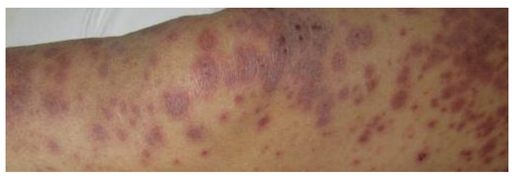
Development of a machine learning polyvariant risk prediction model for severe cutaneous adverse drug reactions to carbamazepine and other aromatic antiseizure medications

Robert Zeibich[1], Alison Anderson[1,2,*], Zhibin Chen[1,2], Yi-Wu Shi[3], Ching-Ching Ng[5], Larry Baum[4], Stacey Cherny[4,6], Pak C. Sham[4], Kheng-Seang Lim[7], Wei-Ping Liao[3], Terence J. O'Brien[1,2,8], Piero Perucca[1,2,8,9,10], Patrick Kwan[1,2,8,11,*]

Introduction

- Carbamazepine (CBZ) is effective in epilepsy as well as pain (e.g., trigeminal neuralgia) and bipolar disease.
- Stevens-Johnson syndrome (SJS) and toxic epidermal necrolysis (TEN) are rare but life-threatening adverse reactions to CBZ treatment.

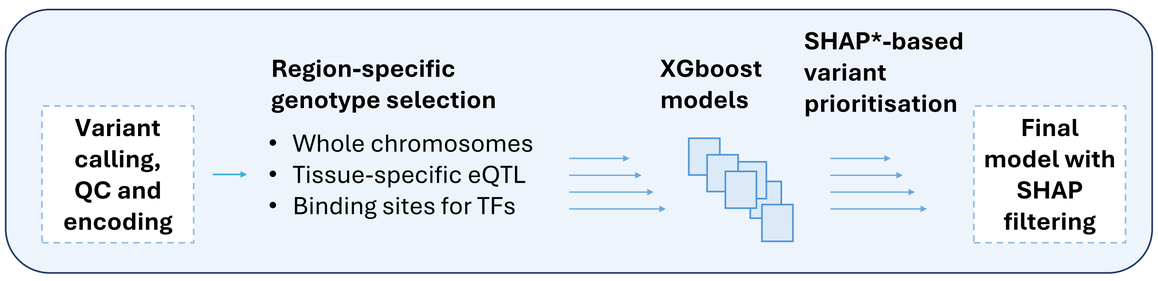


- In Asians, the *HLA-B*15:02* allele is associated with the risk of CBZ-induced SJS/TEN, however, the positive predictive value (PPV) is only 0.07[1], indicating that other genetic variants may modify the risk.

Objective: Identify additional risk-modifying variants.

Methods

Polyvariant risk identification workflow



*SHapley Additive exPlanations

Data: Whole genome data was available from Han Chinese individuals with CBZ-induced SJS/TEN (79 carriers and 33 non- carriers) or CBZ-tolerance (28 carriers, 52 non-carriers).

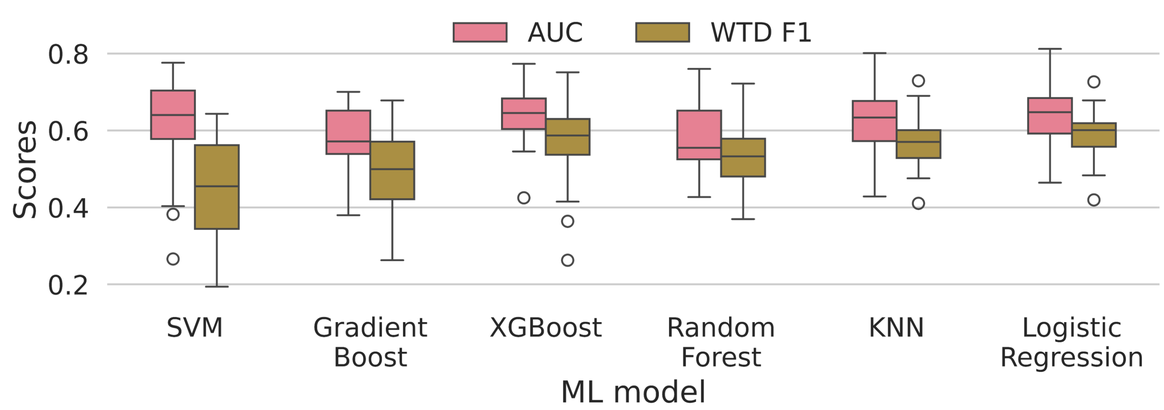
Step 1: Best predictive model for known risk regions
Six machine learning (ML) models were evaluated to determine their capacity to identify the best Area Under the Curve (AUC) and weighted F1 score (WTD F1) for the known risk region (chromosome six, containing *HLA-B*).

