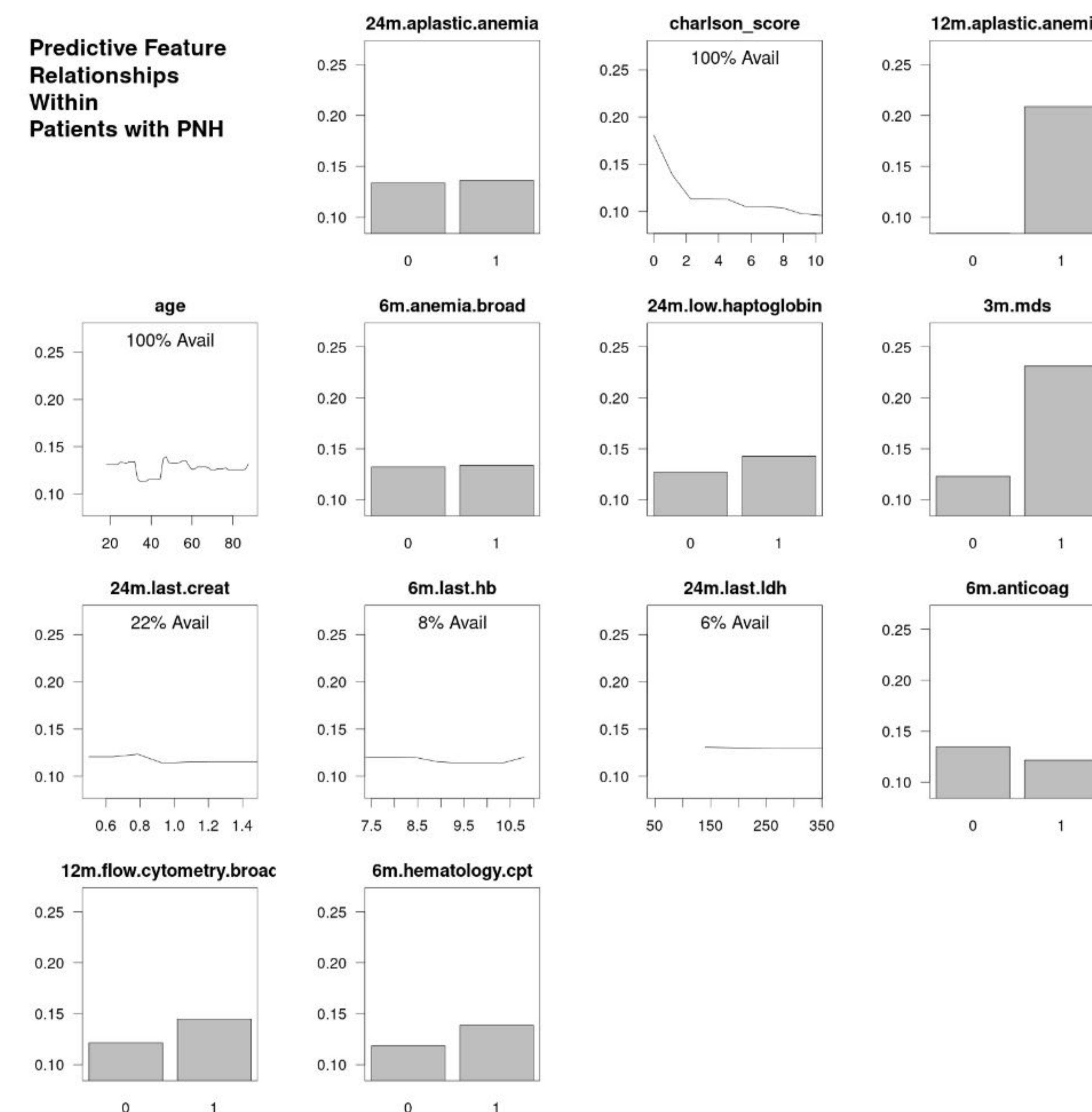


Figure 1: Partial dependence plots for patients with a PNH diagnosis 3 to 12 months after a hemoglobin lab test. As expected, anemia, low haptoglobin, having flow cytometry, and having hematology procedures were positively associated with PNH diagnosis. Increasing disease burden (Charlson Score) was negatively associated with PNH diagnosis.

## CONCLUSIONS

This machine learning algorithm performs sufficiently to be deployed, offering the potential to reduce diagnostic delays for patients with PNH. Setting the threshold for further workup should be tested within a local environment due to variation in lab testing (i.e., Hb) patterns.

