ESC European Society of Cardiology

European Heart Journal - Digital Health (2021) **2**, 635–642
doi:10.1093/ehjdh/ztab082

ORIGINAL ARTICLE

A data mining-based cross-industry process for predicting major bleeding in mechanical circulatory support

Susanne E.A. Felix [1]*, Ayoub Bagheri[1,2], Faiz R. Ramjankhan[3], Marco R. Spruit[4], Daniel Oberski[2], Nicolaas de Jonge[1], Linda W. van Laake[1], Willem J.L. Suyker[3], and Folkert W. Asselbergs[1,5,6]

[1]Department of Cardiology, Division Heart & Lungs, University Medical Centre Utrecht, University of Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands; [2]Department of Methodology and Statistics, Faculty of Social Sciences, Utrecht University, Utrecht, the Netherlands; [3]Department of Cardiothoracic Surgery, University Medical Centre of Utrecht, University of Utrecht, Utrecht, the Netherlands; [4]Department of Information and Computing Sciences, Utrecht, Utrecht, the Netherlands; [5]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK; and [6]Health Data Research UK London and Institute of Health Informatics, University College London, London, UK

Received 13 December 2020; revised 17 February 2021; editorial decision 13 September 2021; accepted 27 September 2021; online publish-ahead-of-print 1 October 2021

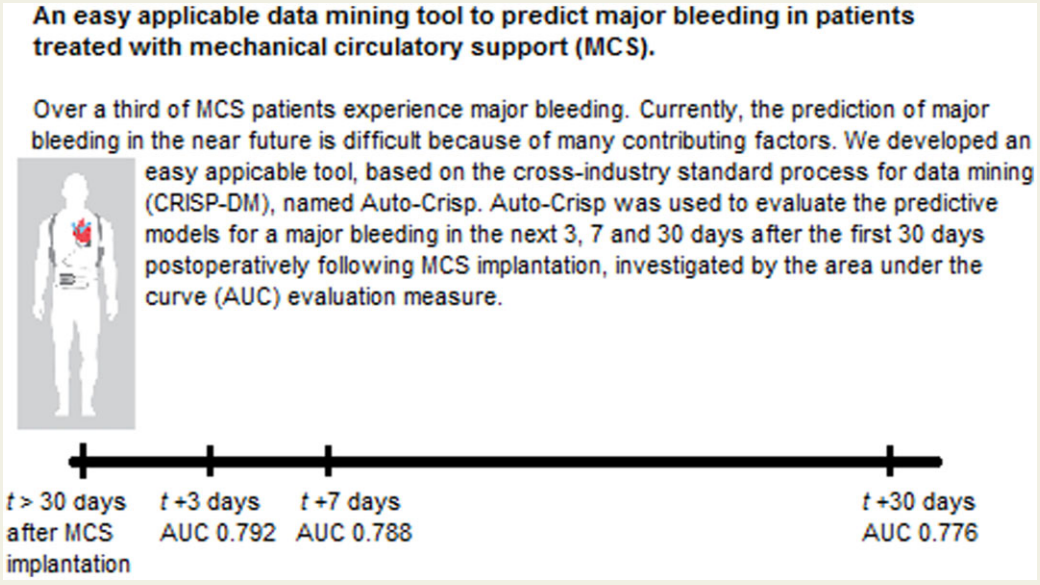
Aims Over a third of patients, treated with mechanical circulatory support (MCS) for end-stage heart failure, experience major bleeding. Currently, the prediction of a major bleeding in the near future is difficult because of many contributing factors. Predictive analytics using data mining could help calculating the risk of bleeding; however, its application is generally reserved for experienced researchers on this subject. We propose an easily applicable data mining tool to predict major bleeding in MCS patients.

Methods and results All data of electronic health records of MCS patients in the University Medical Centre Utrecht were included. Based on the cross-industry standard process for data mining (CRISP-DM) methodology, an application named Auto-Crisp was developed. Auto-Crisp was used to evaluate the predictive models for a major bleeding in the next 3, 7, and 30 days after the first 30 days post-operatively following MCS implantation. The performance of the predictive models is investigated by the area under the curve (AUC) evaluation measure. In 25.6% of 273 patients, a total of 142 major bleedings occurred during a median follow-up period of 542 [interquartile range (IQR) 205–1044] days. The best predictive models assessed by Auto-Crisp had AUC values of 0.792, 0.788, and 0.776 for bleedings in the next 3, 7, and 30 days, respectively.

Conclusion The Auto-Crisp-based predictive model created in this study had an acceptable performance to predict major bleeding in MCS patients in the near future. However, further validation of the application is needed to evaluate Auto-Crisp in other research projects.

* Corresponding author. Tel: +31887555555. Email: S.E.A.Felix-2@umcutrecht.nl

Graphical Abstract



Keywords Mechanical circulatory support • Cross-industry standard process for data mining (CRISP-DM) • Bleeding • Data mining • Prediction

Introduction

Patients with refractory advanced (end-stage) heart failure may be treated with mechanical circulatory support (MCS) by a continuous-flow left ventricular assist device (cf-LVAD), which is inserted into the heart and takes over left ventricular heart function (*Figure 1*). These devices are implanted as a bridge to heart transplantation, bridge to candidacy (e.g. to evaluate the reversibility of organ failure), or destination therapy (if heart transplantation is contra-indicated).[1–4]

