Enhancing Transvenous Lead Extraction Risk Prediction: Integrating Imaging Biomarkers into Machine Learning Models

Short Title: Using imaging to predict risk following TLE.

Authors:

Vishal S Mehta MBBS [1,2] – corresponding author

YingLiang Ma PhD [2,3]

Nadeev Wijesuriya MBBS [1,2]

Felicity DeVere MBBS [1,2]

Sandra Howell MBBS [1,2]

Mark K Elliott MBBS PhD [1,2]

Nilanka N Mannkakara MBBS [1,2]

Tatiana Hamakarim[1,2]

Tom Wong PhD [1,2,4]

Hugh O'Brien PhD [2]

Steven Niederer DPhil [2,4]

Reza Razavi PhD [1,2]

Christopher A Rinaldi MD FHRS [1,2]

Affiliations:

[1] Cardiology Department, Guy's and St Thomas' NHS Foundation Trust, London, UK

[2] School of Biomedical Engineering and Imaging Sciences, King's College London, UK

[3] School of Computing Sciences, University of East Anglia, Norwich, UK

[4] National Heart and Lung Institute, Imperial College London, Hammersmith Hospital, London, UK

Correspondence:

Dr Vishal Mehta

School of Biomedical Engineering and Imaging Sciences

St Thomas' Hospital

London, SE1 7EH, U.K.

Email: vishal.mehta@kcl.ac.uk

Keywords

Transvenous lead extraction; TLE; Complications; Machine Learning; Artificial intelligence

Sources of funding

The authors are supported by the Wellcome/EPSRC Centre for Medical Engineering (WT203148/Z/16/Z).

Disclosures

NW receives fellowship funding from the British Heart Foundation (FS/CRTF/22/24362). VM has received fellowship funding from Siemens and Abbott. NNM is in receipt of fellowship funding from Heart Research UK (grant no. RG2701) and Abbott. YM receives research funding from UK Engineering and Physical Sciences Research Council (EP/X023826/1). SAN acknowledges support from the UK Engineering and Physical Sciences Research Council (EP/M012492/1, NS/A000049/1, and EP/P01268X/1), the British Heart Foundation (PG/15/91/31812, PG/13/37/30280, SP/18/6/33805), US National Institutes of Health (NIH R01-HL152256), European Research Council (ERC PREDICT-HF 864055) and Kings Health Partners London National Institute for Health Research (NIHR) Biomedical Research Centre. CAR receives research funding and/or consultation fees from Abbott, Medtronic, Boston Scientific, Spectranetics and MicroPort outside of the submitted work.

Abstract – 250 words

Background:

Machine learning (ML) models have been proposed to predict risk related to transvenous lead extraction (TLE).

Objective:

We tested if integrating imaging data into an existing ML model increases its ability to predict major adverse events (MAE: procedure-related major complications and procedure-related deaths) and lengthy procedures (≥100 minutes).

Methods:

We hypothesised certain features: i) lead angulation ii) coil percentage inside the superior vena cava (SVC), and iii) number of overlapping leads in the SVC, detected from a pre-TLE plain anterior-posterior (AP) chest x-ray (CXR) would improve prediction of MAE and long procedure times. A deep-learning convolutional neural network was developed to automatically detect these CXR features.

Results:

1050 cases were included, with 24 (2.3%) MAEs. The neural network was able to detect: i) heart border with 100% accuracy, ii) coils: 98% accuracy, iii) acute angle in the right ventricle and SVC: 91% and 70% accuracy respectively. The following features significantly improved MAE prediction: i) ≥50% coil within the SVC, ii) ≥2 overlapping leads in the SVC, and iii) acute lead angulation. Balanced accuracy (0.74 to 0.87), sensitivity (68% to 83%), specificity (72% to 91%), and area under the curve (AUC) (0.767 to 0.962) all improved with imaging biomarkers. Prediction of lengthy procedures also improved: balanced accuracy (0.76 to 0.86), sensitivity (75% to 85%), specificity (63% to 87%), and AUC (0.684 to 0.913).

Conclusion:

Risk prediction tools integrating imaging biomarkers significantly increases the ability of ML models to predict risk of MAE and long procedure time related to TLE.

Abbreviations

ML - Machine Learning

TLE - Transvenous lead extraction

SVC - Supervior Vena Cava

AP - Anetrior Posterior

CXR - Chest X-Ray

MAE - Major Adverse Event

AUC - Area Under the Curve

CIED - Cardiac Implantable Electronic Device

DL - Deep Learning

CNN - Convolutional Neural Network

GSTT - Guy's and St Thomas' Trust

ELECTRa - European Lead Extraction ConTRolled

IQR - Interquartile Range

ADASYN - Adaptive Synthetic

JSC - Jaccard Similarity Coefficient

SNN - Spiking Neural Network

SVM - Support Vector Machines

RF - Random Forest

ICD - Implantable Cardioverter Defibrillator

Introduction

The progressive rise of cardiac implantable electronic devices (CIEDs) over the last 20 years has been mirrored by an increased requirement for lead extraction. Transvenous lead extraction (TLE) remains a safe procedure in the majority of cases, with the incidence of procedure related major complications quoted as 1.6% in a systematic review of 18,433 patients [1]. Whilst procedural success is high, there are an increasing number of patients at risk of adverse outcomes due to a high burden of comorbidities and multiple previous device interventions[2]. Improved risk stratification of patients based on likelihood of risk of adverse outcomes, and procedural complexity, is important for informed patient consent and appropriate resource allocation basis[3].

Multiple risk scores have been developed; however, their predictive capability is constrained by the rarity of adverse events, the binary nature of pre-procedural risk factors, and variation in methodology[4,5]. We previously published a risk assessment model to predict risk of adverse events utilising machine learning methods. The study used patient-level data derived from the ESC EORP Electra registry of 3,555 patients undergoing TLE in 73 centres across Europe [6]. This ML model's predictive capability was tested against an independent registry of >1,000 patients, and its performance compared to an established clinical risk score – the Electra Registry Outcome Score (EROS)[3].

We hypothesised that geometric features extracted from a pre-procedure plain anterior-posterior (AP) chest radiograph (CXR) would improve this ML model's predictive capability. In this study, we have the following aims: i) develop a deep learning (DL-) based, convolutional neural network (CNN) to automatically detect geometric features relevant to risk of TLE from pre-procedure CXRs, ii) evaluate if these features improve the previously published ML model's performance, and iii) assess whether the same methods could predict long procedure times (which acted as a proxy for procedure complexity).

Methods

The database collection and analysis were approved by the Institutional Review Board of Guy's and St Thomas' Hospital.

Study Populations and outcome definitions

A prospectively collected database of all consecutive patients undergoing TLE between October 2000 and November 2019 were recorded on a computer database at Guys' and St Thomas' NHS Foundation Trust (GSTT), a high-volume UK TLE centre was used for chest radiograph data. 1171 patients underwent TLE in this period. The intended outcome from the ML model was pre-procedural identification of patients at high risk of negative outcomes and complex procedures. The outcome of major adverse event – defined as a composite of procedure related major complication and procedure related death – was selected as the ground truth to capture this goal. These were defined in both datasets by the consensus statement in the ELECTRa dataset[3,8] (supplementary table 1). Each MAE was only counted once per patient (i.e. if a major complication and death were recorded for the same patient, this was only counted as 1 MAE). Prolonged procedure time, defined as time taken from procedure start to removal of targeted leads lasting >100 minutes. This cut-off was chosen as it was the 75th percentile of procedural duration (i.e., 25% of the cases were longer than 100 minutes).

