XXIX IFSO World Congress
1-4 September 2026 | Toronto, Canada

Prediction of Weight Loss and Comorbidity Resolution After Bariatric Surgery Using Machine Learning with the UK National Bariatric Surgery Registry

Mr Shivam Bhanderi, Dr Swarna Kempe Gowda, Dr Adithya Srinivas,
Prof Omar Khan, Mr Chandra Cheruvu

Disclosures

- No disclosures

Acknowledgements

- Dendrite Clinical Systems

Background

- Outcomes from bariatric surgery vary widely
- MDT decisions rely on clinical experience
- The NHS is different compared to other nations in patient selection and aims of surgery
- Growing interest in AI models and outcome prediction across surgery[1]
  - SOPHIA prediction tool[2,3]

1. Vitiello et al. Artificial intelligence and machine learning in bariatric surgery: a comprehensive systematic review. Langenbecks Arch Surg 2026. doi: 10.1007/s00423-026-04042-1
2. Bariatric Weight Trajectory Prediction. https://bariatric-weight-trajectory-prediction.univ-lille.fr
3. Saux P, Bauvin P et al. Development and validation of an interpretable machine learning-based calculator for predicting 5-year weight trajectories after bariatric surgery: a multinational retrospective cohort SOPHIA study. Lancet Digit Health 2023; 5(10):e692-e702.

NBSR

- National Bariatric Surgery Registry[1,2]
- Entries expected for all NHS/private primary and revisional bariatric procedures
- Annual report published in collaboration with BOMSS[3]

1. National Bariatric Surgery Registry. https://nbsr.e-dendrite.com
2. Miras AD, Kamocka A, Patel D, et al. Obesity surgery makes patients healthier and more functional: real world results from the United Kingdom national bariatric surgery registry. Surg Obes Relat Dis. 2018;14(7):1033–40
3. British Obesity & Metabolic Specialist Society. https://bomss.org

What is Machine Learning?

- Machine Learning (ML):
  - A way for computers to learn patterns of data and make predictions without being explicitly programmed.
- Potential in healthcare to:
  - Identify trends
  - Predict outcomes
  - Support clinical decision making

XGBoost

- E**X**treme **G**radient **Boost**ing
- Starts with a decision tree and tests it
- Evaluate the error
- Build a new decision tree that refines the error
- Repeats based on starting parameters specified, including:
  - Tree complexity
  - Learning rate
  - Predictors to use and outcome to predict

1. XGBoost. https://xgboost.ai

Study Aim

To assess the performance of a machine learning model to predict %EWL, T2DM resolution and OSA resolution trained on data from the UK NBSR.

Methods: Data Extraction

- Relevant registry entries between 2018-2023
  - More modern era of bariatric surgery
  - Avoid time period where registry in infancy
  - Allow time for follow up data to be available
- Primary bariatric surgery, excluding endoscopic interventions
- Follow up data taken from closest to 12m post op (9-15 months)
- Exclusion:
  - Suspected duplicate registry entries (matching DOB, gender, operation)
  - No operation data available
  - No follow up data available

Methods: Data Validation

- Surveillance for erroneous data values:
  - 15y < Age < 90y
  - 120cm < Height < 215cm
  - 50 < Weight < 325kg
  - 15 < BMI < 200
  - -50% < %WL < 100
  - -100% < %EWL < 300%
- Reproducible recoding of registry variables for analysability
- Direct calculation of height, weight, BMI, %WL, %EWL based on available data
- Machine learning with XGBoost using R 4.3.3

Methods: ML Model

- 80% training, 20% test
- 5-fold cross validation
- Hyperparameter tuning using grid search
- Early stopping to prevent overfitting
- Separate models per outcome

Results: Data Extraction

- 49,812 unique patient identifiers
- 1375 suspected duplicates excluded
- 19760 entries excluded as no follow up data available
- 14657 entries excluded as no follow up data within 9-15months
- 244 entries excluded after erroneous value screening

- 13,976 eligible primary bariatric operations with reported outcome data 9-15months

Results: Demographics

Total number of patients/operations	13976
Mean Age (SD)	44.79 (11.28)
SG	6216
RYGB	5421
OAGB	1418
Band	920
DS/SADI/BPD	1
Mean Pre-op BMI (SD)	44.54 (7.43)
Pre-op T2DM	3381
Pre-op OSA	2746
Median F/U months (IQR)	11.99 (11.4-12.71)
Mean %WL (SD)	27.26 (10.07)
Mean %EWL (SD)	66.18 (26.75)
Mean F/U BMI (SD)	32.23 (6.42)
F/U T2DM	1069
F/U Persistent OSA	1495