As a result of the shortage of donor hearts and the use as destination therapy, the duration of MCS has increased in the recent years. The use of cf-LVADs has potential disadvantages in terms of adverse events, affecting a substantial percentage of the patients, both in short-term and long-term support. For example, bleeding occurs in 35% of the patients.[4] Bleedings are most frequently located in the gastro-intestinal tract, often requiring blood transfusion and readmission.[5–7] Also bleedings located elsewhere, for example, intracranial haemorrhage, result in significant morbidity and an increased risk of mortality.[8,9] The occurrence of bleeding is associated with the use of anticoagulation, which is required to prevent thrombosis, provoked by the foreign material. Anticoagulation consists of a vitamin K antagonist, mostly in combination with a thrombocyte aggregation inhibitor. Furthermore, after cf-LVAD implantation, bleeding risk is increased because the majority of patients have an acquired coagulopathy disorder, caused by several factors including platelet dysfunction and an impaired function of the von Willenbrand factor.[10,11]

Current knowledge on the risk of bleeding during MCS is based on baseline and peri-operative data using 'conventional' regression methods.[1,12] Pre-operative risk factors for bleeding after discharge include an age above 65 years old, female gender, ischaemic aetiology, and the lowest quartile haematocrit (<31%).[13]

However, the risk of bleeding differs over time.[14] In addition, it has been observed that some patients require recurrent hospitalization for gastro-intestinal bleedings, while others are admitted once.[15] Therefore, an accurate prediction of a bleeding in the short term at any moment during MCS is required to optimize the outcome. Identification of a patient 'at risk' could result in closer monitoring and/or change in the anticoagulation regimen to limit or even prevent the actual occurrence of bleeding. At present, no predictive models are available to accompany this. As the risk of bleeding changes over time, we may overlook a (number of) risk factor(s). The addition of data collected during follow-up might improve the predictive performance. Furthermore, using data mining can contribute in the identification of risk factors.[16]

Cross-industry standard process for data mining (CRISP-DM) is a standard approach which will help to translate the research question into data mining tasks, to select appropriate data transformations and data mining techniques, and to provide means for evaluating the effectiveness of the results.[17,18] However, for most healthcare professionals, this method remains a major challenge due to limited knowledge on data mining. We propose a data mining-based approach according to the CRISP-DM methodology to predict a future major bleeding in MCS patients, excluding the initial post-operative phase.

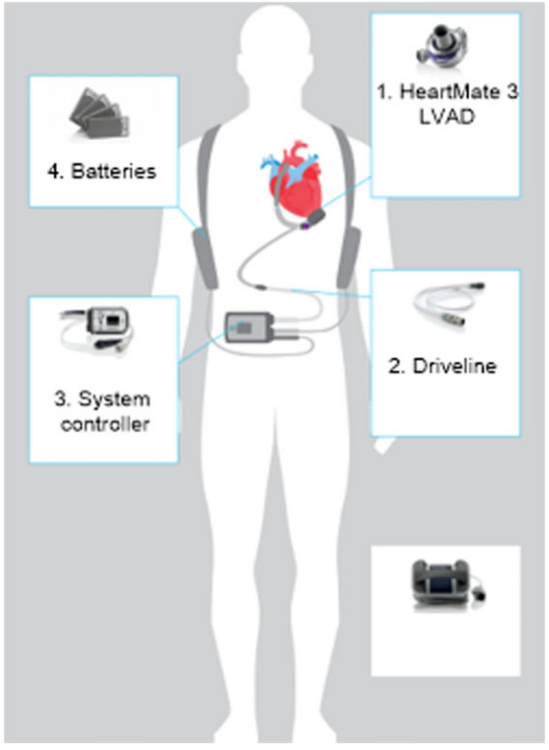


Figure 1 Continuous-flow left ventricular assist device (cf-LVAD), including all parts of the device. 1. HeartMate 3 (HM 3) LVAD. 2. Driveline, connecting the HM 3 to the system controller. 3. System controller, which senses the function of the LVAD and controls the power. 4. Batteries.

Methods

In this section, the study data are introduced and the methodology of this study is described in summary. In the Supplementary material online, *Appendix*, detailed information and background information on the application, named Auto-Crisp, is described.

Study data

Data of all patients who received a cf-LVAD at the University Medical Centre of Utrecht between 2006 and 2018 were collected from their electronic health records (EHRs). These comprise laboratory results, medication and baseline data registered in the local EHRs and results of all international normalized ratio (INR) values from the anticoagulation clinics during MCS (the latter was available in 77 patients). The ethics committee of the University Medical Centre Utrecht approved this study.

Electronic health records contain heterogeneous data originating from multiple sources. To address this, we distinguished medical data into the following types: time point data, time interval data, baseline data, and event data. Time point data contain data about a dynamic variable at one time point (e.g. INR), whereas time interval data contain data about a time period of a dynamic variable (e.g. medicine X from March to June). Baseline data include static variables without a temporal aspect (e.g. sex). Lastly, event data are time point data which contain the events to be predicted, also known as class labels (e.g. mortality or, in this study, major bleeding). The end point of the study was the occurrence of a major bleeding, that occurred beyond 30 days after the implantation of the cf-LVAD, thereby excluding surgery-related bleedings, as these probably have another aetiology. Major bleeding was defined according to the definition of the Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS): an episode of suspected internal or external bleeding that results in death, re-operation, hospitalization, and/or transfusion of red blood cells.[1]

Auto-Crisp

CRISP-DM is a widely used open standard process model with which a data mining project is built. As shown in Supplementary material online, *Appendix Figure S1*, CRISP-DM consists of six steps: business understanding, data understanding, data preparation, modelling, evaluation, and deployment. In business understanding, the project plan is produced, following the objective of the study, inventory of the resources, and the determination of the data mining goals. Data understanding is the step in which initial data are collected, a description report is provided, data are explored, and the quality of the data is verified. Then, data are prepared by the selection of data, data cleaning, construction, integration, and formatting of the data. In the modelling step, the modelling techniques are selected, a test design is generated, and the model is built. Thereafter, the models are evaluated and the process is reviewed. The deployment step comprises the production of the final report and monitoring and maintenance of the plan.

Auto-Crisp guides users through these steps of the data mining project, providing sensible defaults for each step and automating most processes that have no or little influence. However, it still provides some design choices. To enable the use of this application for other researchers, the complete source code is available at https://github.com/Susannefelix/AutoCrisp.

After the modelling step, the output of the built model is depicted in a confusion matrix out of which the sensitivity and specificity, the positive and the negative predictive values are visualized. Furthermore, a receiver operating characteristic (ROC)-curve is used with corresponding area under the curve (AUC) to assess the performance of the model. We report on the best functioning model in different settings.

Auto-Crisp was used to create models to predict a major bleeding from any moment during MCS (t_i), excluding the first 30 post-operative days. The different settings refer to the time frames (3, 7, and 30 days) in which the major bleeding was predicted. For the occurrence of a major bleeding at t_i and within the timeframes of 3, 7, and 30 days, an AUC of 0.5–0.6 was interpreted as a poor diagnostic value, 0.6–0.7 as fair, 0.7–0.8 as acceptable, 0.8–0.9 as excellent, and 0.9–1 as outstanding.[19]

Results

Between 2006 and 2018, 273 patients (69% male) received a cf-LVAD. Follow-up was completed for all patients for a median period of 542 (IQR 205–1044) days. In total, 633 487 time points were included, on average 2320 laboratory values per patient. The time interval dataset consisted of 54 385 time intervals and thus an average of 1367 medications were registered per patient. During a total follow-up period of 510 patient-years (i.e. the duration of support in years per patient, multiplied by the number of patients), 142 bleedings beyond 30 days after cf-LVAD implantation occurred. Major bleeding affected 25.6% of the patients, of which 76% were male with a mean age of 52.3 years old. In Auto-Crisp, an overview of these descriptive data is shown (*Figure 2*). Insight is given concerning the quality of the data, including the amount of missing data and, in this study,

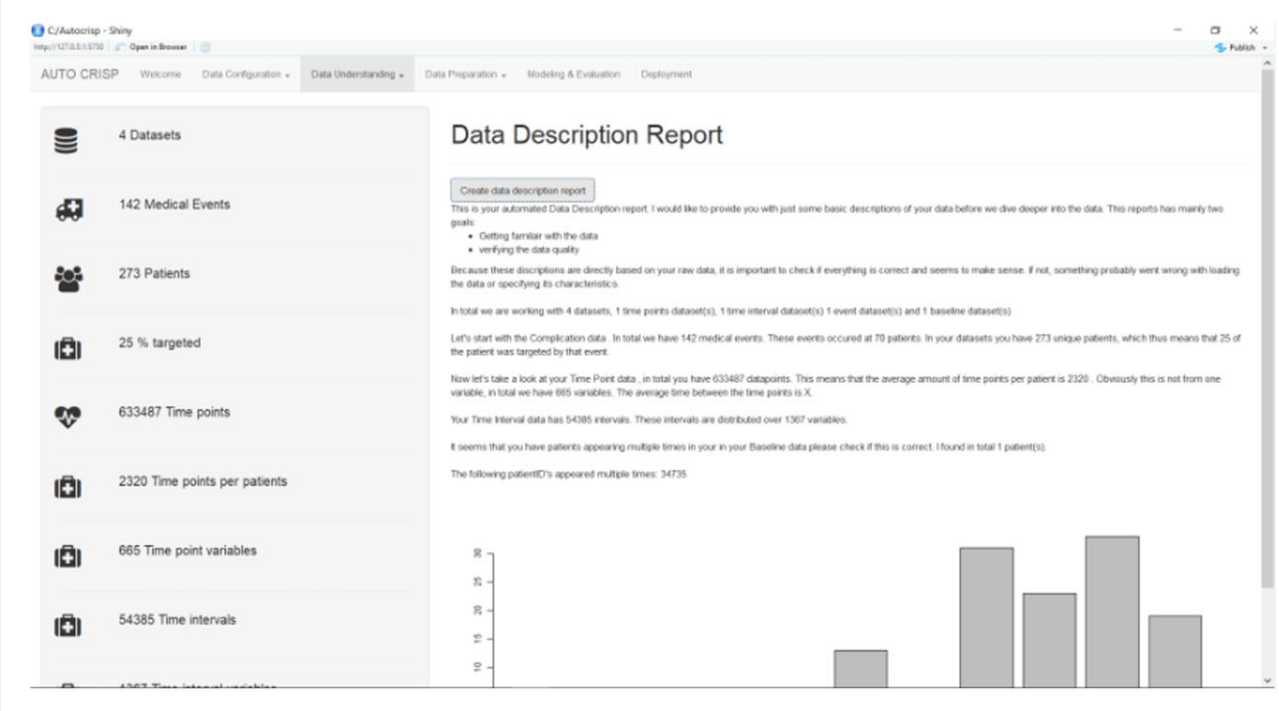


Figure 2 Data description report. The data description report reveals an overview of the descriptive statistics from the different datasets.

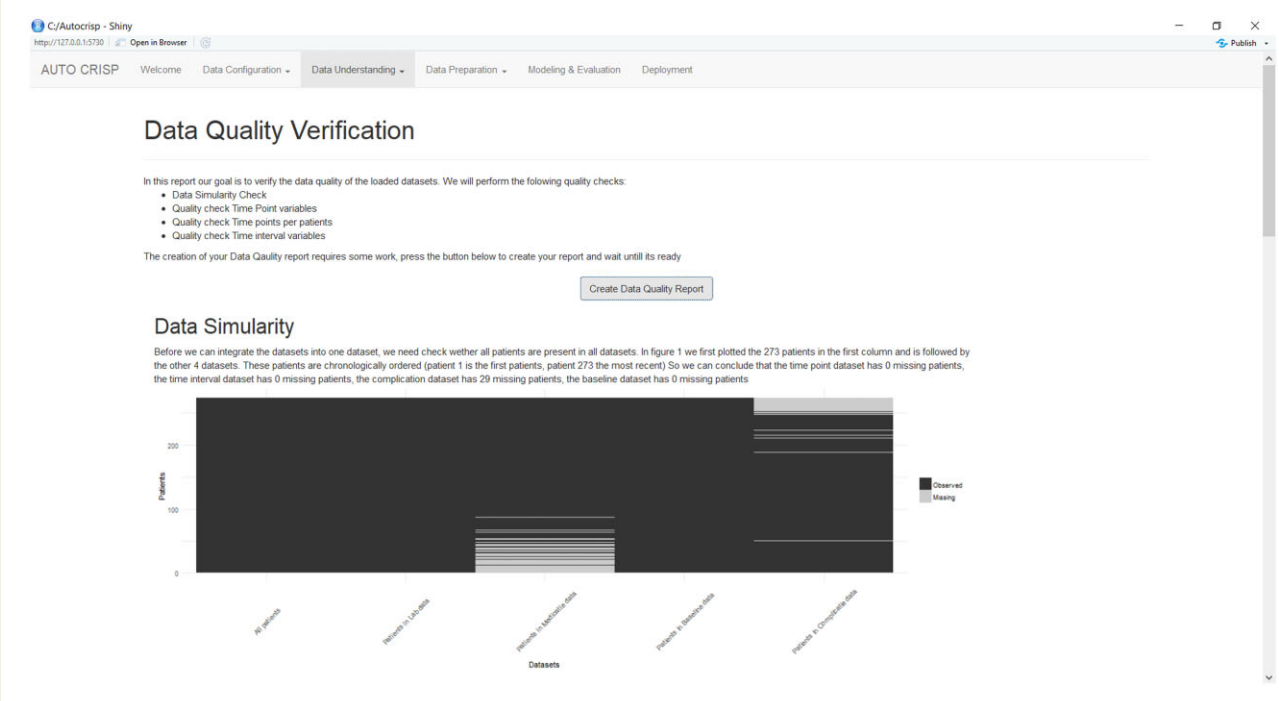


Figure 3 Data quality verification. The data quality verification tab reveals information on the completeness of the datasets, for example, if all patients are represented in each dataset.

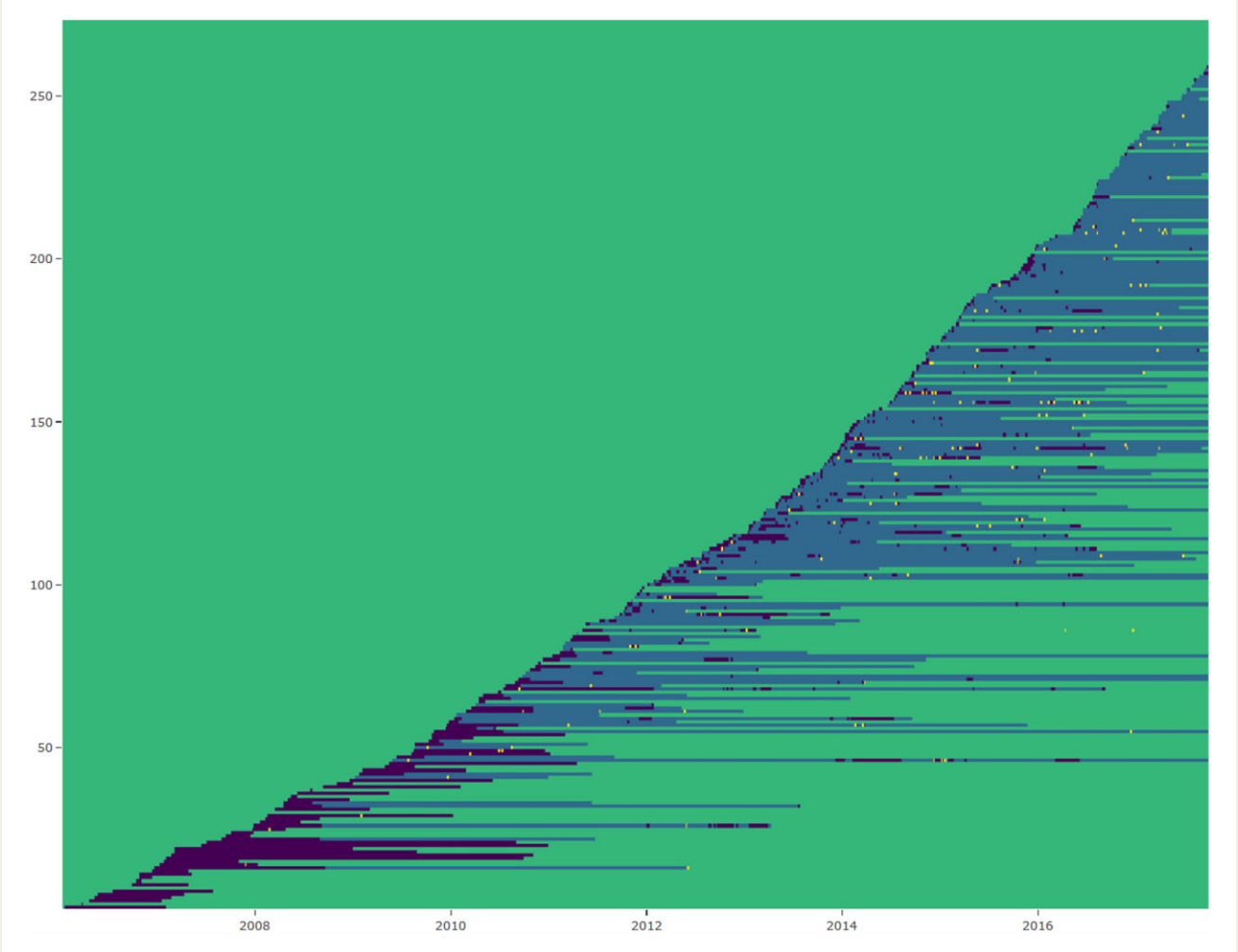


Figure 4 The use of oral anticoagulation and events. A graphical overview to visualize the use of anticoagulation during the study period and the moment(s) of a bleeding event. Purple: inclusion in cohort, no oral anticoagulation. Blue: inclusion in cohort, use of oral anticoagulation. Yellow: bleeding complication.

Table 1 AUC-values of the model in predicting a major bleeding >30 days after cf-LVAD implantation, with and without the data 7 days after the bleeding event

Setting	Full window	Cleaned dataset
1: 0	0.799	0.824
2: 1–3	0.779	0.792
3: 1–7	0.779	0.788
4: 1–30	0.769	0.776

Table 2 Feature extraction

Setting	Last value	Statistics	Summary	TVA	Pattern mining
1: 0	0.824	0.817	0.830	0.820	0.820
2: 1–3	0.792	0.785	0.780	0.787	0.828
3: 1–7	0.788	0.799	0.796	0.777	0.794
4: 1–30	0.776	0.803	0.796	0.780	0.721
Baseline	0.594	0.582	0.664	0.595	Not applicable

TVA,: trend-based approximation.

an illustrative overview of the duration of MCS together with the duration of anticoagulation and the occurrence of a bleeding (*Figures 3* and *4*).

The haemoglobin level was lower in the patients with a bleeding in comparison to the total population, while the creatinine and the INR were higher in patients with a bleeding, though not significantly different ($P = 0.36$ and $P = 0.25$, respectively). The level of lactate dehydrogenase (LD) showed a peak approximately 2 weeks prior to the bleeding event, though had a broad standard deviation, as a result of very high LD-levels in a few patients.

Table 3 Classifier algorithm

Setting	Random forest	Naïve Bayes	Decision tree	Logistic regression	Support vector machine
1: 0	0.824	0.690	0.501	0.793	0.772
2: 1–3	0.792	0.773	0.494	0.787	0.701
3: 1–7	0.788	0.773	0.524	0.754	0.657
4: 1–30	0.776	0.770	0.490	0.705	0.708

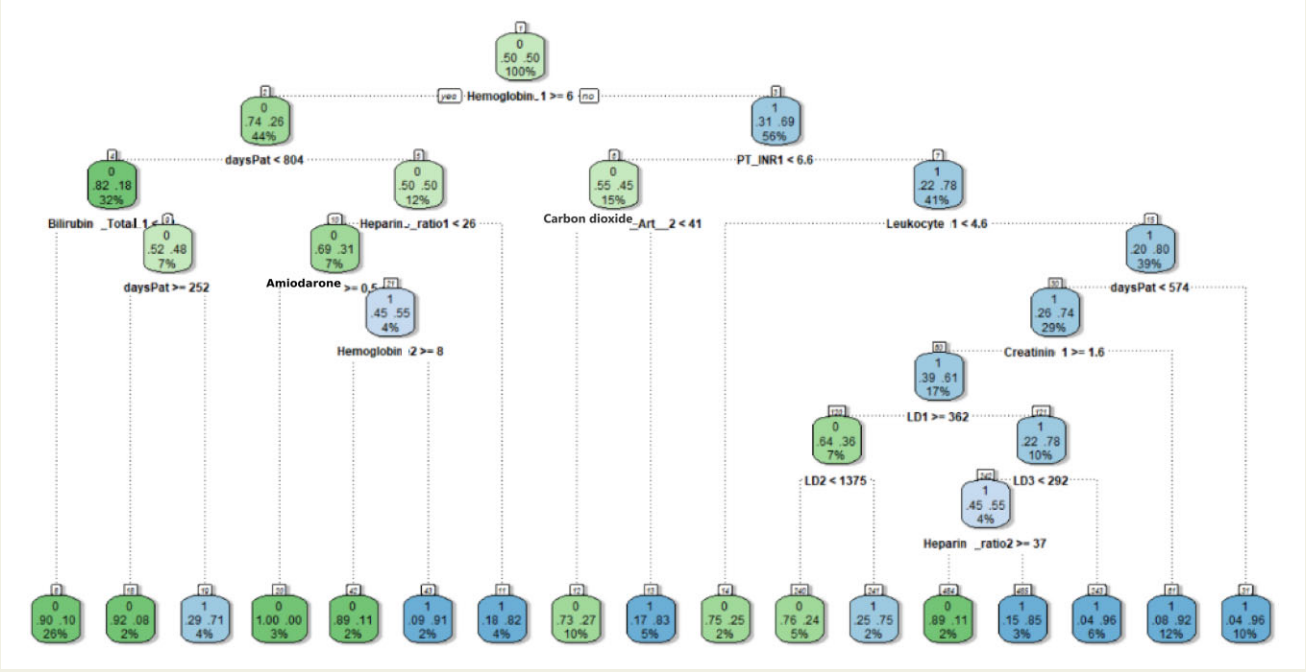


Figure 5 Decision tree in setting 3, predicting a bleeding within 7 days from t_i, using undersampling. An example of a decision tree, which first splits on the most recent haemoglobin level. The plot mostly splits on the most recent measurements (e.g. haemoglobin_1), which implies that these are most important for the tree than older measurements '1' refers to the most recent result, '2' to the second-to-last result et cetera. At the top, the decision tree is split into two classes: class 0 (green) and class 1 (blue), which contain both 50% of the data. On the left, the boxes are green because the majority (74 vs. 26%) belong to class 0. On the right, the boxes are blue because the majority (69 vs. 31%) belong to class 1. The percentages in the boxes refer to the percentage of data which will be included in the left or right side of the nodes; for example following the first split, 56% of data will go to the right and 44% to the left. t_i, any moment during MCS after the first 30 post-operative days; haemoglobin, haemoglobin level in mmol/L; Carbon dioxide_Art, arterial carbon dioxide tension in mmHg; DaysPat, days on mechanical circulatory support; LD, lactate dehydrogenase level in U/L; Leukocyte, number of leukocytes $\times 10^9$/L; PT_INR, international normalized ratio.

Data preparation

A plot of the time interval data set (medication) showed an increase in the mean amount of medication prescribed from 2011. The mean number of medications per patient prior to this period was too low to represent the truth, as in general, the mean number of medication has not changed over the years. The low number of medication prior to 2011 was probably related to the limited digital medication prescription. Therefore, these medication files were excluded.

We cleaned the data by excluding the data of 7 days following the bleeding event, because these are not relevant in predicting the event and result in the noise of the model. In *Table 1*, the AUC-values of the model are shown, both with and without cleaning the data as described.

Feature extraction and selection

In Auto-Crisp, different feature extraction methods can be selected (see the Supplementary material online, *Appendix* for a detailed explanation). In this study, the Summary technique achieved the highest AUC in most settings, where for example the frequent sequential pattern mining method did not improve the accuracy of the model despite its complex approach (*Table 2*).[20]

Table 4 Significant predictors on logistic regression analysis

Coefficients	Estimate	Standard error	Z value	Pr (>|z|)
Gender 1	-0.8200176	0.3872655	-2.117	0.034221
Sodium 1	-0.0549085	0.0275277	-1.995	0.046080
Haemoglobin 1	-1.4089387	0.3863737	3.647	0.000266
Creatinine 1	0.6028400	0.2155308	2.797	0.005158
Creatinine 2	0.8917792	0.3157466	2.824	0.004738
Creatinine 3	-1.4654553	0.3553743	-4.124	3.73e-05
Carbon dioxide_art_1	0.1010064	0.0506766	1.993	0.046244
Days on MCS	0.0016145	0.0003339	4.835	1.33e-06
Calcium carbonate	0.9008932	0.3248505	2.773	0.005550
Clopidogrel	1.1135802	0.4966252	2.242	0.024942

'1': most recent laboratory result, '2': second-to-last, et cetera.
Carbon dioxide_art, arterial carbon dioxide tension in mmHg; MCS, mechanical circulatory support.

For feature selection, both Boruta and recursive feature elimination (RFE) did not improve the AUC, while both were very time consuming.[21] Therefore, we did not exclude any features.

Modelling

Despite unbalanced datasets in this study, where only 2% of the data were from patients with a major bleeding, none of the data sampling techniques revealed improvement of the AUC-level of the model. The random forest classifier provided the best performance in comparison to logistic regression and decision tree (*Table 3*). As an example of the various models, one of the decision trees is shown, using undersampling and prediction of a bleeding within 7 days is visualized in *Figure 5*. The main determinants of the logistic regression analysis are shown in *Table 4*.

The overall accuracy of the best predictive model of a major bleeding in the next 3, 7, and 30 days corresponds to an AUC-level of 0.792, 0.788, and 0.776, respectively.

Discussion

In MCS patients, the substantial amount of adverse events is still the Achilles heel of this therapy. Amongst the adverse events, bleeding has been shown to be related to both the use of anticoagulation and acquired coagulopathies in these patients. However, the prediction of a major bleeding over time is difficult. In this study, including all available data out of the EHR, we evaluated predictive factors for a major bleeding after the first postoperative phase. By implementing data mining, based on the CRISP-DM methodology into a semi-automated tool called Auto-Crisp, we demonstrated acceptable prediction of a major bleeding in a MCS patient from any moment following the first 30 days after implantation. The models developed were applicable for risk assessment in several time frames, with a maximum of 30 days.[19] This is a promising tool with a relatively easy applicability, even for clinicians and healthcare professionals. Auto-Crisp contains automated steps in the data mining process, though does reveal insight into the quality of the data at different steps in an intuitive manner. As well, as a result of the automated steps, it is less time consuming. Further validation of the application is needed to evaluate Auto-Crisp in other research projects.

