

Background

Sickle Cell Disease (SCD) is a genetic disorder caused by an *HBB* gene mutation and resulting structural variant haemoglobin (HbS) that polymerises upon deoxygenation, leading to a sickled morphology of red blood cells (*Chakravorty et al., Arch Dis Child 2015*). This causes acute pain episodes, known as vaso-occlusive crises (VOC), that can result in severe end-organ damage in the long-term and a shortened lifespan (*Jang et al., J Transl Med 2021*). Responsible for an estimated 95% of SCD hospitalisations and as a key predictor of death (*Darbari et al., Eur J Haematol 2020, Ballas et al., Am J Hematol 2005*), it is critical to find new ways to drive early detection of VOCs in order to support preventative interventions.

Aims

This work aimed to develop a machine learning (ML) algorithm capable of predicting the potential onset of a VOC, using physiological data captured longitudinally by a wearable smartwatch, as well as patient-reported outcomes (PROs) entered via a specialised mobile phone application.

Methods

Following informed patient consent, participants gained access to a mobile phone application, or "digital patient wallet", encompassing their health data and access to a PRO entry portal (*Summers et al., HemaSphere 2023*). This allowed for daily recording of PROs including EQ-5D-5L, pain, mood, and fatigue scores, alongside self-reported VOCs. Participants were also provided with a Withings ScanWatch to be worn day and night, capturing physical activity, sleep quality, and heart rate data in an automated manner. Medical record data including healthcare utilisation, pathology, and demographic data was obtained through participant completion of a Subject Access Request form.

Data for a snapshot cohort of 186 patients over a 1-month period (May-June 2023) were analysed through an ensemble of ML models: gradient boosting

Sign in via your Institution

Potential Articles of Interest

ELIPSIS: developing tools to better understand VOC in SCD

Nirmish Shah, Blood, 2021

Validation of Patient-Reported Vaso-Occlusive Crisis Day As a Potential Endpoint in Studies of Sickle Cell Disease

Karin Coyne et al., Blood

Incidence of Vaso-Occlusive Crisis Does Not Increase with Achieving Higher Hemoglobin Levels on Voxelotor Treatment or after Discontinuation: Analyses of the HOPE Study

Increased Patient Activation Is Associated with Fewer Emergency Room Visits and Hospitalizations for Pain in Adults with Sickle Cell Disease

Robert M Cronin et al., Pain Medicine, 2019

Epidural analgesia in a child with sickle cell disease complicated by acute abdominal pain and priapism

F. Labat et al., BJA: British Journal of Anaesthesia, 2001

Selenium Deficiency in a Mouse Model of Sickle Cell Disease Resulted in Increased Oxygen Consumption and Aberrant Mitochondrial Retention (OR11-05-19) machine, neural network, and k-means clustering. All collated metrics were fed into the model, with the exclusion of hospitalisation as an *'a posteriori'* variable, to ensure predictions were based on appropriate variables and not strengthened by self-confirming factors. The resulting predicted likelihood of a VOC for each participant was calculated on each day and given as a scale from 0-100%, with a likelihood of 75% or higher deemed to be a predicted VOC. This was then assessed against the true incidence of the participants' self-reported VOCs.

Results

The mean age was 36 (SD 11.6) years, and most patients were female (70%). The HbSS genotype was most common (75%), followed by HbSC (16%), and HbS Beta + Thalassemia (5%). 82,804 datapoints were collated for all participants across this period, encompassing 90 different variables. 69 instances of a VOC were self-reported by patients during the period of analysis, 58 (84%) of which were accurately predicted by the algorithm. 2,366 instances of no VOC were self-recorded, giving a negative predicted value of 83% (1,958/2,366 predicted). Further breakdown by the predicted likelihood score is displayed in Figure 2, highlighting the spread of individually calculated VOC likelihood scores per day against the incidence of patient-reported VOCs.

Summary/Conclusions

This work provides an important translation of daily trackable metrics through a remote, digital ecosystem, to the real-world prediction of potential VOC events in SCD. The purpose of this algorithm is, at its core, to act as a pre-emptive warning system from which proactive behavioural changes or interventions can be enacted. As such, the prediction of additional potential risk scores above the 75% threshold reflects an approach of 'over-predicting' considered to be a preferable strategy for the algorithm, given the severe nature of VOCs and the health impacts of missing a potential event.

Our ongoing aim is to further develop the algorithm and its contributing metrics against medical data, in addition to patient self-reported events, to further optimise the predictor. This includes the curation of core metrics requiring minimal patient input, in order to reduce missing datapoints across those that feed into the algorithm and thereby reduce unpredicted VOCs. Ultimately, we hope to provide an early alert system that will help patients to identify their VOCs sooner, and thus action the next steps that may prevent possible hospital admissions.

Disclosures

Summers: Pfizer Ltd: Other: Employed by Sanius Health, who have a financial work contract with the company. Agrippa: Pfizer Ltd: Consultancy. Lugthart: GBT: Other: Conference attendance. Anie: Pfizer Ltd: Honoraria. Telfer: Terumo: Speakers Bureau; Kyowa Kirin Limited: Other: Investigator-led funding; Napp Pharmaceuticals: Other: Clinical trial activity; Celgene: Other: Clinical trial activity; Apopharma: Other: Clinical trial activity, Speakers Bureau; Pfizer: Other: Advisory board; Clinical trial activity; Data monitoring committee; Novartis: Other: Advisory board; Clinical trial activity; bluebird bio: Other: Advisory board; Investigator-led funding;.

Figure 1

Hong et al., Current Developments in Nutrition, 2019



related cookies across the TrendMD network (widget, website, blog). Learn more

Yes	No
-----	----

View Metrics

Cited By

Google Scholar



Article Activity Alert First Edition Alert Latest Issue Alert Figure 1. Overview of the 10 variables with the highest importance score within the predictive algorithm. This is based on how much the variable is used in the neural net and gradient boosting machine.



Figure 2. Summary of the predicted scores of VOC likelihood against actual patient-reported instances of a VOC.



VIEW LARGE

Collections

Abstracts

Authors

DOWNLOAD SLIDE

Sign in via your Institution

Add comment

Advertisement



Submit to BloodPublic AccessContact UsAbout BloodPermissionsNewsroomSubscriptionsAlertsBlood Classifieds

- Current Issue First edition
- All Issues
- Advertising in Blood
- Terms and Conditions
- Twitter