Prediction of Antipsychotic Associated Weight Gain in Children and Adolescents Taking Second Generation Antipsychotics: A Machine Learning Approach

MSR98

College of Pharmacy
UNIVERSITY OF HOUSTON

Ning Lyu[1], Ying Lin[2], Paul J Rowan[3] , Susan Abughosh[1], Tyler J Varisco[1], Hua Chen[1]

1. University of Houston College of Pharmacy, TX, USA 2. University of Houston College of Engineering, TX, USA 3. The University of Texas at Houston School of Public, TX, USA

Contact Information:
Ning Lyu, Ph.D.
University of Houston
Email: nlyu@uh.edu

BACKGROUND

Second Generation Antipsychotics (SGA) are associated with serious cardiometabolic side effects in children and adolescents, especially antipsychotic-associated weight gain (AAWG). However, no study has focused on real-time AAWG prediction among children and adolescents.

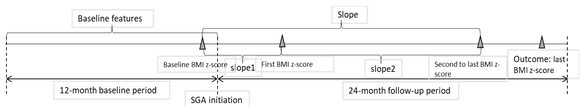
OBJECTIVE

This study aims to develop a prognostic-based machine learning algorithm that would dynamically predict the real-time risk of antipsychotic-associated weight gain (AAWG).

METHODS

Study design

- Using an 80% random sample of the study cohort to predict the last weight measure recorded during up to 24 months of the SGA treatment.
- The potential features were extracted from the 12-month baseline and 24-month follow-up period.
- The best-performed model was identified by comparing these models' Area Under the Curve (AUC) and other evaluation metrics using a 20% random sample of the study cohort.



Slope= (Second to last BMI z-score - baseline BMI z-score)/Duration between them
Slope1= (First BMI z-score - baseline BMI z-score)/Duration between them
Slope2= (Second to last BMI z-score - First BMI z-score)/Duration between them
Acceleration= Slope2 - Slope1

Study cohort

- Aged 6 to 19 years at the SGA initiation.
- Continuously prescribed SGA for a minimum of 90 days.
- Being active during the 12 months before SGA initiation.
- Possessed one BMI z-score measure within the 30-day window before SGA initiation and at least two BMI z-score measures during an up to 24-month follow-up period since the SGA initiation.

METHODS

Data source

- IQVIA Ambulatory EMR- US database between 2016 and 2021

Study outcome

- Severe weight gain: ≥0.5 BMI z-score increase
- Moderate weight gain: ≥0.25 and <0.5 BMI z-score increase
- Minor weight change: <0.25 BMI z-score increase

Statistical analysis

- Multiclass Logistic Regression (MLR) model
- Multiclass Classification and Regression Trees (CART) model
- Multiclass Random Forest (MRF) model
- Multiclass Vector Generalized Additive Model (VGAM) model
- Multiclass Extreme Gradient Boosting (Xgboost) model
- Sensitivity analysis was performed with more intensive data points (>= 3 BMI z-scores during the follow-up period.

RESULTS

- In **Figure 1**, a total of 10,997 patients who met the eligibility criteria were identified, of which 80% (8,798) of the SGA recipients were designated for training data, and the remaining 20% (2,199) were earmarked for testing purposes.
- In **Table 1**, the study cohort exhibited a mean age of 11.7 years (SD: 3.3), with 59% (n=6,522) being male and 67% (n=7,404) being Non-Hispanic White, and 50% (n=5,543) from the Southern region of the United States.
- In **Table 2**, the Xgboost model exhibited the highest AUC (0.921), accuracy score (0.814), F1 score (0.798), and overall sensitivity (0.665) and specificity (0.873) among the five multiclass predictive models.
- In **Table 3**, Important features identified for AAWG prediction include BMI z-score slope, baseline BMI z-score, SGA duration, duration between last measure and outcome, and counts of BMI z-score.

RESULTS

Figure 1. Schematic diagram

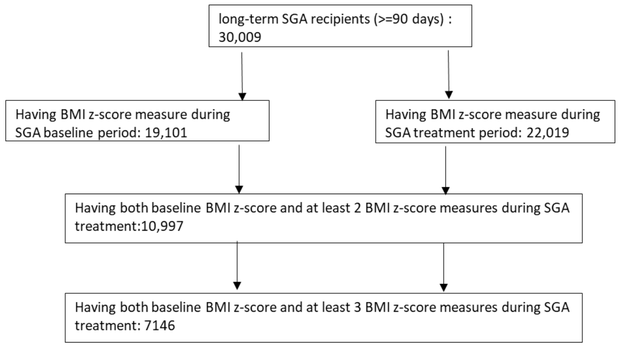


Table 1. Characteristics of weight gain groups

	Minor weight gain	Moderate weight gain	Severe weight gain	Chi-square
	(N= 7041)	(N=1134)	(N=2822)	
	N (%)	N (%)	N (%)	P value
Sex				<0.001
Male	4059 (57.7)	637 (56.2)	1826 (64.7)	
Female	2972 (42.2)	497 (43.8)	995 (35.2)	
Race				0.023
Non-Hispanic White	4657 (66.1)	766 (67.6)	1981 (70.2)	
Non-Hispanic Black	780 (11.1)	111 (9.8)	272 (9.6)	
Hispanic	35 (0.5)	8 (0.7)	9 (0.3)	
Asian	55 (0.8)	7 (0.6)	24 (0.9)	
Others	308 (4.4)	56 (4.9)	119 (4.2)	
Region				0.456
South	3587 (50.9)	541 (47.7)	1415 (50.1)	
Midwest	1859 (26.4)	304 (26.8)	748 (26.5)	
Northeast	882 (12.5)	158 (13.9)	367 (13.0)	
West	713 (10.1)	131 (11.6)	292 (10.4)	

Table 2. Performance evaluation

At least 2 BMI z-scores at follow-up period (main analysis)	MLR model	CART model	MRF model	VGAM model	Xgboost model
AUC	0.826	0.853	0.851	0.848	0.921
Accuracy	0.75	0.798	0.794	0.776	0.814
F1 score	0.787	0.783	0.755	0.748	0.798
Sensitivity	0.52	0.64	0.579	0.573	0.655
Specificity	0.809	0.872	0.85	0.845	0.873
At least 3 BMI z-scores at follow-up period (Sensitivity analysis)					
AUC	0.884	0.899	0.895	0.894	0.946
Accuracy	0.797	0.842	0.815	0.809	0.851
F1 score	0.839	0.823	0.778	0.789	0.841
Sensitivity	0.569	0.673	0.606	0.627	0.716
Specificity	0.849	0.894	0.873	0.876	0.899

Table 3. Important features for AAWG

	Feature	Gain
1	BMI z-score Slope	0.6330
2	Baseline BMI z-score	0.1997
3	SGA duration	0.0606
4	Duration between Last measure to outcome	0.0481
5	Counts of BMI z-score at follow-up period	0.0151

CLINICAL INSIGHTS

- Accurate real-time prediction of AAWG holds significant clinical importance in identifying high-risk patients and tailoring personalized monitoring schedules and timely interventions in pediatric recipients of Second-Generation Antipsychotics (SGA).

LIMITATIONS

- Our evaluation of model performance relied on an internal validation dataset.
- It's essential to acknowledge the absence of several sociodemographic, clinical, and behavioral attributes within data. Variables such as insurance status, dietary patterns, and genetic factors, which can contribute to weight gain, were not available.

CONCLUSIONS

- The Xgboost model developed in this study holds promise for accurately predicting AAWG in children and adolescents undergoing treatment with SGA.
- The precision of AAWG prediction may be further improved in future endeavors by incorporating more intensive data points into the analysis.