The prediction of a bleeding in the near future could result in a change of the anticoagulation regimen. For example, target INR could be decreased or platelet inhibition interrupted, weighing the risk of thrombosis on the contrary. Recently, an analysis was performed on the safety of a lower target INR range in patients implanted with the newest device [HeartMate 3 (HM 3)], as the HM 3 had already demonstrated to have less thrombo-embolic complications in comparison with its predecessor, the HM II.[22] Using a lower target INR did not result in thrombo-embolic events over 6 months in a small number of HM 3 patients ($n = 15$), major bleeding in only one patient.[23] In addition, a lower dose of Aspirin (81 instead of 325 mg) did not change the rates of bleeding or thrombosis during 2-year follow-up of HM 3 patients.[24,25] However, for the individual patient identified as a patient 'at risk' of major bleeding, all contributing factors should be taken into account to make the decision to lower the intensity of the anticoagulation.[26]

Because EHRs usually have sparsely sampled data, resulting in a lot of missing data points, imputation of the data is useful to evaluate as much data as available. In this study, there was no significant effect of the type of imputation technique in terms of the accuracy of the model. However, some parameters might be underexposed because of their disproportional measurements. This might be the explanation of the competitive results of the AUC-level between the random forest classifier algorithm and the 'conventional' logistic regression analysis.

Furthermore, the completeness of the data is of utmost importance. There is an international registry in which all baseline, peri-procedural and adverse event data are registered for all MCS patients. This registry is suitable for analysing patterns in adverse events, for example.[27,28] However, individual prediction of a specific (type of) adverse events is difficult due to a lot of missing data.

In MCS patients, the addition of (continuous) functional data of the pump, e.g., power, flow, and speed, might contribute to the ability to predict certain adverse events, for example, pump thrombosis (then, power is increased and flow diminished). Currently, the actual values of these data are continuously visualized on the screen of the controller of the LVAD, and data of a longer period are administered by the manufacturer. If the temporal changes of these data will be collected and visualized, e.g., by an application, it might detect pump thrombosis at an early stage.[29] For future perspectives, the integration of risk models into the EHR could identify a patient at risk of an adverse event, ideally resulting in an intervention to prevent or limit the consequences.

Conclusion

Auto-Crisp is an application that enables data mining for healthcare professionals with less experience in data mining. Using Auto-Crisp, a model was created to predict major bleeding in MCS patients in the near future with acceptable accuracy. We proposed validation and further development of Auto-Crisp to provide the application of data mining for healthcare professionals with less experience in data science.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Funding

This project has received unrestrictive funding from Abbott. Folkert W. Asselbergs is supported by UCL Hospitals NIHR Biomedical Research Centre.

Conflict of interest: none declared.