## REFERENCES

Worker A, Mahon H, Sams J, Boardman-Pretty F, Marchini E, Dubis R, et al. A machine learning algorithm for the detection of paroxysmal nocturnal haemoglobinuria (PNH) in UK primary care electronic health records. *Orphanet J Rare Dis.* 2024 Oct 13;19(1):378. doi:10.1186/s13023-024-03406-4. PMID: 39396996; PMCID: PMC11479535.

## ACKNOWLEDGEMENTS

This study was conducted in collaboration with and supported by Novartis Precision Health.

## CONTACT INFORMATION

Saurabh Gombar, MD, PhD - saurabh@atroposhealth.com

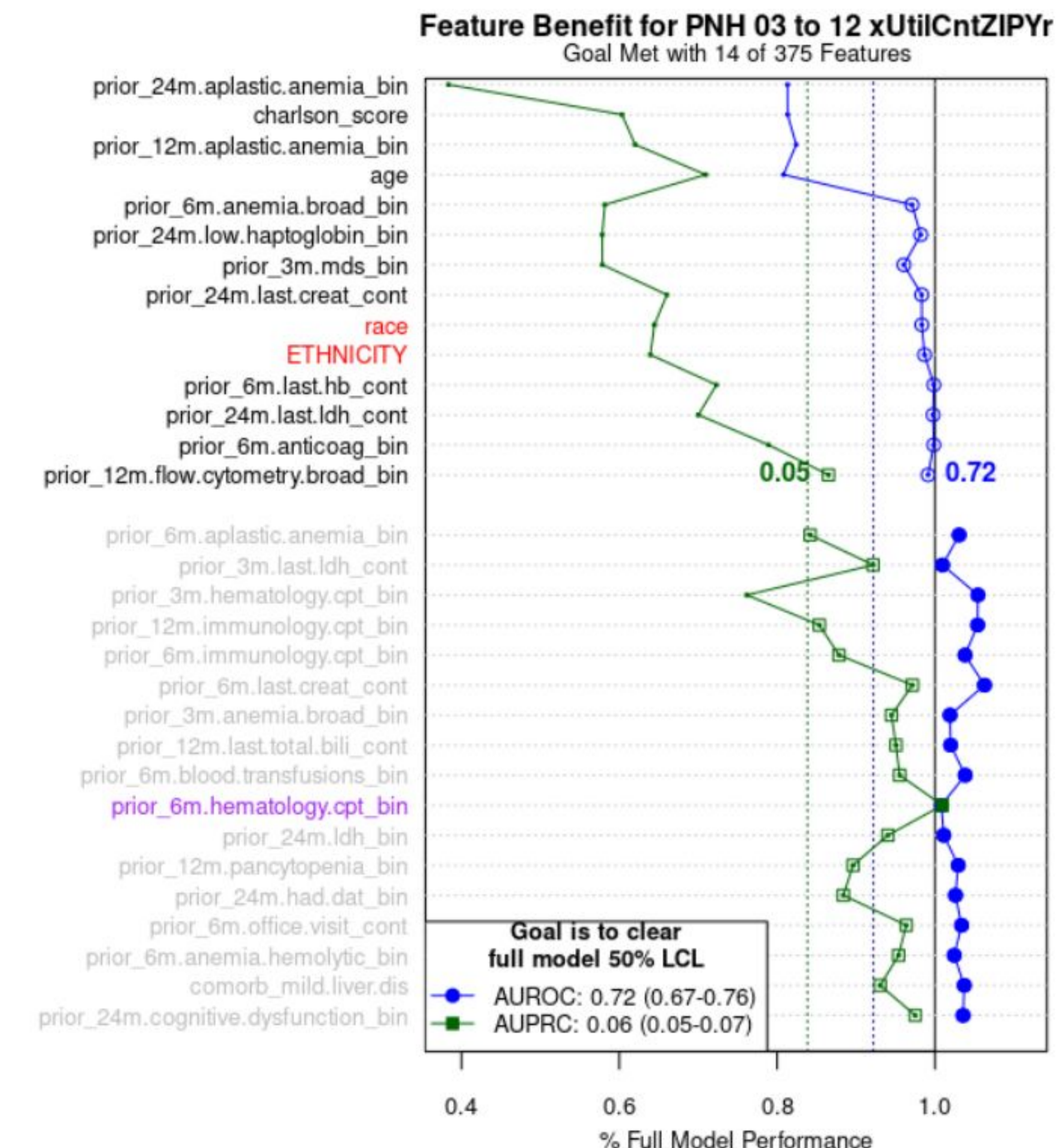


Figure 2: Features included in the maximally efficient PNH identification model are indicated in black (auto-selected) and purple (manually included) text. Manually removed features are in red. Gray features were not included, but were potential high contributors. The horizontal axis indicates the performance of the model relative to the full, 375-feature model for area under the precision recall (green) and receiver operating characteristic curves (blue).