Statistics

Categorical variables were compared with a Chi-squared test or Fisher's exact test. Normally distributed data was analysed using independent samples t-test. Non-normally distributed continuous data was analysed using the Kruskal-Wallis one-way analysis of variance test. The results are presented as mean±standard deviation for normally distributed variables and median [interquartile range (IQR)] for

non-normally distributed variables. Categorical variables are presented as number of patients (% of group). These analyses were performed using R version 1.3.1093.

The definitions of the performance measures used to describe the accuracy of the neural network to detect the geometric features, and the ML- model to predict MAE, are described in table 1. The Jaccard similarity coefficient (JSC) was used as a simple statistical measure for testing the correlation between an identified individual geometric feature on a CXR and the binary outcome of MAE[9]. As the dataset is known to be imbalanced[6], an Adaptive Synthetic (ADASYN) sampling technique was used to balance the datasets and calculate the JSC[10]. The relative importance of each feature in the final model is represented by an F-score, which is a way to rank the features based on their contribution to the final prediction, i.e. the higher the F-score value, the greater that particular feature contributes to the model's prediction of risk. The F-score for each feature should be considered in relation to the score of other features contributing to the model. Balanced accuracy, sensitivity, specificity, and receiver operating characteristic curves were used to assess the predictive capability of the ML-model.

Overview of original machine learning study

Previously, we have published a a ML-based risk stratification tool trained using the ELECTRa registry to predict risk of MAEs in 3555 patients undergoing TLE and tested this on an independent registry of 1171 patients [6]. ML models were developed, including a self-normalising neural network (SNN), stepwise logistic regression ("stepwise model"), support vector machines (SVM) and random forest (RF) model. These were compared to the Electra Registry Outcome Score (EROS) for MAEs [3]. ML techniques were similar to EROS by balanced accuracy (ML-model: 0.74 vs EROS: 0.70) and superior by area under the curve (ML-model: 0.764 vs EROS: 0.677). Whilst there was an improvement in accuracy, this was incremental, and hypothesised that integrating imaging biomarkers into this model would improve its predictive capability.

Rationale on radiograph feature selection

A combination of known risk factors and senior clinical opinion was considered when choosing appropriate features from a plain CXR. In previously published data, the presence of an ICD lead, an increased number of leads extracted[3,11], and likelihood of lead encapsulation [12] are established risk factors that could be detected from a plain CXR. The most common cause of procedure related major complication or death is a tear in the SVC[13] and therefore identifying accurate proxies for a model of the SVC and the location of leads and coils in relation to this was considered essential. It was also hypothesised that lead-to-lead interaction (measured by whether the leads were overlapping), or an acute (i.e., a <90 degree) angulation of the lead within the SVC or the myocardium were important factors to consider. An acute angle was defined as an angle between 0 and 90 degrees of the lead itself at the locations of the SVC or the right ventricle (supplementary figure 1).

These geometric features were combined with the baseline characteristics of the patients as illustrated in table 2 to feed into a ML model to predict risk of MAE. The model was based on the pre-procedural features in a previously published ML-model trained on the ELECTRa registry and tested on the GSTT database [6]. All features considered and included in the final ML models, including and excluding geometric features are in supplementary table 2. An XGBoost classifier was chosen due to its high performance with imbalanced datasets and binary classifications, and is an ensemble learning method that combines the predictions of multiple models to produce a stronger prediction [14]. Due to the relative infrequency of MAEs observed, leave-one-out cross validation was used[15]. For comparison, the added value of the geometric features was tested against the model using only the pre-procedure characteristics in table 2 (i.e., excluding the geometric features from the CXR).

Overview of imaging biomarker detection framework

The deep learning framework to automatically identify the geometric features has been previously been described in detail[16], a summary of the process is illustrated in figure 1 and outlined below:

I) Identification of the approximate location of the SVC

The SVC is not readily identifiable on a CXR, unless contrast is injected, however, it is possible to determine an approximate course in relation to heart anatomy. Using known common approximate models of heart anatomy, we were able to approximate the location of the SVC. The SVC location is approximately half the height and one third of the width of the heart region (supplementary figure 2). This was validated by overlaying 3D anatomy models extracted from 20 pre-procedure CT scans. In order to use this automated SVC location in a CXR, the heart region needed to be detected. A transfer learning approach was used to detect the heart region which was based on a modified VGG16 model. 70% of the CXRs were used for training and 30% were used for accuracy testing.

II) Detection of lead and coils

To segment the leads and coils from the CXRs, a U-Net convolutional network was trained and tested on the CXR dataset. To semi-automate this, a vessel enhancement filter[17] was used to extract all wire-like objects and the resultant image was automatically binarized. An experienced clinician manually removed features not leads or coils, which acted as the "ground truth" for the U-net model to train on 737 CXRs to automatically extract coils and leads, whilst ignoring other objects such as ECG leads, ribs, and the generator. To improve detection with poor quality images, the contrast or brightness of these images were reduced by a factor between 0.6 and 0.9 to create an additional 1474 training images.

III) Extracting geometric features

A centreline extraction method was used to precisely identify the lead or coil detected that was present in the SVC[18]. In this process, a binarized image of the detected coils and leads was skeletonised and a one-pixel wide object is created. A contour finding algorithm and wire reconstruction method was used to determine the course of the lead and compute location and angulation of the leads or coil, and determine if they were overlapping[19,20]. The two points of angulation interest were at the point of entry into the SVC or within the RV (supplementary figure 1). The ground truth was the manual labelling of all angles of the leads in two positions: the entry point of the SVC and inside the RV. Overlapping leads was defined as the centrelines of the leads overlapping with each other at least once.

Results

Baseline characteristics

The overall cohort included 1,151 patients, of which 1,050 had pre-procedure chest radiographs available for use in the current analysis (table 2). There were 24 (2.3%) MAEs, of which the most common major complication was cardiac avulsion (n=12, 1.2%) (supplement table 3). In total, there were 484 cases with complete procedure duration data available, of these 123 (25.4%) had a long procedure time. Overall, 73.5% (n=772) were male with mean age at explant of 65.3±14.5 years and mean LVEF of 44.8±14.3%. Median lead dwell time was 5.30 (1.80-10.00) years and the plurality of devices explanted were dual chamber pacemakers (n=370, 35.3%) followed by ICDs (n=297, 28.3%). In total 2,246 leads were explanted, with a mean of 2.14 leads per procedure. 443 defibrillator leads were explanted, of which 219 (49.4%) were dual coil leads. The majority of extractions were for infective indications (n=562, 54.1%), of which most were for local infections (n=372, 52.6%). The most common comorbidities were heart failure (n=414, 40.7%) followed by hypertension (n=406, 40.4%). In total 206 (20.1%) patients had CKD, with a median baseline creatinine of 92.00 [77-119.3] md/dl.

Automated detection performance

The performance of the DL neural network to automatically detect the geometric features is described in table 3. This was excellent in all areas with the exception for acute angle detection in the SVC.

Geometric feature selection and deep-learning model for risk assessment

The geometric features assessed and their respective JSC values are illustrated in table 4. 143 cases had an acute angle in the RV. Greater than 50% of a coil in the SVC combined with ≥2 overlapping leads in the SVC had a very sensitive JSC value, and resulted in a much higher performance and contribution overall to the model to predict MAE rather than separating the features (see supplementary table 4). Therefore, the following geometric features were included in the final model: i) >50% coil in the SVC and ≥2 overlapping leads in the SVC, and ii) acute angle in the RV.

Ability of model to predict major adverse event and long procedure time.

The most important features in the model excluding geometric features are dwell time, LVEF and eGFR (supplementary figure 3) for predicting MAE. The presence of a dual coil lead was only the 10th most important feature in the model excluding geometric features. In comparison, two geometric features were the 2nd and 12th most important features in in the model including the CXR data (figure 2A). The most important pre-procedural clinical factor remained lead dwell time, however the presence of ≥50% of a coil in the SVC and ≥2 overlapping leads had a similar predictive value for MAE as dwell time (F-score: 19.7 vs 18.7). An acute angle in the RV was almost as important as LVEF (F-score: 7.16 vs 8.53) and more important than male gender (F-score: 7.07). Infective factors were of high relevance (endocarditis F-score: 16.85, local infective indication F-score: 15.21).

There was a significant improvement in the model's ability to predict MAE with CXR data (table 5). Balanced accuracy improved from 0.75 to 0.87, sensitivity improved from 69% to 83%, specificity improved from 72% to 91%, and AUC improved from 0.767 to 0.962 (figure 2B). There was no significant difference in model performance when using the ground truth data compared to automated CXR feature detection (supplementary figure 4).

With respect to long procedure time, the most important pre-procedural clinical factors were endocarditis (F-score 12.95), respiratory disease (F-score: 11.77) and local infection (F-score: 7.01) (figure 3A). The most predictive feature was >50% of coil in the SVC and ≥2 overlapping leads from the pre-procedural CXR of (F-score: 48.11). Balanced accuracy improved from 0.76 to 0.86, sensitivity improved from 75% to 87%, specificity improved from 63% to 85%, and AUC improved from 0.684 to 0.913 (figure 3B).

We explored whether inclusion of lateral chest radiographs increased the predictive capability of the model, however there was no significant improvement, with balanced accuracy increasing from 0.91 to 0.92 and the AUC from 0.962 to 0.969 (see supplementary file 2).

Discussion

This is the first deep learning derived model integrating imaging biomarkers to predict risk following TLE. main findings are:

1. Convolutional neural networks can automatically, and accurately, detect key geometric features relevant to lead extraction, particularly the SVC location, the course of a lead, lead angulation, percentage of coil in the SVC, and whether leads overlap.
2. Imaging biomarkers integrated into an independently tested ML-model significantly improved its predictive capability.
3. ML can improve prediction of potentially lengthy procedures.

Relevance of risk stratification

Undertaking a lead extraction procedure with the appropriate expertise, tools and in the correct setting is one of the cornerstones of reducing procedural risk. Extensive data has shown that high volume lead extraction centres perform better compared with low volume centres and patients with known risk factors who have a lead extraction in a low volume centre have poorer outcomes reinforcing the fact that high volume operators should perform higher risk procedures[1,21]. TLE is a resource intensive and expensive procedure, which is often poorly reimbursed relative to their economic costs to the health institution[22]. Within this context, risk stratification scores are a helpful tool for the heart team when triaging patients appropriately pre-procedure.

There have been several published risk stratification tools, primarily using a combination of expert consensus and logistic analysis. Bontempi et al developed the "LED risk score", which identified number of extracted leads within a procedure, lead age, dual-coil lead and presence of vegetation as increasing the likelihood of a complex procedure[23]. Kancharla et al derived a risk score including an ICD lead >5 years old, and pacemaker lead >10 years old. The tool was prospectively validated on whether patients should have TLE performed in an operating room or device laboratory[4]. The EROS score developed by Sidhu et al [7] categorised patients as low (EROS 1), intermediate (EROS 2) and high risk (EROS 3). EROS 3 risk was heavily dependent on lead dwell time – a pacemaker lead >15 years, and ICD lead >10 years from implant. We previously compared the EROS score to a ML-based risk score trained on the ELECTRa database and tested on the dataset used in this study with an AUC of 0.764, however the study concluded that more nuanced and richer pre-procedure data was required to make ML a beneficial tool[6]. This current study has built on that conclusion and shown features extracted from a plain CXR – a simple investigation that should be performed in all patients undergoing lead extraction – significantly improves risk prediction. This is particularly marked in the geometric features related to the SVC and the congestion of leads within it. TLE has a high procedural success rate and complications remain low, so an excellent AUC is necessary

to provide the necessary sensitivity and specificity for added value. The AUC improvement from 0.76 to 0.96 in this study demonstrates such an improvement. An AUC score of between 0.9-0.99 is considered excellent[24].

Clinical perspectives

Better risk stratification tools have the potential to ensure stronger mitigation and minimise the risk of MAE. The value of such a risk tool may be useful for both infective and non-infective indications of lead extraction. For infective indication, there is always a class I indication to extract infected material[25], however it is not always clear who is likely to be at significantly higher risk of an adverse event. For non-infective indications, the evidence for lead extraction is less strong with clinical consensus and patient preference primarily determining whether a malfunctioning lead should be extracted or left redundant when a new lead is implanted. Currently, the data regarding the potential harms of abandoning leads in such circumstances is unclear, and weighing benefits and risks is a largely a joint decision between the cardiologist and the patient[26]. Having a highly sensitive and specific test would help quantify that risk better and allow the decision to be made more objectively. The current study provides the framework for such a test.

In addition, the ability of the current model to determine those at higher risk of a long procedure times has clinical utility. Traditional risk factors for complex procedures, in particular the presence of an ICD, markers of infection and respiratory disease, were strong influencers in the model. It is likely that these patients had more complex general anaesthetic requirements in view of their endocarditis and respiratory disease, and the presence of a coil increased the number of tools and time required to extract the material. These factors are important when determining resource and time allocation, and whether the procedure should occur in the operating room, hybrid or device laboratory.

Future perspectives

A previous study by Howard et al has shown the utility of neural networks to accurately identified the model of pacemakers and defibrillators[27]; however, this is the first deep learning framework to automatically detect CXR and lead related features increasing the risk of MAE following TLE. We have demonstrated that novel methods can improve the prediction of an adverse outcome by integrating a ubiquitous and inexpensive investigation, such as a CXR. All patients should have a CXR pre-TLE, and this data, combined with baseline clinical features, can be routinely integrated into pre-procedure planning. Such techniques can also potentially be used in cross-sectional imaging which have greater spatial resolution of the three-dimensional course of a lead and coil, particularly in relation to the SVC; however, contrast enhanced computerised tomography (CT) imaging using current deep learning techniques is challenging due to the high intensity of artefacts from the lead and coils. Real time fluoroscopy and motion tracking could also give added information on lead-to-lead or lead-to-SVC interaction giving a far greater understanding on how lead kinetics affect procedural risk [20].

Overall, the whole process is computationally inexpensive. The computer vision algorithm works within 500 milliseconds (ms) to identify the coil and lead route, as well calculating angles. The prediction algorithm (XGBoost) performs the prediction in 1000ms to produce the prediction once the relevant baseline demographic and imaging data are uploaded. We anticipate that this tool can be used before or during the "heart team" meeting, where a decision is made on the appropriateness and location of any lead extraction procedure.

Limitations

The high imbalance between the MAE (2.4%) and non-MAE (97.6%) cases means much larger datasets are required to increase the confidence in our results. To mitigate this, leave-one-out validation, and ML-algorithms (i.e., XG Boost and ADASYN sampling technique) which perform well with imbalanced data were used. Certain low frequency but traditional high-risk features of patients undergoing TLE, such as lead calcification were not included in the analysis as they were not consistently recorded. The risk

prediction model was applied retrospectively which may introduce bias, and a prospective validation study would mitigate this. Whilst the neural network to detect most geometric features was robust, it was sub-optimal in detecting acute angles in the SVC due to the low frequency of such cases, and this feature was therefore not included in the final model. A plain AP CXR was used, however for 3-dimensional distances between individual leads, a higher volume of lateral CXRs may be included in future analyses.

Conclusion

This is the first study to use imaging biomarkers to assess risk related to TLE. We have developed a high-performing automated deep learning algorithm to accurately extract geometric features from a plain anterior posterior chest radiograph. By integrating this imaging data to an existing ML-model, we have significantly improved prediction of adverse events and potentially complex procedures related to TLE.

References

1. Pelargonio G, Narducci ML, Manzoli L, et al.: Safety of transvenous lead extraction according to centre volume: A systematic review andmeta-analysis. Europace 2014; 16:1496–1507.
2. Mehta VS, Elliott MK, Sidhu BS, et al.: Long-term survival following transvenous lead extraction: Importance of indication and comorbidities. Heart Rhythm 2021; 18:1566–1576.
3. Bongiorni MG, Kennergren C, Butter C, et al.: The European Lead Extraction ConTRolled (ELECTRa) study : a European Heart Rhythm Association (EHRA) Registry of Transvenous Lead Extraction Outcomes. 2017; :2995–3005.
4. Kancharla K, Acker NG, Li Z, et al.: Efficacy and Safety of Transvenous Lead Extraction in the Device Laboratory and Operating Room Guided by a Novel Risk Stratification Scheme. JACC Clin Electrophysiol [Internet] Elsevier, 2019; 5:174–182. Available from: https://doi.org/10.1016/j.jacep.2019.01.001
5. Jacheć W, Polewczyk A, Polewczyk M, Tomasik A, Kutarski A: Transvenous Lead Extraction SAFeTY Score for Risk Stratification and Proper Patient Selection for Removal Procedures Using Mechanical Tools. J Clin Med 2020; 9:361.
6. Mehta VS, O'Brien H, Elliott MK, et al.: Machine learning–derived major adverse event prediction of patients undergoing transvenous lead extraction: Using the ESC EHRA EORP European lead extraction ConTRolled ELECTRa registry. Heart Rhythm Elsevier B.V., 2022; 19:885–893.
7. Sidhu BS, Ayis S, Gould J, et al.: Risk stratification of patients undergoing transvenous lead extraction with the ELECTRa Registry Outcome Score (EROS): an ESC EHRA EORP European lead extraction ConTRolled ELECTRa registry analysis. EP Europace [Internet] 2021; 23:1462–1471. Available from: https://doi.org/10.1093/europace/euab037
8. Maytin M, Jones SO, Epstein LM: Long-term mortality after transvenous lead extraction. Circ Arrhythm Electrophysiol United States, 2012; 5:252–257.
9. Chung NC, Miasojedow BZ, Startek M, Gambin A: Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. BMC Bioinformatics BioMed Central Ltd., 2019; 20.

10. Institute of Electrical and Electronics Engineers: Neural Networks, 2008, IJCNN 2008, (IEEE World Congress on Computational Intelligence), IEEE International Joint Conference on : date, 1-8 June 2008.
11. Segreti L, Rinaldi CA, Claridge S, et al.: Procedural outcomes associated with transvenous lead extraction in patients with abandoned leads: An ESC-EHRA ELECTRa (European Lead Extraction ConTRolled) Registry Sub-Analysis. Europace 2019; 21:645–654.
12. Kołodzińska A, Kutarski A, Koperski Ł, Grabowski M, Małecka B, Opolski G: Differences in encapsulating lead tissue in patients who underwent transvenous lead removal. Europace Oxford University Press, 2012; 14:994–1001.
13. Tułecki Ł, Polewczyk A, Jacheć W, et al.: A study of major and minor complications of 1500 transvenous lead extraction procedures performed with optimal safety at two high-volume referral centers. Int J Environ Res Public Health MDPI, 2021; 18.
14. Chen T, Guestrin C: XGBoost: A Scalable Tree Boosting System. 2016; . Available from: http://arxiv.org/abs/1603.02754
15. Sengupta PP, Shrestha S, Berthon B, et al.: Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. JACC Cardiovasc Imaging. Elsevier Inc., 2020, pp. 2017–2035.
16. Ma, Y., Mehta, V.S., Rinaldi, C.A., Hu, P., Niederer, S., Razavi, R. (2023). Automatic Detection of Coil Position in the Chest X-ray Images for Assessing the Risks of Lead Extraction Procedures. In: Bernard, O., Clarysse, P., Duchateau, N., Ohayon, J., Viallon, M. (eds) Functional Imaging and Modeling of the Heart. FIMH 2023. Lecture Notes in Computer Science, vol 13958. Springer, Cham. https://doi.org/10.1007/978-3-031-35302-4_32
17. Frangi AF, Niessen WJ, Vincken KL, Viergever MA: Multiscale vessel enhancement filtering. In Wells WM, Colchester A, Delp S, eds: Medical Image Computing and Computer-Assisted Intervention — MICCAI'98 Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 130–137.

18. Ma YL, Alhrishy M, Narayan SA, Mountney P, Rhode KS: A novel real-time computational framework for detecting catheters and rigid guidewires in cardiac catheterization procedures. Med Phys John Wiley and Sons Ltd, 2018; 45:5066–5079.

19. Suzuki S, be K: Topological structural analysis of digitized binary images by border following. Comput Vis Graph Image Process [Internet] 1985; 30:32–46. Available from: https://www.sciencedirect.com/science/article/pii/0734189X85900167

20. Ma Y, Zhou D, Ye L, Housden RJ, Fazili A, Rhode KS: A Tensor-Based Catheter and Wire Detection and Tracking Framework and Its Clinical Applications. IEEE Trans Biomed Eng IEEE Computer Society, 2022; 69:635–644.

21. Sidhu BS, Gould J, Bunce C, et al.: The effect of centre volume and procedure location on major complications and mortality from transvenous lead extraction: an ESC EHRA EORP European Lead Extraction ConTRolled ELECTRa registry subanalysis. Europace. 2020 Nov 1;22(11):1718-1728. doi: 10.1093/europace/euaa131. PMID: 32688392.

22. Gould J, Sidhu BS, Porter B, et al.: Financial and resource costs of transvenous lead extraction in a high-volume lead extraction centre. Heart 2020; l:931–937.

23. Bontempi L, Vassanelli F, Cerini M, et al.: Predicting the difficulty of a lead extraction procedure: The LED index. Journal of Cardiovascular Medicine 2014; 15:668–673.

24. Carter J V., Pan J, Rai SN, Galandiuk S: ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery (United States) Mosby Inc., 2016; 159:1638–1645.

25. Kusumoto FM, Schoenfeld MH, Wilkoff BL, et al.: 2017 HRS expert consensus statement on cardiovascular implantable electronic device lead management and extraction. Heart Rhythm Elsevier B.V., 2017; 14:e503–e551.

26. Poole JE, Gleva MJ, Mela T, et al.: Complication rates associated with pacemaker or implantable cardioverter-defibrillator generator replacements and upgrade procedures: Results from the REPLACE registry. Circulation 2010; 122:1553–1561.

27. Howard JP, Fisher L, Shun-Shin MJ, et al.: Cardiac Rhythm Device Identification Using Neural Networks. JACC Clin Electrophysiol Elsevier Inc, 2019; 5:576–586.

28. García V, Mollineda RA, Sánchez JS: Index of balanced accuracy: A performance measure for skewed class distributions. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2009; 5524 LNCS:441–448.

Figures

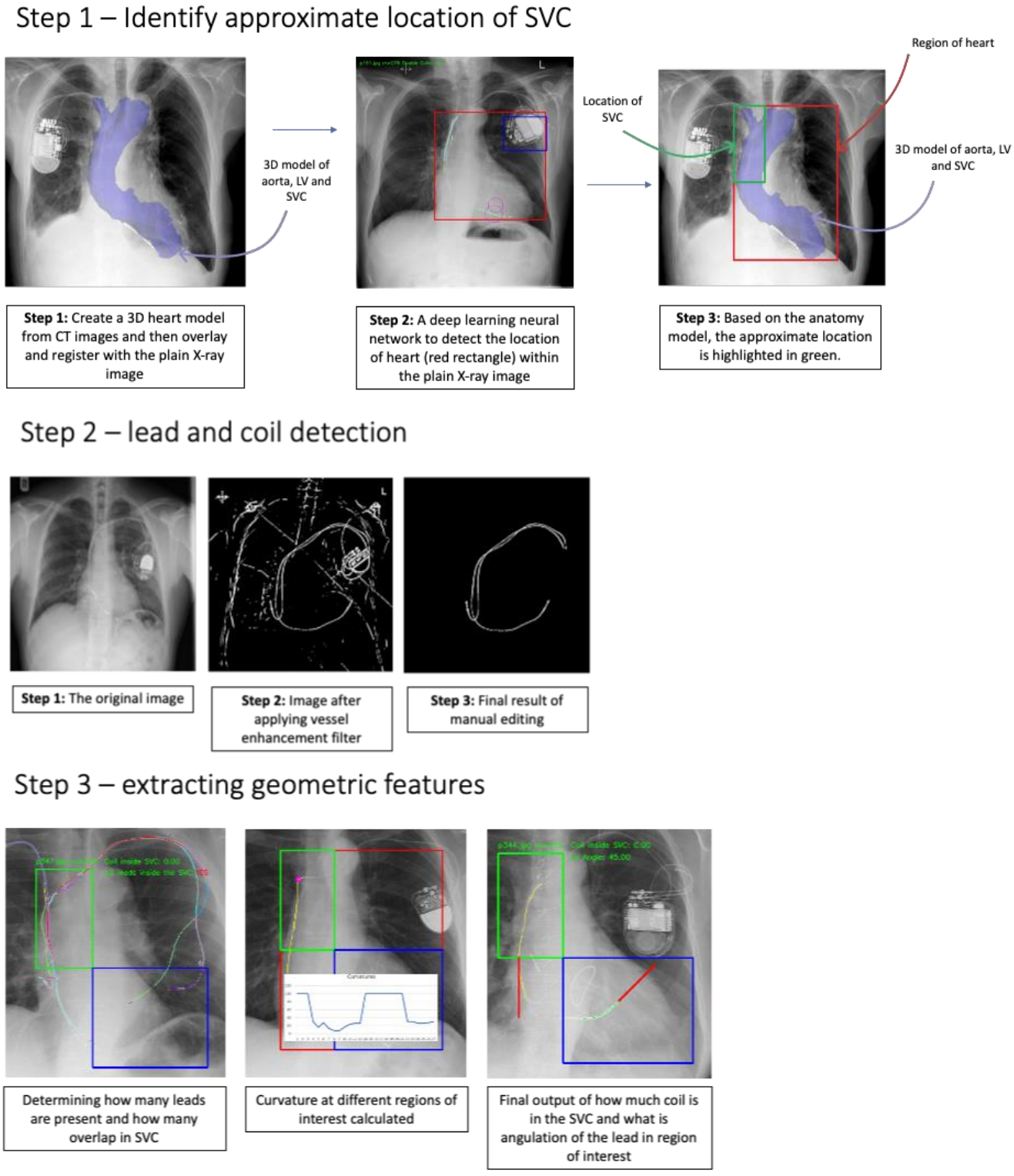
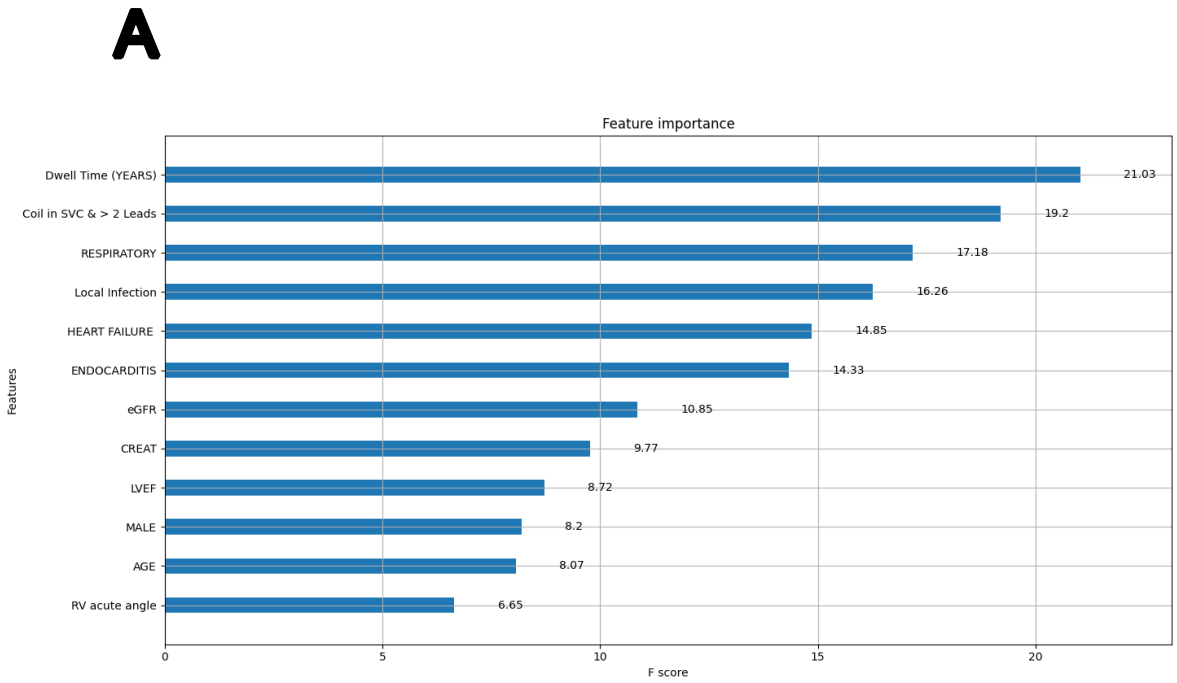


Figure 1: An overview of the deep learning framework to identify geometric features on a plain anterior-posterior chest radiograph.



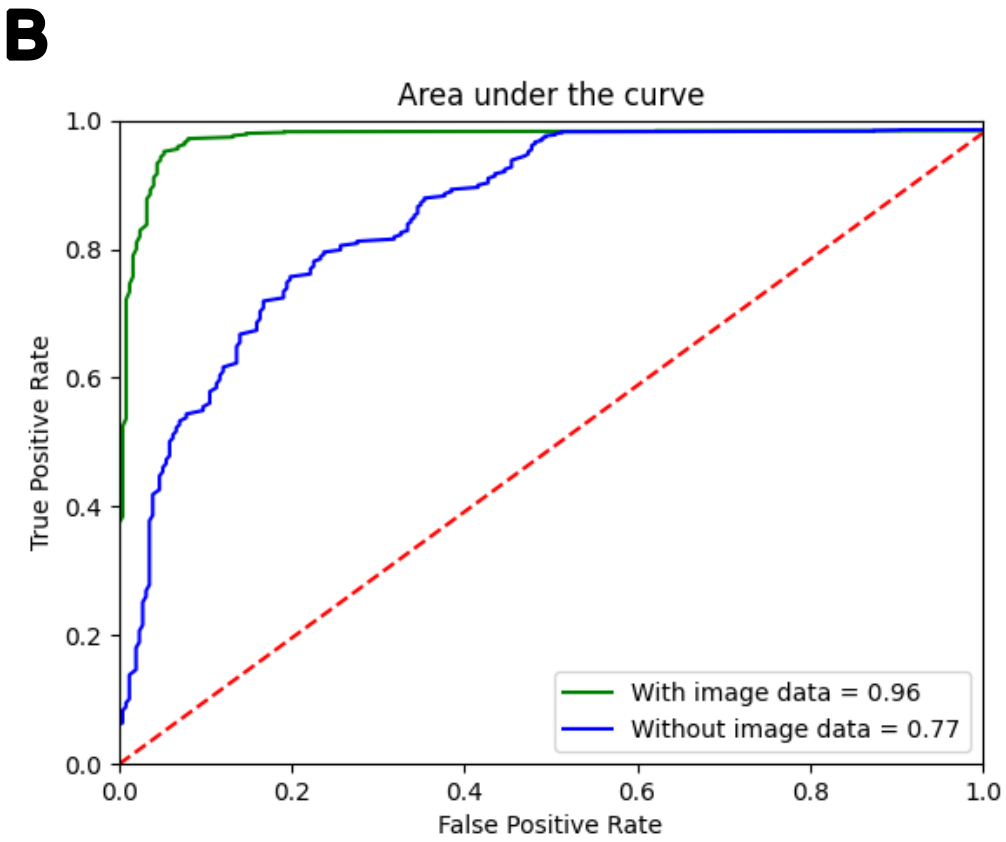
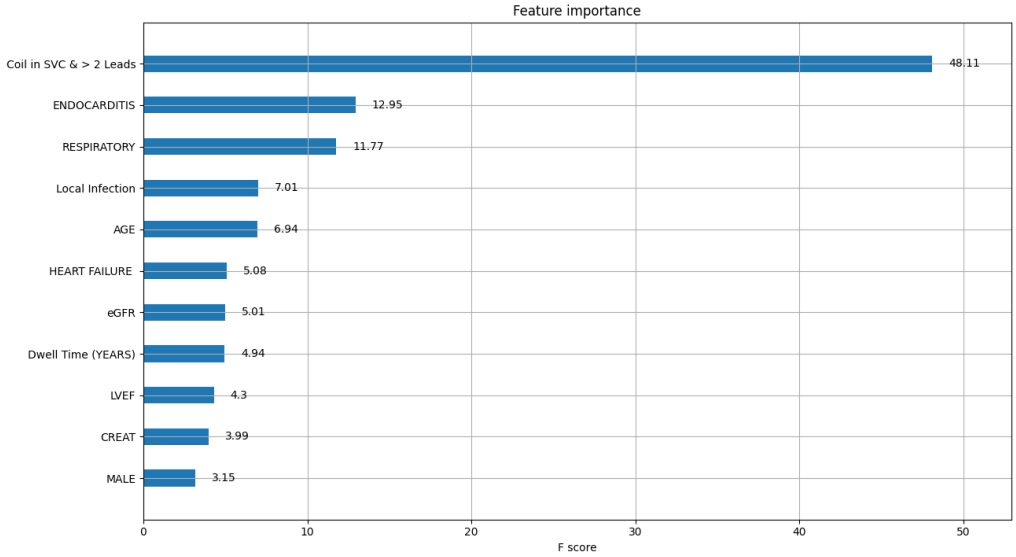


Figure 2: Machine learning model for predicting MAE. A: Features and their respective weights for predicting MAE. B: The receiver operating characteristic (ROC) curve of the ML model for predicting major adverse event (MAE). Blue – the ROC curve using a model without imaging data. Green – the ROC curve using a model with imaging.

A



B

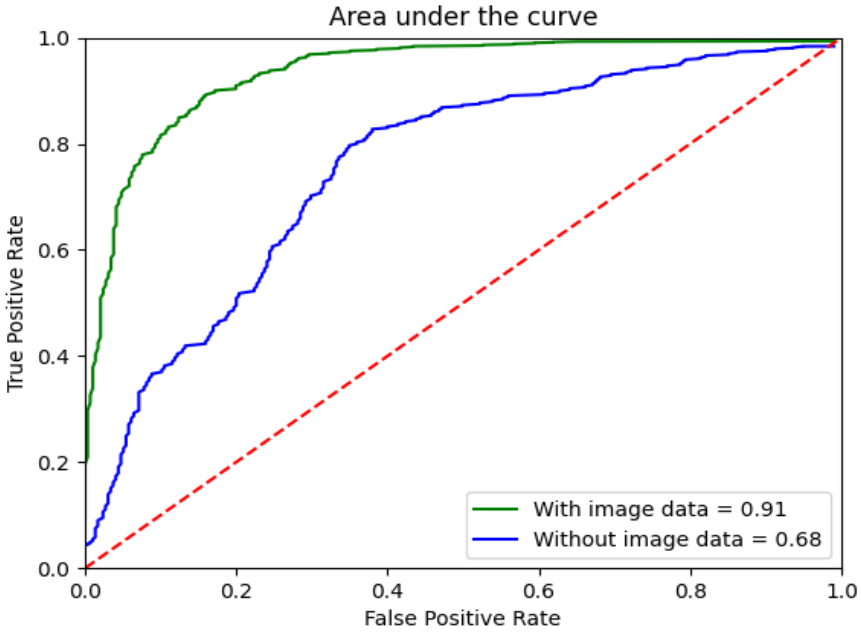


Figure 3: Machine learning model for predicting long procedure times. A: Features and their respective weights for predicting MAE. B: The receiver operating characteristic (ROC) curve of the ML model for predicting major adverse event (MAE). Blue – the ROC curve using a model without imaging data. Green – the ROC curve using a model with imaging.

Tables

Table 1: Performance measures used to describe results of the neural network and machine learning model to predict major adverse event (MAE).

Performance measure	Definition	Formula
Balanced Accuracy	Balanced accuracy, rather than total accuracy, was utilised as the primary performance measure in view of the skewed distribution of outcomes with no MAEs vs those with MAEs – i.e., there were very few cases involving MAEs in both datasets[28]. Using accuracy alone, would over-estimate the ability of the risk stratification method in appropriately identifying the patient risk. Balanced accuracy is the mean accuracy for each class.	$\frac{TPR + TNR}{2}$
Precision	Precision is the ability of a model to identify only the relevant objects – i.e., precision is how good the model is at accurately identifying a geometric feature	$\frac{TP}{TP+FP} = \frac{TP}{\text{All detections}}$
Recall	Recall is the ability of a model to find all the relevant cases – i.e., how many times the model was able to accurately identify a geometric feature	$\frac{TP}{TP+FN} = \frac{TP}{\text{All ground truths}}$
F1 Score	The F1 score is a method of calculating a mean of precision and recall.	$2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$

Jaccard Similarity Coefficient	JSC is a statistical measure for correlating the similarity between binary data samples (defined as sample A and B). M_{11} represents the total number of attributes where sample A and B both have a value of 1. M_{01} represents the total number of attributes where the attribute of A is 0 and the attribute of B is 1. M_{10} represents the total number of attributes where the attribute of A is 1 and the attribute of B is 0. It represents the percent of characteristics found in both samples and found in only one sample, i.e., a Jaccard Similarity of 0.3 means 30% of the characteristics were found in both samples, and 70% were found in only one of the two samples.	$JSC = \frac{M_{11}}{M_{01} + M_{10} + M_{11}}$

TPR – True Positive Rate. TNR – True Negative Rate. TP – True Positive. FP – False Positive. FN – False Negative.

Table 2: Baseline characteristics of all subjects in test dataset.

	Overall
n	1050
Demographics	
Male (%)	772 (73.5)

Explant Age in Years (mean (SD))	65.30 (14.49)
Device type	
Single chamber PPM (%)	74 (7.1)
Dual Chamber PPM (%)	370 (35.3)
Implantable Cardioverter Defibrillator (%)	297 (28.3)
CRT-Pacemaker (%)	67 (6.4)
CRT-Defibrillator (%)	240 (22.9)
Leads	
Dwell Time in Years (median [IQR])	5.30 [1.80, 10.00]
Active RV lead (%)	546 (24.4)
Passive RV lead (%)	207 (19.7)
Active RA lead (%)	534 (50.9)
Passive RA lead (%)	201 (19.1)
LV lead (%)	234 (22.3)
Single Coil Defibrillator Lead (%)	224 (21.3)
Dual Coil Defibrillator Lead (%)	219 (20.9)
Total number of leads extracted (%)	
1	294 (28.0)
2	444 (42.3)
3	216 (20.6)
4 or more	96 (9.1)
Side of explant	
Left sided (%)	856 (81.5)
Right sided (%)	131 (12.5)
Both sides (%)	63 (6.0)

Extraction Indication	
Non-infective Indication (%)	482 (45.9)
Local Infection (%)	372 (35.5)
Systemic Infection (%)	196 (18.7)
Extraction History	
History of Previous Extraction (%)	118 (11.2)
Co-Morbidities	
IHD (%)	402 (39.6)
Valvular Disease (%)	100 (9.9)
Heart Failure (%)	414 (40.7)
Diabetes Mellitus (%)	174 (17.3)
Hypertension (%)	406 (40.4)
Respiratory Disease (%)	145 (14.4)
CKD (%)	206 (20.1)
ESRF (%)	10 (1.0)
Cardiac Function	
LVEF (mean (SD))	44.77 (14.34)
Biochemistry	
Creatinine (μmol/L) (median [IQR])	92.00 [77.00, 119.25]
eGFR (ml/min/1.73 m2) (mean (SD))	67.47 (20.75)
eGFR<60 (ml/min/1.73 m2) (%)	391 (37.2)
Peak CRP (mg/L) (median [IQR])	6.00 [2.00, 18.00]

505

506 *CRP – C-reactive protein. eGFR – estimated glomerular filtration rate. LVEF – left ventricular ejection fraction.*

507 *ESRF – end stage renal failure. CKD – chronic kidney disease. IHD – ischaemic heart disease. LV – left ventricle. RA*

508 *– right atrium. RV – right ventricle. CRT – cardiac resynchronisation therapy. PPM – permanent pacemaker. IQR –*

509 *Interquartile Range.*

510

511 **Table 3:** The performance of the neural network to detect the feature on the CXR in comparison to the

512 ground truth. * Ability to detect >75% of the true heart border

513

Geometric Feature	Accuracy	Precision	Recall	F1 score
Detection of heart border*	NA	0.818	1	NA
Detection of coils	0.98	0.98	0.99	0.98
Detection of leads	0.782	0.879	0.836	0.857
Detect acute angle in SVC	0.7	0.833	0.714	0.769
Detect acute angle in RV	0.905	0.969	0.923	0.945

514

515 *SVC – superior vena cava. RV – right ventricle.*

516

517

518 **Table 4:** The geometric feature detected and its respective Jaccard Similarity Coefficient (JSC)

519

Geometric Features	**JSC**
50% of coil and ≥2 overlapping leads inside the SVC	0.35
The acute angle in the RV	0.03
The acute angle near the entry point of the SVC	0.01

520

521 *SVC – superior vena cava. RV – right ventricle.*

522

523

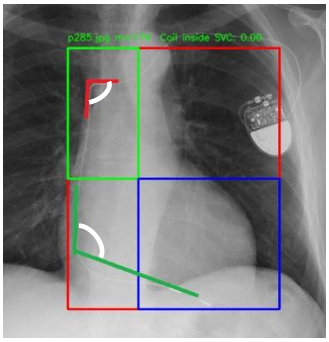
524 **Table 5:** The predictive capability of the machine learning model to predict a major adverse event or
525 long procedure excluding and including the geometric features from the pre-procedure chest radiograph.
526

Outcome	Results	Balanced accuracy	Specificity	Sensitivity	Area Under Curve
Major Adverse Event	Excluding Imaging	0.74	0.63	0.62	0.77
	Including Imaging	*0.87*	*0.91*	*0.83*	*0.96*
Long procedure time	Excluding Imaging	0.76	0.63	0.75	0.684
	Including Imaging	*0.86*	*0.87*	*0.85*	*0.913*

527
528
529

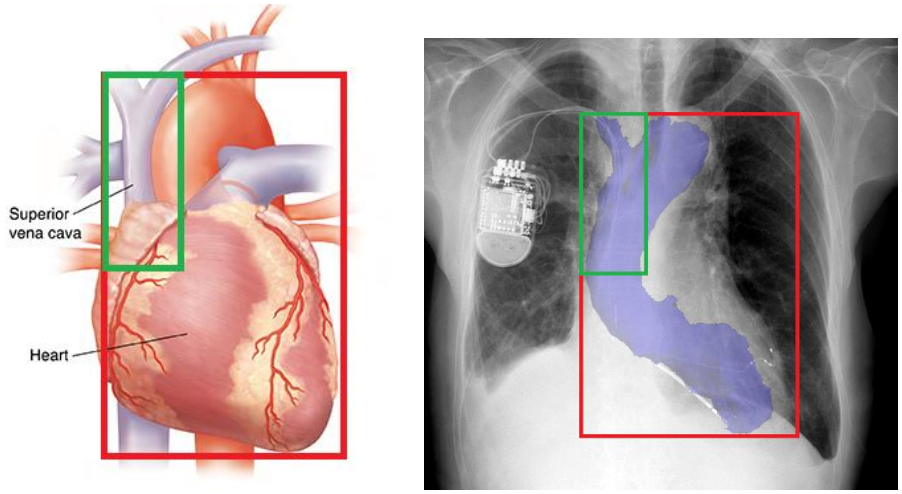
Supplemental Materials

Supplementary figure 1: Angle location



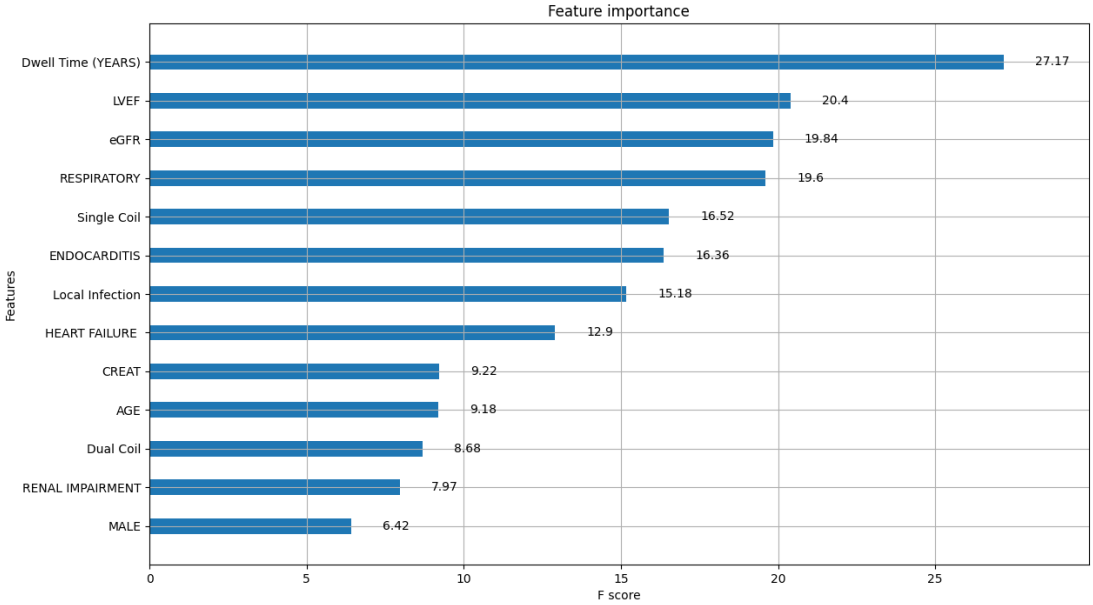
The red angle is the angulation of the lead near the entry point of the SVC and the green lead angle is the one within the RV.

Supplementary figure 2: (a) A medically accurate illustration of the anatomy of the heart. The red box is the region of the heart, and the green box is the location of the SVC. (b) Overlaying 3D anatomy models with the X-ray image. The blue shadow is the 3D model of aorta, left ventricle and the SVC. The red box is the region of the heart, and the green box is the location of the SVC.

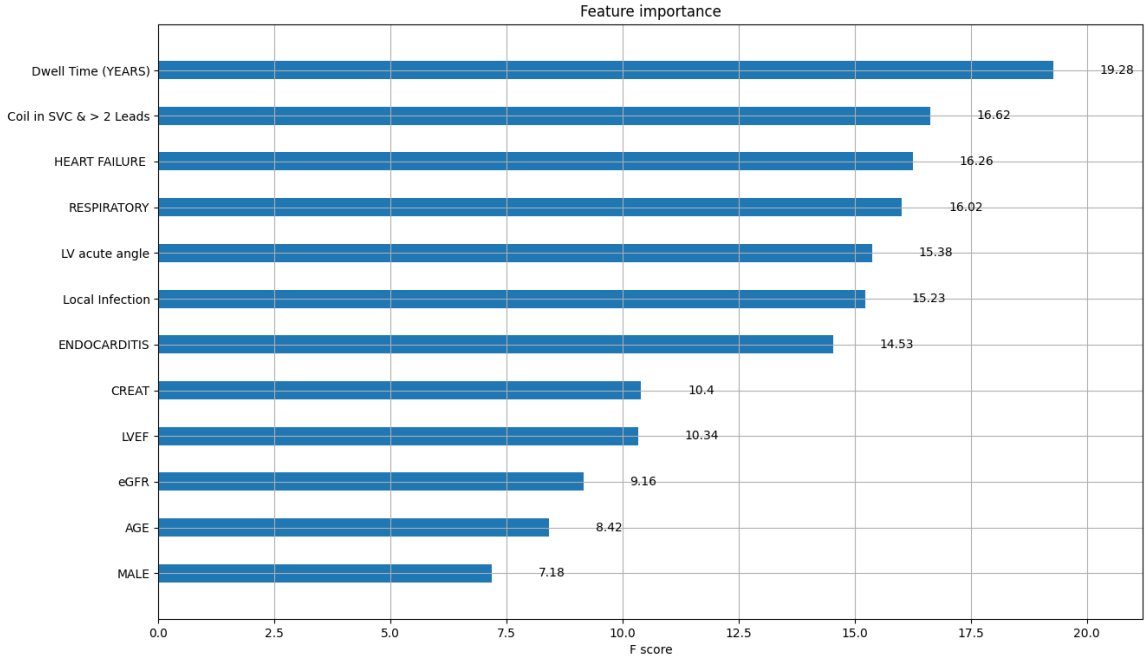


A B

Supplementary Figure 3: The features for the model excluding imaging data for predicting MAE.



Supplementary Figure 4: The features for the model excluding imaging data for predicting MAE using the ground truth data.



(a) Using ground truth data

Supplement Table 1: classification of complications

Classification of Complication	Examples
Major Complication	Death
	Cardiac avulsion or tear requiring thoracotomy, pericardiocentesis, chest tube, or surgical repair
	Vascular avulsion or tear (requiring thoracotomy, pericardiocentesis, chest tube, or surgical repair)
	Pulmonary embolism requiring surgical intervention
	Respiratory arrest or anaesthesia related complication leading to prolongation of hospitalization
	Stroke
	Pacing system related infection of a previously non-infected site
Minor Complication	Pericardial effusion not requiring pericardiocentesis or surgical intervention
	Haemothorax not requiring a chest tube
	Hematoma at the surgical site requiring reoperation for drainage
	Arm swelling or thrombosis of implant veins resulting in medical intervention
	Vascular repair near the implant site or venous entry site
	Hemodynamically significant air embolism
	Migrated lead fragment without sequelae
	Blood transfusion related to blood loss during surgery
	Pneumothorax requiring a chest tube
	Pulmonary embolism not requiring surgical intervention

Supplement Table 2: Pre-implant features considered in machine learning model excluding and including geometric features from chest radiographs.

X – not included in final model. ✓ - was included in final model.

All Features Considered	Included in ML Model (excluding geometric features)	Included in ML Model (including geometric features)
Age at Explant	✓	✓

Male Sex	✓	×
Lead dwell time (Months)	×	×
Lead dwell time (years)	✓	✓
Single Coil Defibrillator Lead	✓	✓
Dual Coil Defibrillator Lead	✓	✓
Local Infection (Infected Device/Erosion)	✓	✓
Systemic infection (Sepsis/Infective Endocarditis)	X	X
Creatinine (mcg/dL)	×	×
C-Reactive Protein (mg/dL)	X	×
eGFR (ml/min/m2)	✓	✓
History of Previous Extraction	X	×
Left Ventricular Ejection Fraction	✓	✓
Diabetes	×	×
Coronary Artery Disease	×	×
Heart Failure	✓	✓
Renal Impairment	✓	×
Chronic Respiratory Disease	✓	✓
Hypertension	×	×
Coronary Artery Bypass Grafting	×	×
Ischaemic Heart Disease	×	×
Peripheral Vascular Disease	×	×
Cerebrovascular Accident	×	×
Cumulative Co-Morbidity Burden	×	×

Total Leads Extracted		×	×
Device Type		×	×
Device Indication		×	×
Pacing lead type extracted - "Atrial Passive", "Atrial Active", "RV Passive", "RV Active", "LV Lead"		×	×
Non-infective extraction indication - "Lead Displacement", "Lead Failure", "Lead Malfunction", "Venous stenosis", "Pain", "Functional Lead Removal", "Other", "Sprint Fidelis lead extraction"		×	×
Chest radiograph feature	Acute RV lead angle	X	✓
	>50% coil in SVC	X	✓
	Acute angle near SVC entry point	X	✓

565

566 *Supplement Table 3: Categorisation of procedure related major adverse events (MAEs).*
567

Major adverse events	**Number (%)**
Death	1 (0.1)
Cardiac avulsion or tear requiring thoracotomy, pericardiocentesis, chest tube, or surgical repair	12 (1.2)
Vascular avulsion or tear (requiring thoracotomy, pericardiocentesis, chest tube, or surgical repair)	10 (1.0)
Stroke	2 (0.2)
Pacing system related infection of a previously non-infected site	2 (0.2)

568

Note – Please note that the total number of major adverse events (MAEs) is greater than total number of cases with MAEs in the analysis, as events are adjusted and only counted once, i.e. a patient may have had two MAEs from the lead extraction procedure, however this is counted as one MAE.

Supplementary table 4: Table outlining the relative contribution of the features to the model's predictive algorithm, and the performance of the model for major adverse events.

Outcome	Feature	F-score	Balanced accuracy	Specificity	Sensitivity	Area Under Curve
Major Adverse Event	Combined feature of ≥2 overlapping leads and >50% coil in the SVC	*19.2*	*0.87*	*0.91*	*0.83*	*0.96*
	≥2 overlapping leads only	*2.4*	*0.8*	*0.78*	*0.79*	*0.798*
	>50% of coil in SVC only	*8.3*	*0.83*	*0.83*	*0.81*	*0.868*

SVC – Superior Vena Cava