References

1. Goldstein DJ, Meyns B, Xie R, Cowger J, Pettit S, Nakatani T, Netuka I, Shaw S, Yanase M, Kirklin JK. Third Annual Report From the ISHLT Mechanically Assisted Circulatory Support Registry: a comparison of centrifugal and axial continuous-flow left ventricular assist devices. *J Heart Lung Transplant* 2019; **38**:352–363
2. Kirklin JK, Pagani FD, Kormos RL, Stevenson LW, Blume ED, Myers SL, Miller MA, Baldwin JT, Young JB, Naftel DC. Eighth annual INTERMACS report: special focus on framing the impact of adverse events. *J Heart Lung Transplant* 2017;**36**: 1080–1086.
3. Teuteberg JJ, Cleveland JC Jr, Cowger J, Higgins RS, Goldstein DJ, Keebler M, Kirklin JK, Myers SL, Salerno CT, Stehlik J, Fernandez F, Badhwar V, Pagani FD, Atluri P. The Society of Thoracic Surgeons Intermacs 2019 Annual Report: the changing landscape of devices and indications. *Ann Thorac Surg* 2020;**109**: 649–660.
4. Kirklin JK, Xie R, Cowger J, de By TMMH, Nakatani T, Schueler S, Taylor R, Lannon J, Mohacsi P, Gummert J, Goldstein D, Caliskan K, Hannan MM. Second annual report from the ISHLT mechanically assisted circulatory support registry. *J Heart Lung Transplant* 2018;**37**:685–691.
5. Forest SJ, Bello R, Friedmann P, Casazza D, Nucci C, Shin JJ, D'Alessandro D, Stevens G, Goldstein DJ. Readmissions after ventricular assist device: etiologies, patterns, and days out of hospital. *Ann Thorac Surg* 2013;**95**:1276–81.
6. Draper KV, Huang RJ, Gerson LB. GI bleeding in patients with continuous-flow left ventricular assist devices: a systematic review and meta-analysis. *Gastrointest Endosc* 2014;**80**:435–446.
7. Goldstein DJ, Aaronson KD, Tatooles AJ, Silvestry SC, Jeevanandam V, Gordon R, et al. Gastrointestinal bleeding in recipients of the HeartWare ventricular assist system. *JACC Heart Fail* 2015;**3**:303–13.
8. Cho SM, Moazami N, Frontera JA. Stroke and intracranial hemorrhage in HeartMate II and HeartWare left ventricular assist devices: a systematic review. *Neurocrit Care* 2017;**27**:17–25.
9. Tsiouris A, Heliopoulos I, Mikroulis D, Mitsias PD. Stroke after implantation of continuous flow left ventricular assist devices. *J Card Surg* 2019;**34**:541–548.
10. Steinlechner B, Dworschak M, Birkenberg B, Duris M, Zeidler P, Fischer H, Milosevic L, Wieselthaler G, Wolner E, Quehenberger P, Jilma B. Platelet dysfunction in outpatients with left ventricular assist devices. *Ann Thorac Surg* 2009; **87**: 131–7.
11. Crow S, Chen D, Milano C, Thomas W, Joyce L, Piacentino V, 3rd, Sharma R, Wu J, Arepally G, Bowles D, Rogers J, Villamizar-Ortiz N. Acquired von Willebrand syndrome in continuous-flow ventricular assist device recipients. *Ann Thorac Surg* 2010;**90**:1263–9; discussion 9.
12. Magnussen C, Bernhardt AM, Ojeda FM, Wagner FM, Gummert J, de By TMMH, Krabatsch T, Mohacsi P, Rybczynski M, Knappe D, Sill B, Deuse T, Blankenberg S, Schnabel RB, Reichenspurner H. Gender differences and outcomes in left ventricular assist device support: the European Registry for Patients with Mechanical Circulatory Support. *J Heart Lung Transplant* 2018;**37**:61–70.
13. Boyle AJ, Jorde UP, Sun B, Park SJ, Milano CA, Frazier OH, Sundareswaran KS, Farrar DJ, Russell SD; HeartMate II Clinical Investigators. Pre-operative risk factors of bleeding and stroke during left ventricular assist device support: an analysis of more than 900 HeartMate II outpatients. *J Am Coll Cardiol* 2014;**63**:880–8.
14. Kormos RL, Cowger J, Pagani FD, Teuteberg JJ, Goldstein DJ, Jacobs JP, Higgins RS, Stevenson LW, Stehlik J, Atluri P, Grady KL, Kirklin JK. The Society of Thoracic Surgeons Intermacs database annual report: evolving indications, outcomes, and scientific partnerships. *J Heart Lung Transplant* 2019;**38**:114–126.
15. Welden CV, Truss W, McGwin G, Weber F, Peter S. Clinical predictors for repeat hospitalizations in left ventricular assist device (LVAD) patients with gastrointestinal bleeding. *Gastroenterol Res* 2018;**11**:100–105.
16. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol* 2017;**69**:2657–2664.
17. Wirth, R., Hipp, J. CRISP-DM: towards a standard process model for data mining. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*. 2000, pp. 29–39.
18. Piatetsky, G. (2014). CRISP-DM, still the top methodology for analytics, data mining, or data science projects. http://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science- projects.html.
19. Hosmer, D., Lemeshow, S., Sturdivant, R. Applied logistic regression. 3rd ed. John Wiley & Sons, Hoboken, NJ; 2013.
20. Batal I., Fradkin D., Harrison J., Moerchen F., Hauskrecht M. Mining recent temporal patterns for event detection in multivariate time series data. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2012, pp. 280–288.
21. Kursa, M. B., Rudnicki, W. R. Feature selection with the Boruta package. *J Stat Softw* 2010;**36**:1–13.
22. Mehra MR, Uriel N, Naka Y, et al. A fully magnetically levitated left ventricular assist device - final report. *N Engl J Med* 2019;**380**:1618–1627.
23. Netuka I, Ivák P, Tučanová Z, et al. Evaluation of low-intensity anti-coagulation with a fully magnetically levitated centrifugal-flow circulatory pump-the MAGENTUM 1 study. *J Heart Lung Transplant* 2018;**37**:579–586.
24. Saeed O, Colombo PC, Mehra MR, Uriel N, Goldstein DJ, Cleveland J, Connors JM, Najjar SS, Mokadam NA, Bansal A, Crandall DL, Sood P, Jorde UP. Effect of aspirin dose on hemocompatibility-related outcomes with a magnetically levitated left ventricular assist device: an analysis from the MOMENTUM 3 study. *J Heart Lung Transplant* 2020;**39**:518–525.
25. Gustafsson F. Is aspirin needed to prevent thrombotic events in patients treated with a HeartMate 3 left ventricular assist device? *J Heart Lung Transplant* 2020;**39**: 526–528.
26. Shah P, Tantry US, Bliden KP, Gurbel PA. Bleeding and thrombosis associated with ventricular assist device therapy. *J Heart Lung Transplant* 2017;**36**: 1164–1173.
27. Movahedi F, Kormos RL, Lohmueller L, Seese L, Kanwar M, Murali S, Zhang Y, Padman R, Antaki JF. Sequential pattern mining of longitudinal adverse events after left ventricular assist device implant. *IEEE J Biomed Health Inform* 2020;**24**:2347–2358.
28. Kilic A, Seese L, Pagani F, Kormos R., Kilic A, et al. Identifying temporal relationships between in-hospital adverse events after implantation of durable left ventricular assist devices. *J Am Heart Assoc* 2020;**9**:e015449.
29. Grabska J, Schlöglhofer T, Gross C, Maw M, Dimitrov K, Wiedemann D, Zimpfer D, Schima H, Moscato F. Early detection of pump thrombosis in patients with left ventricular assist device. *ASAIO J* 2020;**66**:348–354.