Results: %EWL Prediction

Number of patients	11709
Mean Pre-op BMI (SD)	44.56 (7.31)
Pre-op T2DM	2916
Pre-op OSA	2316
Median F/U months (IQR)	12.02 (11.33-12.78)
Mean %WL (SD)	27.28 (10.04)
Mean F/U BMI (SD)	32.24 (6.32)
F/U T2DM	996
F/U Persistent OSA	1156

Top predictors:

Pre-op BMI

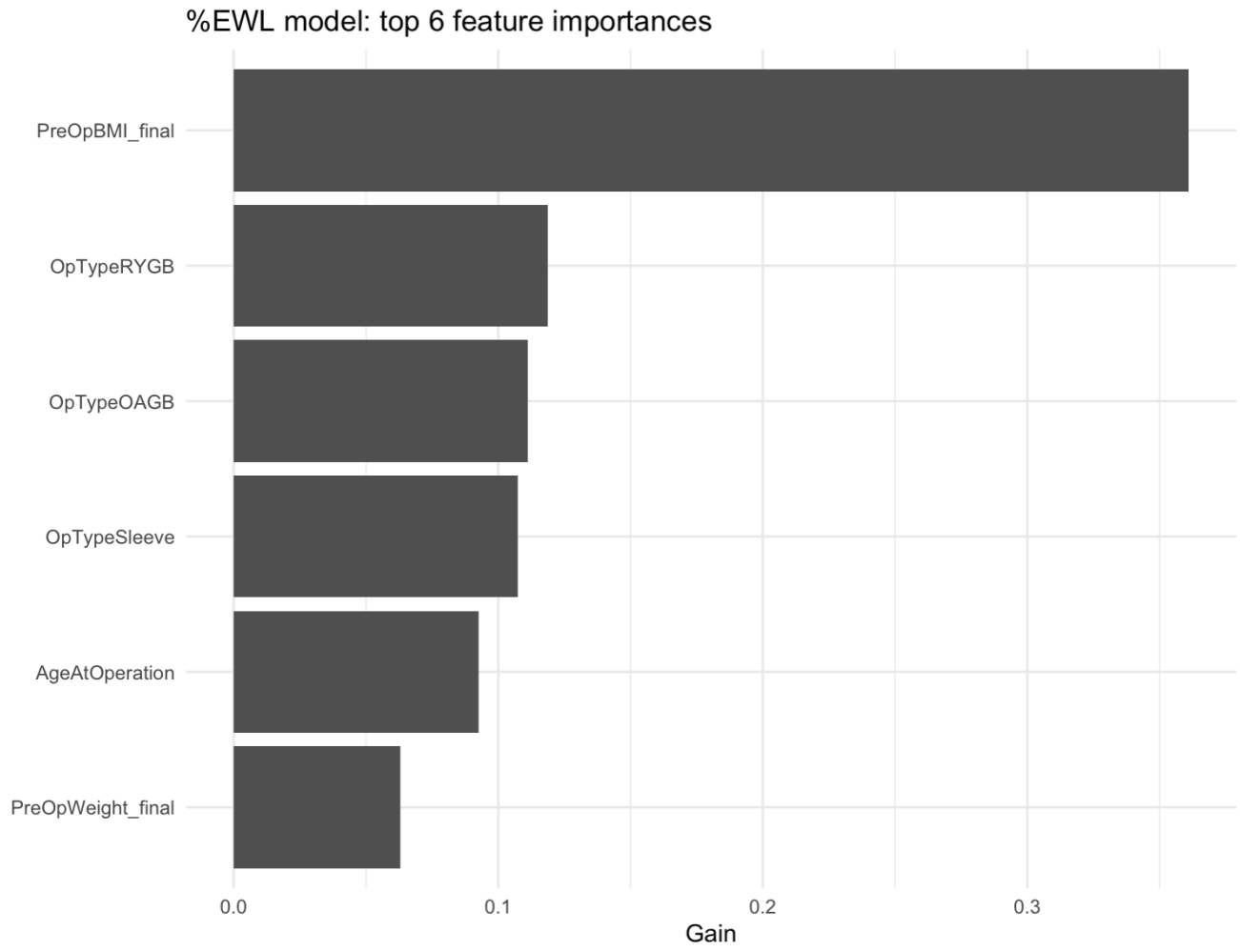
Operation Type

Age

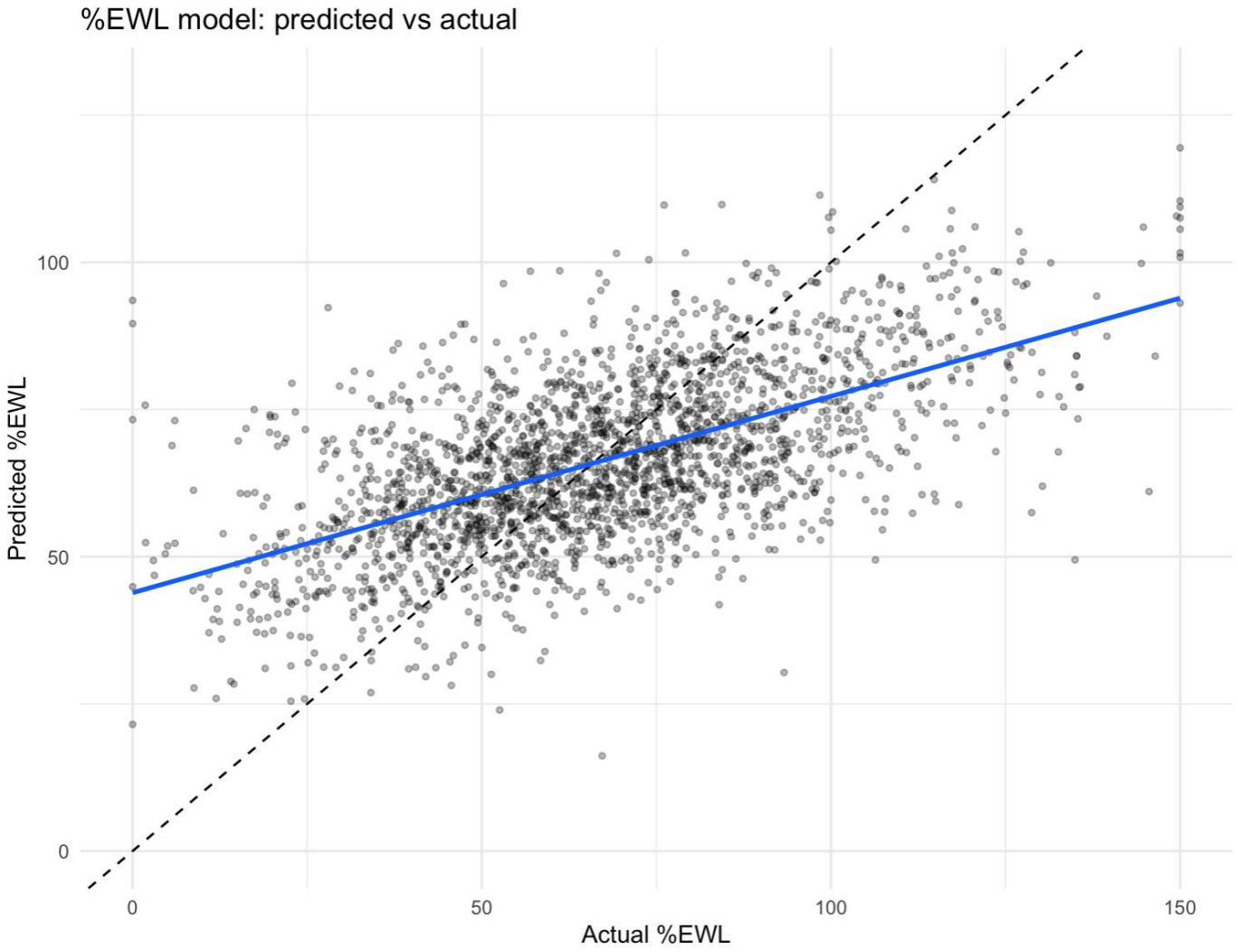
Diabetes severity

Model Performance	
RMSE	21.57
MAE	16.54
R^2	0.348

Results: %EWL Prediction



Results: %EWL Prediction



Results: T2DM prediction

Number of patients	3378
Mean Pre-op BMI (SD)	44.01 (7.36)
Pre-op T2DM	3378
Pre-op OSA	1094
Median F/U months (IQR)	12.02 (11.43-12.78)
Mean %WL (SD)	24.52 (10.25)
Mean %EWL (SD)	60.65 (27.03)
Mean F/U BMI (SD)	32.98 (6.36)
F/U T2DM	986
F/U Persistent OSA	611

Top predictors:

Pre-op T2DM Severity

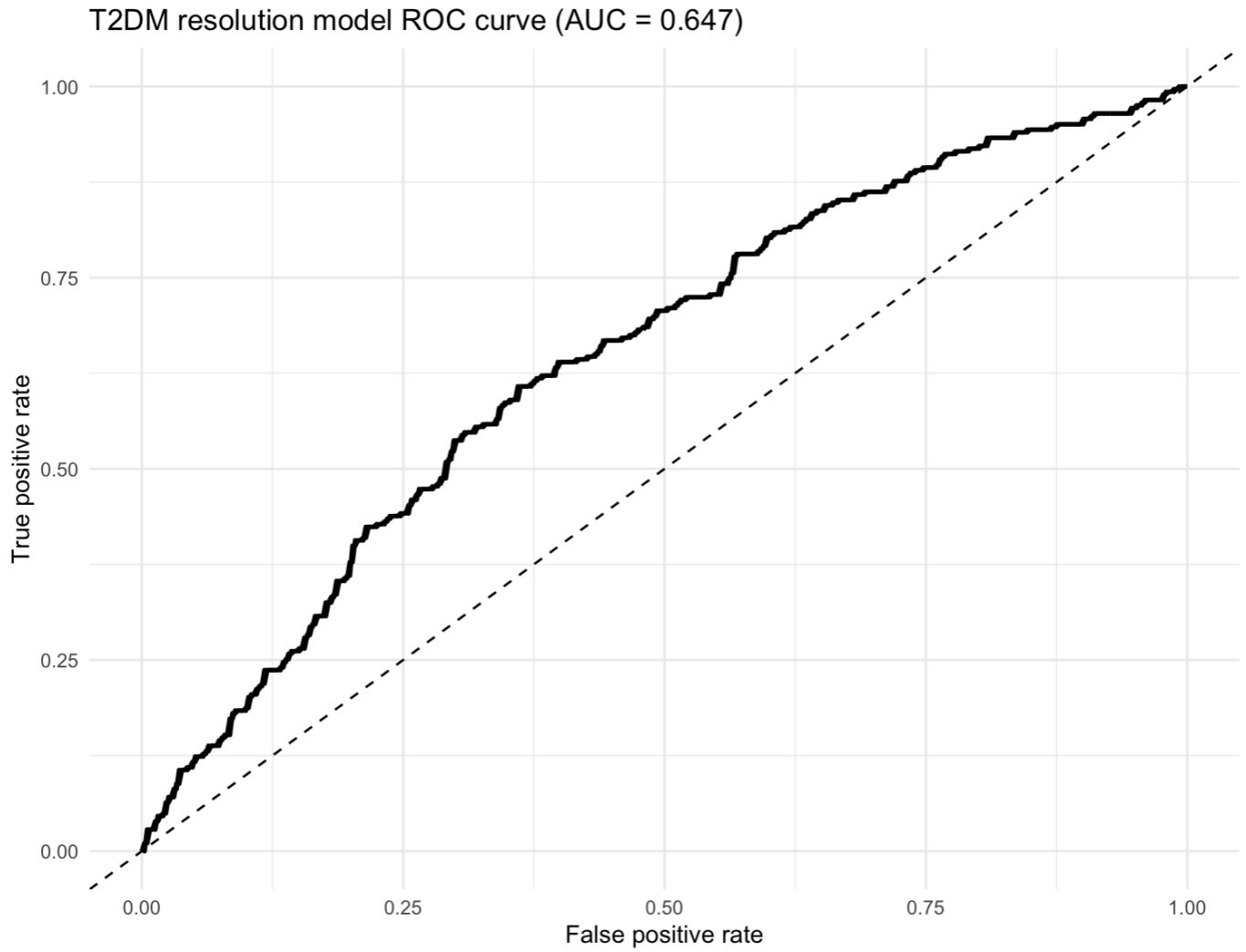
Pre-op HbA1c

Pre-op BMI

Operation Type

Model Performance	
AUC	0.647
Accuracy	0.584
Sensitivity	0.721
Specificity	0.485
Precision	0.502

Results: T2DM Prediction



Results: OSA prediction

Number of patients	2744
Mean Pre-op BMI (SD)	47.07 (8.25)
Pre-op T2DM	1094
Pre-op OSA	2744
Median F/U months (IQR)	12.02 (11.4-12.78)
Mean %WL (SD)	26.07 (10.02)
Mean F/U BMI (SD)	34.48 (7.21)
F/U T2DM	356
F/U OSA	1495

<u>Top predictors:</u>
Pre-op BMI
Functional Status
Age

Model Performance	
AUC	0.589
Accuracy	0.551
Sensitivity	0.617
Specificity	0.498
Precision	0.495

Overall Model Performance

- Machine learning is finding a credible signal in NBSR data
  - Strongest signal is for predicting %EWL
  - T2DM remission is partly predictable
  - OSA resolution is poorly predictable
- Most important predictors are clinically plausible
- No signs in model performance metrics of gross overfitting

Conclusions & Limitations

- Voluntary registry => relative lack of follow up data
- ML models capture predictive signals for %EWL
- Predicting human behaviour / Missing model features
- Registry design to facilitate easier evaluation of ML models
- Strengths:
  - Specific to UK's NHS
  - No gross overfitting detected
- Limitations
  - Lack of follow up data => inherent selection bias for more complete records
- Future work: adapting the registry design, adding model features

THANK YOU

XXX IFSO WORLD CONGRESS

"Advancing Multimodal Obesity Care"

24-27 August 2027

ifso2027.org

IFSO PARIS 2027

See you in

PARIS