Step 2: Systematic screening for additional risk regions
Used the best model to systematically investigate 1,390 regions of the DNA:

- **Whole chromosomes** (to compare with chromosome six)
- **Transcription Factor Binding Sites (TFBSs)**
  - TFs bind to DNA and regulate gene expression, which is modified by genetic variants.
- **Expression Quantitative Trait Loci (eQTL) variants** for 17 tissues
- ***HLA*-regions** (eQTL variants, TFBSs specific to *HLA-B/HLA-A*, coding regions of *HLA-B*/*HLA-A*)

Step 3: Evaluated the performance of putative predictive variants
Compared the estimated PPV from the polyvariant model with that obtained from the *HLA-B*15:02* carrier status

Step 1 results: Best predictive model



XGBoost performed best with a median AUC of 0.71 and a median weighted F1 score of 0.65.

Step 2 results: Top 10 risk modifying regions

Region group	Variants	AUC	WTD F1
TF: ZNF366	38,145	0.73	0.65
TF: JUND	35,020	0.73	0.64
TF loci within 5,000 bp before and after *HLA-B*	344	0.72	0.65
Skin-Not-Sun-Exposed-Suprapubic	950,142	0.72	0.65
TF: BACH2	3,494	0.72	0.63
Colon-Sigmoid	603,214	0.72	0.63
TF: HEYL	2,145	0.71	0.66
Small-Intestine-Terminal-Ileum	300,969	0.71	0.66
Pituitary	508,964	0.71	0.66
Variants within protein-coding region of *HLA-B*	361	0.71	0.65

Step 3 results: Final model

- Identified 465 putative variants.
  - The model increased the PPV from 0.07 to 0.20.
  - The model achieved an AUC of 0.93.
- Risk beyond chromosome six/*HLA* regions
- 115 of 465 variants were classified as highly confident based on SHAP analysis.
  - 60% of these identified in TFBSs
  - Four variants modified the binding of BACH2, which is known for regulating CD8+ T cell differentiation.
  - Three eQTL variants that modify the expression of *HLA-B* also modify the expression of genes associated with skin inflammation.

Limitation: Small sample size, but high number of *HLA-B*15:02* carriers

Conclusion

- Proof of concept of a methodology for detecting polyvariant risk
- By improving the precision of genetic prediction for these potentially life-threatening adverse drug reactions, our model may enable more *HLA-B*15:02* carriers to receive CBZ treatment safely.
- Validation in an independent cohort is warranted.

Acknowledgement

We thank the patients, clinicians and researchers for participating and making the study possible.

This study was partly funded by the Australian National Health & Medical Research Council (NHMRC) (GNT1103979), Hong Kong Health and Medical Research Fund (01120086), and Science and Technology Project of Guangzhou Fund (201904010275 and 201904020028). PK is supported by an NHMRC Investigator Grant (GNT2025849).

References

1. Hung, S. I., Chung, W. H. & Chen, Y. T. HLA-B genotyping to detect carbamazepine-induced Stevens Johnson syndrome: implications for personalizing medicine. Per Med 2, 225-237, doi:10.2217/17410541.2.3.225 (2005).
2. SJS image from https://dermnetnz.org/topics/sjs-ten-images licensed under Creative Commons Attribution-NonCommercial-NoDerivs 3.0 (New Zealand)

Affiliations

1. Department of Neuroscience, School of Translational Medicine, Monash University, Melbourne, Australia; 2. Departments of Medicine and Neurology, Royal Melbourne Hospital, University of Melbourne, Australia; 3. Department of Neurology, Institute of Neuroscience, Key Laboratory of Neurogenetics and Channelopathies of Guangdong Province and the Ministry of Education of China, The Second Affiliated Hospital, Guangzhou Medical University, Guangzhou, China; 4. Department of Psychiatry, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; 5. Microbiology and Molecular Genetics, Institute of Biological Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur, Malaysia; 6. Department of Epidemiology and Preventive Medicine, Tel Aviv University, Tel Aviv, Israel; 7. Division of Neurology, Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; 8. Department of Neurology, Alfred Health, Melbourne, Australia; 9. Bladin-Berkovic Comprehensive Epilepsy Program, Department of Neurology, Austin Health, Melbourne, Australia; 10. Epilepsy Research Centre, Department of Medicine, Austin Health, The University of Melbourne, Melbourne, Australia; 11. Department of Neurology, The First Affiliated Hospital of Chongqing Medical University, Chongqing Key Laboratory of Neurology, Chongqing 400016, China; *Corresponding author