Machine learning algorithm to predict >5% weight gain in PWH switching to INSTI

Guaraldi G[1,2], Motta F[3], Milic J[1], Barbieri S[3], Gozzi L[2], Aprile E[2], Belli M[2], Venuta M[2], Cuomo G[2], Carli F[2], Dolci G[2], Iadisernia V[2], Burastero G[2], Mussini C[1,2], Mandreoli F[3]

[1]Department of Surgical, Medical, Dental and Morphological Sciences, University of Modena and Reggio Emilia, Italy; [2] Department of Infectious Diseases, Azienda Ospedaliero-Universitaria Policlinico of Modena, Modena, Italy; [3] Department of Physical, Computer and Mathematical Sciences, University of Modena and Reggio Emilia, Italy

597

Background

Weight gain (WG) is a well-described phenomenon in PWH starting or switching ART. Machine learning (ML) methods are suitable tools to assess P4 medicine (Predictive, Preventive, Personalized & Participatory) and can generate models able to identify patients at risk of WG.

The objective of the study was to develop a ML algorithm that predicts percentage weight gain (WG%) in a given interval of time in PWH switching antiretroviral therapy (ART) including integrase inhibitors (INSTI) with/without TAF.

Methods

This was an observational study that comprised ART-experienced PWH attending Modena HIV metabolic clinic (MHMC) from 2001 to 2021. Data were routinely assessed at each patient visit (Figure 1).

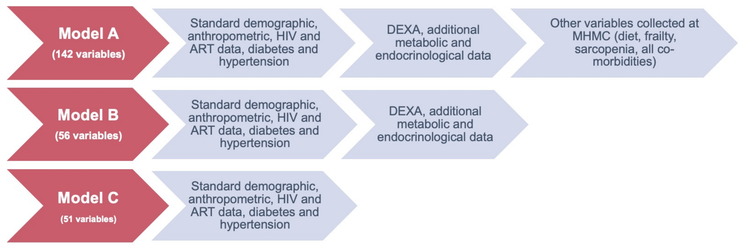


Figure 1. *Variables with availability ≥80% were considered as valuable observations and accounted for 142 variables used in model A. Two additional models were trained in order to test ML performance on parsimonious datasets available in tertiary level clinics (model B) and standard HIV out-patient clinics (model C).*

Data were partitioned in an 80/20 training/test set to generate predictive models. The study outcome was the prediction of weight change % at any given follow-up. A clinically meaningful WG cut-off was set at 5%, at the following annual visit.

A robust implementation of linear regressor algorithms was able to predict weight gain/loss while tolerating missing data. Intelligible explanations were obtained through Shapley Additive exPlanations values (SHAP), which quantified the positive or negative impact of each variable included in the model on the predicted outcome.

Results and discussion

A total of **3516 patients generated 18874 observations** and three predictive models with different sets of variables were trained (models A, B and C; Figure 1; Table 2).

At last observation, median age was 50 years; 70% were male. Median nadir CD4 was 194 cells/μL, current CD4 was 645 cells/μL. 93% had undetectable HIV RNA and time since HIV diagnosis was 19.7 years. Median BMI was 23.5 and 7.5% had obesity.

Machine learning (ML) algorithm built from a rich universe of variables had an excellent performance (>90%) in predicting WG in terms of accuracy, sensitivity, specificity and ROC AUC **(models A and B, Figure 1, Table 1).**

ML algorithm restricted to routinely collected variables in HIV clinics, although highly specific, is insufficient to reliably predict WG **(model C, Figure 1, Table 1).**

The contribution of INSTI with/without TAF had limited impact on WG prediction in comparison to metabolic, anthropometric and lifestyles variables **(Figure 2).**

The following ART switch observations were registered in the dataset: from TDF to TAF - 304; from non-INSTI to INSTI without TAF - 3656; from non INSTI to INSTI with TAF – 293; from EFV to INSTI – 3128.

Table 1 *depicts performance metrics of the ML model with regards to the 5% WG threshold. Out of 3776 observations in the test set (16.7% with WG≥5%), 596 correctly predicted with WG≥5% (true positive – TP)), 35 overestimated WG≥5% (false positive – FP), 35 underestimated WG<5% (false negative – FN); and 3110 correctly predicted with WG<5% (true negative – TN).*

Metric name	Meaning	Definition	Model A	Model B	Model C
Sensitivity (recall, true positive rate)	Ability to correctly identify patients **with WG** who **do** experience WG	TP/(TP+FN)	94.5%	93.0%	10.0%
Specificity (selectivity, true negative rate)	Ability to correctly identify patients **without WG** who **do not** experience WG	TN/(FP+TN)	98.9%	98.6%	96.1%
Positive predictive value (precision)	Probability of **experiencing WG** if the model classifies a patient as **"positive"**	TP/ (TP+FP)	94.5%	93.0%	33.9%
Negative predictive value	Probability of **not experiencing WG** if the model classifies a patient as **"negative"**	TN / (FN+TN)	98.9%	98.6%	84.2%
Accuracy	Closeness of the predicted value to a observed value	(TP+TN)/(TP+FN+FP+TN)	98.1%	97.7%	81.7%
$F_{\beta=0.5}$	Measure of accuracy, taking into an account sensitivity and positive predictive value	$((1+\beta^2)\,TP)/((1+\beta^2)\,TP+\beta^2\,FN+FP)$	94.5%	93.0%	22.9%
ROC AUC	Ability to identify a patient with or without WG with proposed ML models	$\int TPR\; d\,(FPR)$	96.7%	95.8%	53.0%

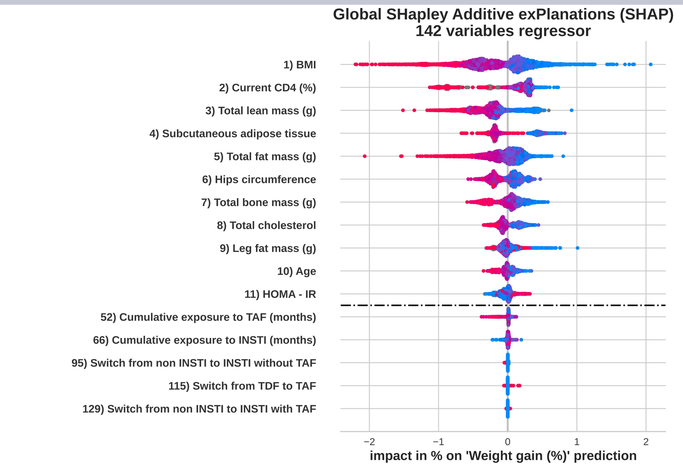
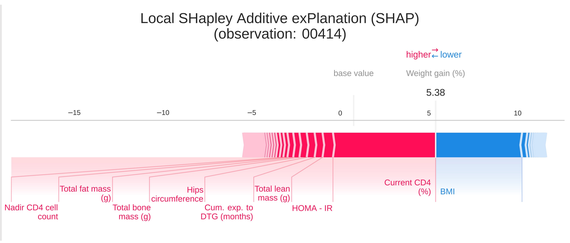


Figure 2. *The global SHAP values for the 11 top variables and ART regimens (out of 142 included in the model A). Values of each variable may have a positive or negative impact, depending on their SHAP value. For instance, high values of BMI in red negatively contributes to WG, while low values in blue contribute strongly to WG.* ***Disclaimer****: SHAP values are not odds ratios with confidence intervals and this figure only qualitatively ranks by importance each variable included in the model. These ML algorithms do not imply causation.*

Figure 3 *depicts positive and negative impact of each variable in a patient with >5% WG.*



Conclusion

- ML models A and B had a remarkable performance in WG prediction, thank to the inclusion of body composition metabolic and endocrinological variables. These models should be deployed in electronic patient charts at metabolic clinic referral centers.
- The parsimonious model (model C) with restricted subset of anthropometric HIV and ART variables, available in standard clinical evaluation, is insufficient to obtain reliable prediction.
- These models stress the multifactorial nature of WG in which the impact of INSTI or/and TAF switch/exposure is diluted in the universe of variables which contribute to WG. The major drivers of WG are modifiable risk factors related to lifestyles.

CROI Conference on Retroviruses and Opportunistic Infections

Corresponding author:
Giovanni Guaraldi, MD - giovanni.guaraldi@unimore.it

Funding
This study is supported by a Gilead grant.

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA