Big Data in cardiac surgery: real world and perspectives

Andrea Montisci1*, Vittorio Palmieri2, Maria Teresa Vietri3, Silvia Sala4, Ciro Maiello2, Francesco Donatelli5,6 and Claudio Napoli7

Abstract
Big Data, and the derived analysis techniques, such as artificial intelligence and machine learning, have been considered a revolution in the modern practice of medicine. Big Data comes from multiple sources, encompassing electronic health records, clinical studies, imaging data, registries, administrative databases, patient-reported outcomes and OMICS profiles. The main objective of such analyses is to unveil hidden associations and patterns. In cardiac surgery, the main targets for the use of Big Data are the construction of predictive models to recognize patterns or associations better representing the individual risk or prognosis compared to classical surgical risk scores. The results of these studies contributed to kindle the interest for personalized medicine and contributed to recognize the limitations of randomized controlled trials in representing the real world. However, the main sources of evidence for guidelines and recommendations remain RCTs and meta-analysis. The extent of the revolution of Big Data and new analytical models in cardiac surgery is yet to be determined.

Keywords: Big Data, Cardiac surgery, Artificial intelligence, Machine learning, Coronary revascularization, Valvular heart diseases, Heart failure, Left ventricular assist devices

Introduction
The combination of data coming from multiple sources, and constituting databases, with significant possibility of integration and complex aggregation and discriminant analyses, defines the so-called Big Data, which intrinsically refers to extensive datasets, widely informative for a large number and variety of persons.

In fact, large and rapidly increasing amount of data (Volume), their multiple sources (i.e., clinical studies, registries, small database, administrative database, patient-reported outcomes, genomic profiles and environmental parameters) (Variety), the rapid data accumulation (Velocity) and their ability to truly represent a specific context (Veracity) characterize Big Data [1].

Analyses of these data may unveil patterns, trends and associations, and define reference models in aggregations of persons [2]. As data digitalization and information technology (IT) are spreading and improving performance [3–5], the use of Big Data is steeply increasing, becoming progressively a reference for many typical processes in medicine, such as the identification of the appropriate therapeutic choice by tailoring therapeutic options, the evaluation of short and mid-term, procedure-related or unrelated, risks of adverse events and the definition of the prognosis. To such an extent, Big Data may generate the basis for precision medicine, as factors impacting event occurrence is progressively available for the single subject and may improve effectiveness of cardiovascular therapies [6].

The place of Big Data is far from being well determined. Figure 1 offers a graphical synthesis of the present and future of Big Data.
Big Data: sources and analysis

The key points of the application of Big Data are the clinical usefulness and the balance between costs of sophisticated data analyses and the expected and real benefits (i) to patients, in terms of quality of care, outcomes and risk prediction; (ii) to operators, in terms of quality and security of processes, from the diagnosis to the choice of therapeutic options, in a perspective of resources saving.

Regarding the sources, Electronic Health Record (EHR) is the main source of Big Data. Administrative data, which are commonly employed for billing purposes in fee-for-service health systems, may turn helpful for a large spectrum of analytic goals, and may fuel risk modeling for clinical and economical purposes [7]. Several approaches are used to enable data aggregation from EHR and to facilitate their contribution to Big Data. The most relevant limitations of the use of these databases are the risk of misclassification and the impact of missing data [8].

Digital imaging significantly contributes to Big Data generation. Today, almost all medical images are stored in pixels or voxels [9], which can be processed by software aiming at improving data quality and diagnostic accuracy [10].

Finally, OMICS datasets, or genome, proteomic, transcriptomic, epigenomic, and metabolomic data, are readily available in digital and structured forms, allowing the recognition of patterns useful for grouping procedures; they are additional important sources for Big Data [11].

In cardiovascular medicine, the contribution of OMICS is enormous and hold a very high potential for Big Data generation and subsequent analysis. A valid example of the OMICS data analysis is that it is able to provide information on myocardial molecular profiles of cardiac surgical patients [12–14]. The variants identification by OMICS techniques allows performing association studies, which may turn useful to risk stratification and outcome prediction in the context of precision medicine [15–19], using appropriate systems for analyses [20–22].

When it comes to analyses of Big Data, classic statistics is limited. Independent storage systems, immediate access and relational databases procedures are crucial for Big Data analyses. Artificial intelligence (AI) is of a particular usefulness, conceived as computer and mathematical concept allowing “machines” to execute learning, problem-solving, patterns recognition, reasoning and planning, in a way that resembles “human thinking”, therefore dealing with uncertainty, projection, and production that are well beyond extrapolating regression models validated in subsets of the general population, called validation samples, to approximate prediction in the general population.

The ability of such a technology to process decisions independently, recognize errors and readjust the decision or prediction models and processes, are the main characteristics of the AI, based on machine learning (ML) and deep learning (DL).

ML is the study of a mathematical algorithm model from sample data which in turn is used to generate predictions or decisions. ML algorithms can be supervised, unsupervised and reinforced [23]. DL is composed of artificial neural networks, with representation learning, to mimic human cognition. DL could be a module to automate predictive analysis, from which data is deduced in a non-linear way. The advantage of a non-linear interpretation is the better ability to identify and interpret more complex characteristics [24] and therefore is linked to a hierarchy of increasing complexity and abstraction [25]. DL is used for image evaluation, such as cardiac magnetic resonance scans; this requires adequate skills and systems [26, 27].

Logistic regression (LR) is a classic classification algorithm that makes a linear combination of input variables and uses the sigmoid function to output a probability.

Neurons in artificial neural networks (ANN) make a linear combination of the output value from the upper layers’ neurons, pass it through sigmoid functions, and finally output a value to the next neurons [28].

The use of predictive models that evaluate the influence of covariates, in the prediction of the results, allows it to identify the patients for whom the intervention will be successful. However, in analyzing non-randomized diagnostic or therapeutic strategies, it is possible to compare non-similar groups, exposing patients to subsequent complications [2].
The present of Big Data in cardiac surgery and the road ahead

While procedures safety and outcomes in cardiac surgery have improved over the years in the majority of elective procedures, cardiac surgery is facing patients with increased complexity due to improved survival owing to refined cardiological, pneumological and oncological therapies. Changes in such a clinical context require rethinking clinical risk assessment and management as well as redefinition of optimal timing for surgical options.

So far, cardiac surgery has relied on logistic models to estimate the risk of events, including mortality associated with cardiac surgery, as essential components of routine clinical management of cardiacl surgical patients. The EUROscore II and the Society of Thoracic Score (STS) scores are the logistic models for cardiac surgery-related risk stratification most commonly employed. Nevertheless, the prediction of those estimates is debated, especially in subsets of patients [29]. AI applied to Big Data has the potential to change the paradigm, from a theoretical and average risk prediction to simulations in single patients, weighting tailored therapeutic options and managing risks, and finally employing sustainable options to improve outcomes.

This is the most appealing application of the new concept of the meta-verse, an aggregate of Big Data, information technologies and AI converging to generate novel approaches to handle reality by navigating virtual near-future in clinical contexts, hopefully yielding time saving, less errors, more precision, variability due to operators, minimize costs and human effort while prolonging life with reasonable quality.

To such an extent, Big Data and AI have been applied in seminal studies in cardiac surgery in the field of myocardial revascularization, valvular heart diseases and end-stage heart failure (Table 1). Table 2 shows the ongoing trials focused on the application of Big Data-derived analysis in cardiac surgery.

Myocardial revascularization

In the field of surgical myocardial revascularization, structured data from EHR managed by the Society of Thoracic Surgeons, the American College of Cardiology (National Cardiovascular Data Registry), and the American Heart Association, represent a significant source for data analysis. Clinical records, and data from imaging, analyzed and interpreted using AI approaches, may integrate original approaches based on clinical registries and regression models (e.g.: neural networks) [9, 30].

One of the major challenging research tasks, to date, is to evaluate whether percutaneous coronary interventions (PCI) are superior over coronary artery bypass surgery (CABG) in specific clinical contexts. Randomized trials (RCTs) on myocardial revascularization have been used to answer this question, and to develop a tool for the identification of patients who may benefit from one therapeutic option over the other or a combination of both (hybrid revascularization), with minimized clinical risk and expenditure. Although RCTs are very effective to control treatment selection bias, they are low-performing in evaluating subgroups and they suffer from inappropriate statistical power in subset and possible post-hoc biases. Observational, nonrandomized data gathered from registries and large multi-sources databases, may be closer to real-world representing a very large majority of patients, and therefore may work better with single patients using AI based analytic modalities. In contrast, those sources of data may be affected by lower data control and potential biases in the outcome definition and assignment [31]. Weintraub and colleagues [32, 33] compared the effectiveness of different myocardial revascularization strategies by linking the American College of Cardiology Foundation (ACCF) National Cardiovascular Data Registry and the Society of Thoracic Surgery (STS) Adult Cardiac Surgery Databases, to data from claims from the Centers of Medicare and Medicaid. They demonstrated that the real-world mortality was not significantly different at 1-year from that anticipated by commonly used scores, while long-term survival was higher in patients receiving CABG as compared with patients who underwent PCI. The method of analyses of these data required the use of probabilistic matching to identify patients throughout databases, adjusting for clinical covariates with the use of inverse probability weighting, and correction of residual confounding by means of a sensitivity analysis.

A valid example of cost-effective applicability of AI to Big Data in the field of myocardial revascularization is the ASCERT study (ACCF and STS Database Collaboration on the Comparative Effectiveness of Revascularization Strategies) [34], where two revascularization strategies (PCI versus CABG) were evaluated in patients suffering from stable ischemic heart disease. Linking large converging databases, both clinical and administrative, to obtain data from 86,244 patients for CABG and 103,549 patients for PCI. Those figures are much larger than any proposed by RCTs. Interestingly, the authors found that patients undergoing CABG had better outcomes than those undergoing PCI, but at the expense of higher costs, allowing the calculation of the indicator of the incremental cost-effectiveness ratio expressed as cost per quality-adjusted life-year gained.

The role of some patients-related characteristics in determining outcomes in myocardial revascularization strategies are highlighted in the large study by Hlatky et al. [35], where authors demonstrated lower long-term
| Author                | Title                                                                 | Year | AI Model                              | Conclusions                                                                                                                                                                           |
|-----------------------|----------------------------------------------------------------------|------|---------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Diller et al. [66]    | Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: Data from a single tertiary centre including 10 019 patients. | 2019 | Deep Learning                         | Prognostication and therapeutic guidance in patients with adult congenital heart disease (ACHD) or pulmonary hypertension                                                   |
| Diller et al. [67]    | Utility of machine learning algorithms in assessing patients with a systemic right ventricle. | 2018 | Deep Learning, Convolutional neural networks | Recognizing transposition of the great arteries (TGA) after atrial switch procedure or congenitally corrected TGA (ccTGA) based on routine transthoracic echocardiograms. Delineation and segmentation of the systemic ventricle |
| Olive et al. [68]     | Current monitoring and innovative predictive modeling to improve care in the pediatric cardiac intensive care unit. | 2018 | AI and machine learning                | Predictive models created by AI and ML may lead to earlier detection of patients at risk for clinical decompensation, improving care for critically ill pediatric cardiac patients. |
| Ruiz-Fernández et al. [69] | Aid decision algorithms to estimate the risk in congenital heart surgery. | 2017 | Multilayer perceptron, self-organizing map, radial basis function networks and decision trees | Feasibility of development of CDSSs using AI algorithms. Such system would help to forecast the level of risk related to a congenital heart disease surgery. |
| Zhong et al. [70]     | Machine learning prediction models for prognosis of critically ill patients after open-heart surgery | 2021 | Extreme gradient boosting, random forest, artificial neural network, and logistic regression | Model to predict 30-days mortality and complications (i.e.: septic shock, thrombocytopenia and liver dysfunction) after open-heart surgery.                                                |
| Meyer et al. [71]     | Machine learning for real-time prediction of complications in critical care: A retrospective study | 2018 | deep learning methods (recurrent neural networks) | Predict severe complications (i.e.: mortality, renal failure with a need for renal replacement therapy, and postoperative bleeding leading to operative revision) during critical care in real time after cardiothoracic surgery. |
| Lei et al. [72]       | Using Machine Learning to Predict Acute Kidney Injury After Aortic Arch Surgery | 2020 | Machine Learning: logistic regression model, support vector machine, random forest, and gradient boosting | Machine learning methods were found to predict AKI after aortic arch surgery significantly better than traditional logistic regression.                                                      |
| Tseng et al. [73]     | Prediction of the development of acute kidney injury following cardiac surgery by machine learning | 2020 | Logistic regression, support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), and ensemble (RF + XGBoost) | AI methods predict cardiac surgery-associated acute kidney injury, which determines risks following cardiac surgery, enabling the optimization of postoperative treatment strategies. |
| Lee et al. [74]       | Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery | 2018 | ML: decision tree, random forest, extreme gradient boosting, support vector machine, neural network classification, and deep learning | Using AI an Internet-based risk estimator was developed to estimate the risk of AKI at the end of surgery.                                                                                       |
| Kilic et al. [75]     | Performance of a machine learning algorithm in predicting outcomes of aortic valve replacement. | 2020 | Extreme gradient boosting (XGBoost)    | Predicting outcomes of surgical aortic valve replacement.                                                                                                                             |
| Wojnarcki et al. [76] | Machine-learning phenotypic classification of bicuspid aortopathy     | 2018 | Random forest analysis                | Three distinct phenotypes of bicuspid valve-associated aortopathy were identified using machine-learning methodology.                                                                 |

Table 1: Artificial intelligence in cardiac surgery and cardiovascular diseases
| Author                  | Title                                                                 | Year | AI Model                                      | Conclusions                                                                                                                                 |
|------------------------|----------------------------------------------------------------------|------|-----------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| Baskaran et al. [77]   | Machine learning insight into the role of imaging and clinical variables for the prediction of obstructive coronary artery disease and revascularization: An exploratory analysis of the CONSERVE study | 2020 | ML: extreme gradient boosting (XGBoost)       | For obstructive CAD, the ML model outperformed CAD consortium clinical score (CAD2). BMI is an important variable, although currently not included in most scores. In this ML model, imaging variables were most associated with revascularization. |
| Cikes et al. [78]      | Machine learning-based pheno-grouping in heart failure to identify responders to cardiac resynchronization therapy             | 2018 | Unsupervised multiple kernel learning algorithm (MKL) | Integrating clinical parameters and full heart cycle imaging data, unsupervised ML can provide a clinically meaningful classification of a phenotypically heterogeneous HF cohort and might aid in optimizing specific therapies. |
| Ambale-Venkatesh et al. [79] | Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis | 2017 | Random survival forest (RF) alone or in combination with other statistical approaches. | Machine learning in conjunction with deep pheno-typing improves prediction accuracy in cardiovascular event prediction in an initially asymptomatic population. |
| Ayers et al. [80]      | Using machine learning to improve survival prediction after heart transplantation                                         | 2021 | Deep neural network, logistic regression, AdaBoost, and random forest | ML techniques can improve risk prediction in OHT compared to traditional approaches. This may have important implications in patient selection, programmatic evaluation, allocation policy, and patient counseling and prognostication. |
Table 2  Ongoing trials on the application of big data and derived analysis techniques to cardiac surgery

| Title                                                                 | Location                                                                 | Status            | Description                                                                                                                                                                                                 |
|----------------------------------------------------------------------|--------------------------------------------------------------------------|-------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Effects of AI Assisted Follow-up Strategy on Secondary Prevention in  | Cardiovascular Institute and Fuwai Hospital, Beijing, Beijing, China     | Not yet recruiting | Assess if AI assisted follow-up strategy will improve secondary prevention in CABG patients                                                                                                                                 |
| CABG Patients                                                        |                                                                          |                   |                                                                                                                                                                                                             |
| Artificial Intelligence Guided Patient Selection for Atrial Fibrillation Catheter Ablation: Randomized Clinical Trial (AI-PAFATrial) | Severance Hospital, Yonsei University Health System Seoul, Korea, Republic of | Not yet recruiting | Prediction of AF catheter ablation (AFCA) efficacy using artificial intelligence (AI)                                                                                                                                 |
| Effect of Artificial Intelligence on Nutritional Status of Children  | Armed Forces Institute of Cardiology and National Institute of Heart Disease Rawalpindi, Punjab, Pakistan Maryam Zahid Rawalpindi, Punjab, Pakistan | Completed         | Assess the effect of artificial intelligence on nutritional status of children post cardiac surgery in comparison to usual care group                                                                                                                                 |
| Post Cardiac Surgery                                                 |                                                                          |                   |                                                                                                                                                                                                             |
| Cloud-based ECG Monitoring and Healthcare Model Building on the Population With Coronary Artery Revascularization | ClinicalTrials.gov Identifier: NCT04485143 | Not yet recruiting | All subjects tracked the occurrence of adverse medical events within one year after discharge from the hospital. Based on the home-based remote personal care model for patients with CABG, a risk prediction model for heart failure and vascular atherosclerosis was established to effectively reduce medical treatment, adverse events, and medical expenditure |
| Machine Learning Predict Acute Kidney Injury in Patients Following  | Chinese PLA General hospital Beijing, Beijing, China                     | Completed         | Several prediction models based on machine learning technique are developed to allow early identification of patients who at the high risk of unfavorable kidney outcomes                                                                                             |
| Cardiac Surgery                                                      |                                                                          |                   |                                                                                                                                                                                                             |
| Machine Learning-Based Prediction of Major Perioperative Allogeneic Blood Requirements in Cardiac Surgery (PREMATRICS) | Kepler University Hospital Linz, Upper Austria, Austria                  | Recruiting        | If an accurate prediction model based on a few features could be created and those patients particularly at risk of massive transfusion of allogenic blood could be identified, it would subsequently be possible to develop an adapted clinical pathway that would allow patient care to be improved and individualized interventions adapted to the situation to be implemented |
| Machine Learning-Based Risk Profile Classification of Patients Undergoing Elective Heart Valve Surgery | ClinicalTrials.gov Identifier: NCT03724123 | Completed         | The investigators investigate the benefit of modern machine learning methods in personalized risk prediction in patients undergoing elective heart valve surgery                                                                                     |
| Remote Monitoring to Improve Physician Monitoring, Patient          | Mayo Clinic in Rochester Rochester, Minnesota, United States             | Completed         | Measure data collected via machine learning algorithms to predict readmission following cardiac surgery                                                                                                                                               |
| Satisfaction, and Predict Readmissions Following Surgery             |                                                                          |                   |                                                                                                                                                                                                             |

We interrogated ClinicalTrials.gov and EUDRACT databases with the following keywords: Coronary Heart Disease AND Artificial Intelligence, cardiac surgery AND Artificial Intelligence, cardiac surgery AND machine learning, cardiac surgery AND big data. A total of 811 studies were found on ClinicalTrials.gov and a total of 10 on EUDRACT; we selected the most appropriated for our purposes.

CABG coronary artery bypass grafting, AI artificial intelligence, AF atrial fibrillation
mortality with CABG rather than with PCI, in an unselected group of patients extracted by the general population of those undergoing those procedures, with outcomes substantially modified by factors such as diabetes, smoking habit, heart failure, and peripheral artery disease.

The fundamental study of Weintraub et al., by linking the STS database to that of the Centers for Medicare and Medicaid Services, showed that in stable patients, older than 65 years with multivessel coronary artery disease, CABG offers an advantage in terms of long-term survival.

While RCTs remain the only accepted evidence-based information gathered in medical guidelines, studies based on Big Data and AI added knowledge on the comparative effectiveness of the two therapeutic strategies. The place of Big Data analysis is therefore yet to be precisely determined.

Valvular heart diseases and cardiac imaging
The identification of significant valvular diseases has the potential to clarify the etiology, and/or reveal a consequence of ventricular failure, with or without dilatation, which may contribute to define the prognosis and to identify elements acting as triggers for worsening heart failure and hence prognosis. Echocardiography is used widely for assessment of cardiac structure and function, with diagnostic accuracy and reliability depending on the operators’ skill, experience and expertise. In contexts such as high volume in busy environment, and emergency, machines enriched with a technology that analyses live imaging in real-time, continuously comparing it with a pool of reference images, may help optimizing the imaging protocol and identifying diseases or patterns of abnormalities, detecting trends over time and evaluating the stability of specific measurements. Those are intriguing perspectives that may be associated with the application of AI in the field of echocardiography. For instance, in mitral valvular regurgitation, recognition of increasing severity of the valvular insufficiency, ventricular dilatation and reduction in chamber shortening, atrial dilatation, as well as the worsening myocardial function assessed by means of semiautomated, relatively load-independent parameters, may parallelly run with the detection of clinical changes and even anticipate overt changes in symptoms and signs of heart failure (HF).

In echocardiography, automated views recognition and structures identification may be considered an initial step toward semi-assisted diagnostic studies. To such an extent, AI-based technology is a key element, as convolutional neural networks may be employed to identify key reference points on images, and then feature specific for diagnostic patterns. Identification of normal patterns, deviations from normal patterns or specific pathologies has been possible in experimental studies in more than 9 cases in 10 evaluated cases [36] in algorithm-based supervised machine learning [37]. A further step in AI-assisted echocardiography may be the identification of deviation from physiology or overt pathological conditions, such as left ventricular hypertrophy, hypertrophic cardiomyopathy [38, 39], ventricular dilatation and reduced chamber and myocardial function [40]. With 2-D speckle-tracking technology, an accurate semi-automated volume and systolic function quantification may be run bed-side and take a few minutes or seconds [41, 42]. In the context of valvular disease associated with ventricular dysfunction, an important role is played by quantification of contractility reserve, which impacts prognosis. Wall motion quantification is relevant to such an extent, with AI operated diagnostic modality pushing accuracy of wall motion quantification as high as 85% [43, 44], helping in the case of evaluation of ventricular contractility reserve and aiding in decision making on the best valvular disease management.

Assessment of mitral valve regurgitation severity may be aided by automated processes based on deep learning machines [45, 46]. With regard to aortic valve disease, identification of trends over time in ascending aorta and aortic root dimensions, ventricular dimensions and shortening, relies on reproducibility of imaging in specific views, and may provide important information impacting preclinical and, timely, clinical decision making. Moreover, aortic annulus sizing represents an important target of quantification, as to date transcatheter procedures for aortic valve replacement are increasing steeply [47].

Beyond echocardiography, the potential revolution of AI may be even more applicable and profound in more standardized diagnostic processes such those applicable to nuclear medicine, computed tomography and magnetic resonance imaging, which may suffer much less variability between-subjects due to body size [48, 49].

HF and mechanical circulatory support: selection of patients, prediction of adverse events and technology development
Heart transplantation (HTx) is the gold standard therapy in advanced HF, defined as persistence of symptoms and significant personal physical limitations despite optimized pharmacological and standard nonpharmacological therapy, associated with recurrent hospitalizations and need for escalation therapy including inotropes [50]. However, the number of organs available does not match the number of patients in need for organs as they slide toward a deterioration of the clinical conditions and require a timely intervention. Left Ventricular assist devices (LVAD) may help as bridge to HTx, or to candidacy or decision, or even as destination therapy in
subjects with advanced or terminal heart failure [51, 52]. Those patients, suffering from end-stage HF, present with unique challenges: frailty [53], end-organ damage, risk for acute decompensations, and high mortality at short-term. Events as cardiogenic shock carries worsened prognosis. The therapeutic strategy shows multiple options, with multiple devices that can be employed in different phases of the clinical course [54].

Because of the high risk of surgery and the patients’ characteristics, the prediction of peri-procedural adverse events and of the long-term complications are critical issues when planning LVAD implantation, impacting benefit and quality of life per costs [55], and healthcare sustainability. Data from registries and discriminant statistics are commonly used to identify potentially life-threatening conditions impacting prognosis and hospital stay duration after LVAD implantation [56].

However, the classic way for risk-prediction estimation is based on statistical methods that implies a proportional and statistically significant participation of several variables in a context where hazard is not proportional over time [57]. On a different approach, AI and Big Data may simulate the effects of decisions, and their interactions with an uncertain environment [58] and facilitate not only the prediction of specific pre-defined events.

Prediction of complications after LVAD implantation is relevant to sustainability of LVAD procedures. AI has been involved to recognize drive-line infections using photographic database as background source [59] and the identification of clusters of variables that predict right ventricular failure, bleeding, infection and pump failure due to pump thrombosis [60].

While statistically-based risk models have proven sub-optimal ability to predict mortality risk in LVAD, by use of Interagency Registry for Mechanically Assisted Circulatory Support (INTERMACS), data from 2006 to 2016, for a total of 16,120 patients included, and bootstrapping with 1000 replications in the testing set, improved 90-day discrimination from 0.707 [0.683–0.730] to 0.740 [0.717–0.762] and 1-year mortality from 0.691 [0.673–0.710] to 0.714 [0.695–0.734] (all \( p < 0.001 \)). The net reclassification rate was up to 49% for 90-day mortality and 37% for 1-year mortality. The findings supported the concept that ML may increase the performance of a risk model for durable LVAD mortality compared to logistic regression-based algorithm [60]. Because continuous blood flow from LVAD is associated with increased risk of complications, as gastrointestinal bleeding, continuous pump speed generating flow is modulated to generate pulsatile flow. More importantly, pump speed of the LVAD may be controlled to assist left ventricles during a single beat to optimize systolic, versus diastolic assistance [61]. Diastolic versus systolic modulation of the pump speed may impact on flow pulsatility and diastolic assistance to reduce external myocardial work [62]. These mathematical models are limited in their applicability because the need for pressure feedback signals from the cardiovascular system require suitable integrated long-term pressure sensors. To date, novel AI based controllers, real-time deep convolutional neural network-based, are tested to estimate left ventricular preload using LVAD flow analyses and a sensorless adaptive control system, trained and evaluated through a number of cross validation settings and physiologic situations in different patient and different conditions, resulting in accurate pre-load evaluation (root mean squared error of 0.84 mmHg, reproducibility coefficient of 1.56 mmHg, coefficient of variation of 14.44%, and bias of 0.29 mmHg for the testing dataset) [63]. The system was able to use LVAD data to measure preload and prevent ventricular suction and pulmonary congestion [64].

Conclusions

After several years of intense research yielding a great number of scientific publications, a major gap exists between practical application of AI applied to Big Data and RCTs to guide practice in real work, in large part because AI and Big Data are yet to become controlled research tools on a large scale. Data quality control, missing data, privacy and potential conflicts of interests in a variety of stakeholders, costs of the technology required, and the need for high-performing information technology are still barriers for a routine and wide use of Big Data in the field of cardiac surgery research and clinical process.

Nevertheless, in the context of an enormous resource re-allocation due to the COVID-19 pandemic, that reduced significantly the research output in other fields of medicine, Big Data and AI may turn to be relevant tools.

The role of hypothesis generation in Big Data science is without doubt, but it should be considered as a complementary mean to obtain evidence [65].

However, the crucial match on usefulness of Big Data and AI in the near future is also played in the side of productivity, simulation, augmented reality aiding diagnostic and clinical decision making, communication with patients and generating precision medicine. These features have the potential to go well beyond the context of knowledge generated from RCTs to prove or unconfirm specific hypotheses, by using strict enrollment criteria to make homogenous the population. Hence, we all need to be familiar with those concepts and tool for the future, which is not that far away from now.

Acknowledgements

None.
References

1. Hillert M. Big Data for development: a review of promises and challenges. Dev Policy Rev. 2016;34:135–74. https://doi.org/10.1111/dpr.1214.
2. Weintraub WS. Role of Big Data in cardiovascular research. J Am Heart Assoc. 2019;8(14):e012791. https://doi.org/10.1161/JAHA.119.012791.
3. De Mauro A, Greco M, Grimaldi M. A formal definition of Big Data based on its essential features. Libr Rev. 2016;65:122 – 35. https://doi.org/10.1108/LR-06-2015-0061.
4. Silverman EK, Schmidt HH-W, Anastasioiu E, Altucci L, Angelini M, Badimon L, Balligand JL, Benincasa G, Capasso G, Conte F, Di Costanzo A, Fabris L, Fricco G, Gatto L, Gentili M, Liscalzo J, Marchese C, Napoli C, Paci P, Pietri M, Quackenbush J, Tien P, Viggiano D, Vilahur G, Glass K, Baumbach J, Badimon L, Balligand JL, Benincasa G, Capasso G, Conte F, Di Costanzo A, Fabris L, Fricco G, Gatto L, Gentili M, Liscalzo J, Marchese C, Napoli C, Paci P, Pietri M, Quackenbush J, Tien P, Viggiano D, Vilahur G, Glass K, Baumbach J. Molecular networks in Network Medicine: Development and applications. Wiley Interdiscip Rev Syst Biol Med. 2020;12(6):e1479. https://doi.org/10.1002/wsbm.1489.
5. Xiong W, Yu Z, Bei Z, Zhao J, Zhang F, Zou Y, et al. A characterization of big data benchmarks. IEEE International Conference on Big Data. 2013; pp. 118–125. https://doi.org/10.1109/BigData.2013.6691707.
6. Weintraub WS, Fahed AC, Rumsfeld JS. Translational medicine in the era of big data and machine learning. Circ Res. 2018;123(11):1202–4. https://doi.org/10.1161/CIRCRESAHA.118.313944.
7. Fizzell JD, Liang L, Schulte PI, Yancey CW, Heidenreich PA, Hernandez AF, Bhatt DL, Fonarow GC, Laskey WK. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. JAMA Cardiol. 2017;2(2):204–9. https://doi.org/10.1001/jamacardio.2016.3996.
8. Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand SL. Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. Circulation. 2007;115(12):1518–27. https://doi.org/10.1161/CIRCULATIONAHA.106.633008.
9. Harvey H, Blocker B. A standardized approach for preparing imaging data for machine learning tasks in radiology. New York: Springer; 2019. p. 61–72. https://doi.org/10.1007/978-3-319-94878-2.
10. TeliTew A, Lippert C, Chang PC, DePristo M. Deep learning of genomic variation and regulatory network data. Hum Mol Genet. 2018;27:R63–R71. https://doi.org/10.1093/hmg/ddy115.
11. Tcheng JE, Kelly J, Wosik J. Big Data: When and How Will it Impact Interventional Cardiology? ACC 2019, Expert Analysis.
12. Manuso E, Giangiacomo G, Dell Mura N, Nicoletti GP, Napoli C. Epigenetic-sensitivliquid biomarkers and personalised therapy in advanced heart failure: a focus on cell-free DNA and microRNAs. J Clin Pathol. 2020 Sep;73(9):535–43. https://doi.org/10.1136/clinpath-2019-206404.
13. Vietri MT, D’Elia G, Benincasa G, Ferraro G, Caliendo G, Nicoletti GP, Napoli C. DNA methylation and breast cancer: a way forward (Review). Int J Oncol. 2021 Nov;59(5):98. https://doi.org/10.3892/ijo.2021.5278.
14. Martin J. Big data, big future. Biotechniques. 2020 Apr;68(4):166–8. https://doi.org/10.2144/0tn-2020-0027.
15. Vietri MT, Molinari AM, Caliendo G, De Paola ML, Giovanni D, Gambardella A, Petronella P, Ciolfi M. Double heterozygosity in the BRCA1 and BRCA2 genes in Italian family. ClinChem Lab Med. 2013 Dec;51(12):2319–24. https://doi.org/10.1515/cclm-2013-0263.
16. Vietri MT, Caliendo G, Casamassimi A, Ciolfi M, De Paola ML, Napoli N, Molinari AM. A novel PALB2 truncating mutation in an Italian family with male breast cancer. Oncol Rep 2015;33(3):1243–7. https://doi.org/10.3892/or.2014.3686.
17. Sarno F, Benincasa G, List M, Barabasi AL, Baumbach J, Ciardiello F, Fletti S, Glass K, Liscalzo J, Marchese C, Maron BA, Parci P, Parini P, Petriollo E, SilvermanEK, Verreri A, Altucci L, Napoli C. International Network Medicine Consortium. Clinical epigenetics settings for cancer and cardiovascular diseases: real-life applications of network medicine at the bedside. Clin Epigenet. 2021;13:166. https://doi.org/10.1186/s13148-021-01047-2.
18. Schiano C, Benincasa G, Franzese M, Della Mura N, Pane K, Salvatore M, Napoli C. Epigenetic-sensitivpathways in personalized therapy of major cardiovascular diseases. Pharmacol Ther. 2020;210:107514. https://doi.org/10.1016/j.pharmthera.2020.107514.
19. Khera AV, Chalfin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Elliott PT, Benjamin EJ, Lubitz SA. Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. Circulation. 2018;137:1027–38. https://doi.org/10.1161/CIRCULATIONAHA.117.031431.
20. Schiano C, Costa V, Aprile M, Grimaldi V, Maiello C, Esposito R, Soricelli A, Colantuoni V, Donatelli F, Ciccodicola A, Napoli C. Heart failure: Pilot transcriptomicanalysis of cardiactissue by RNA-sequencing. Cardio J. 2017;4(5):539–53.
21. Schiano C, Franzese M, Garaci F, Zanardino M, Maiello C, Palmieri V, Soricelli A, Grimaldi V, Cosciroli E, Salvatore M, Napoli C. Machine Learning and Bioinformatics Framework Integration to Potential Familial DCM-Related Markers Discovery. Genes (Basel). 2021;12(12):1946.
22. Krittanawong C, Zhang H, Wang Z, Ayyar M, Kiti T. Artificial intelligence in precision cardiovascular medicine. J Am Coll Cardiol. 2017;69(21):2657–64. https://doi.org/10.1016/j.jacc.2017.03.571.
23. Johnson KW, Torres Soto J, Glicksberg BS, Shreiner K, Miotto R, Ali M, Ashley E, Dudley JT. Artificial intelligence in cardiology. J Am Coll Cardiol. 2018;71:2668–79. https://doi.org/10.1016/j.jacc.2018.03.521.
24. Khalsa RK, Khaskhusha A, Zaidi S, Harky A, Bashir M. Artificial intelligence and cardiac surgery during COVID-19 era. J Card Surg. 2021;36(5):1729–33. https://doi.org/10.1111/jocs.15417.
25. Khalsa RK, Khashkhusha A, Zaidi S, Harky A, Bashir M. Artificial intelligence and cardiac surgery during COVID-19 era. J Card Surg. 2021;36(5):1729–33. https://doi.org/10.1111/jocs.15417.
26. Hammond G, Carter M, Gentleman R, Ihaka R, Hornik K. The R language. J Stat Softw. 2005;15(9):1–28. https://doi.org/10.18637/jss.v015.i09.
27. Infante T, Cavallere C, Punzo B, Grimaldi V, Salvatore M, Napoli C. Radiogenomics and artificial intelligence approaches applied to cardiac
56. Soliman Oll, Akin S, Muslem R, Boersma E, Mainstved VCL, Krabatsch T, Gummert JF, de By TMMH, Bogers AJJC, Zijlstra F, Mohaci P, Caliskan K. EUROMACS Investigators. Derivation and validation of a novel right-sided heart failure model after implantation of continuous flow left ventricular assist devices: The EUROMACS (European Registry for Patients with Mechanical Circulatory Support) right-sided heart failure risk score. Circulation. 2018;137(9):891–906. https://doi.org/10.1161/CIRCULATIONAHA.117.030543.

57. Zhang Z, Reinikainen J, Adeleke RA, Peterse ME, Groothuis-Oudshoorn CGM. Time-varying covariates and coefficients in Cox regression models. Ann Transl Med. 2018;6(7):121. https://doi.org/10.21037/atam.2018.02.12.

58. Karvounis EC, Tispouros MG, Tzallas AT, Kateriadis NS, Stefanoiu K, Goletsis Y, Frigerio M, Verde A, Caruso R, Meyns B, Terrovitis J, Trivella MG, Fotiadis DI. A decision support system for the treatment of patients with ventricular assist device support. Methods Inf Med. 2014;53(2):121–36. https://doi.org/10.3414/ME13-01-0047.

59. Lüneburg N, Reiss N, Feldmann C, van der Meulen P, van de Steeg M, Schmidt T, Wendt R, Jansen S. Photographic LVAD Driveline Wound Infection Recognition Using Deep Learning. Stud Health Technol Inform. 2019;260:192–9.

60. Kilic A, Macikova J, Duan L, Movahedi F, Seese L, Zhang Y, Jacoski MV, Griffin B, Bähringer J, Blackstone EH, Svensson LG, Lyle BW. Machine-learning phenotypic classification of bicuspid aortopathy. J Thorac Cardiovasc Surg. 2018;155(2):461–9. https://doi.org/10.1016/j.jtcvs.2017.08.12.028.

61. Ogawa D, Kobayashi S, Yamazaki K, Motomura T, Nishimura T, Shimamura J, Tsukita Y, Mizuno T, Takewa Y, Tatsmusu E. Mathematical evaluation of cardiac beat synchronization control used for a rotary blood pump. Artif Organs. 2019;42(4):276–85. https://doi.org/10.1111/aor.13171.

62. Ogawa D, Kobayashi S, Yamazaki K, Motomura T, Nishimura T, Shimamura J, Tsukita Y, Mizuno T, Takewa Y, Tatsmusu E. Evaluation of cardiac beat synchronization control for a rotary blood pump on valvular regurgitation with a mathematical model. Artif Organs. 2021;45(2):124–34. https://doi.org/10.1111/aor.13795.

63. Fetanat M, Stevens M, Hayward C, Lovell NH. A SensorControl system for an implantable heart pump using a real-time deep convolutional neural network. IEEE Trans Biomed Eng. 2021;68(10):5309–28. https://doi.org/10.1109/TBME.2021.3061405.

64. Fetanat M, Stevens M, Jain P, Hayward C, Meijsing E, Lovell NH. Fully Elman neural network: a novel convective recurrent neural network optimized by an improved Harris Hawks algorithm for classification of pulmonary artery wedge pressure. IEEE Trans Biomed Eng. 2021. https://doi.org/10.1109/TBME.2021.3194559.

65. Hulsen T, Jamuar SS, Moody AR, Karnes JH, Varga O, Hedensted S, Spreafico R, Hafler DA, McKinney EF. From Big Data to precision medicine. Front Cardiovasc Med. 2019;6:34. https://doi.org/10.3389/fcvm.2019.00034.

66. Diller GP, Kempny A, Babu-Narayan SV, Henrichs M, Brida M, Uebing A, Lammers AE, Baumgartner H, Li W, Wort SJ, Dimopoulos K, Gatzoulis MA. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre involving 10 019 patients. Eur Heart J. 2019;40(13):1069–77. https://doi.org/10.1093/eurheartj/ehr623.

67. Diller GP, Kempny A, Babu-Narayan SV, Henrichs M, Brida M, Uebing A, Lammers AE, Baumgartner H, Li W, Wort SJ, Dimopoulos K, Gatzoulis MA. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease and revascularization: an exploratory analysis of the CONSERVE study. PLoS ONE. 2020;15(6):e0233791. https://doi.org/10.1371/journal.pone.0233791.

68. Kilic A, Goyal M, Miller JK, Gleason TG, Dubrawski A. A performance of a machine learning algorithm in predicting outcomes of aortic valve replacement. Ann Thorac Surg. 2021;111(2):503–10. https://doi.org/10.1016/j.athoracsur.2020.05.107.

69. Wojnarowski CM, Roselli EE, Ieddes JI, Zhu Y, Barnes TA, Lowry AM, Coller Ph, Griffin B, Bähringer J, Blackstone EH, Svensson LG, Lyle BW. Machine-learning phenotypic classification of bicuspid aortopathy. J Thorac Cardiovasc Surg. 2018;155(2):461–9. https://doi.org/10.1016/j.jtcvs.2017.08.12.028.

70. Baskaran L, Ying X, Xu Z, Al’Aref SJ, Lee BC, Lee SE, Danad I, Park HB, Bathina R, Baggianno A, Beltrama V, Cerò R, Choi EY, Choi JH, Choi SY, Cole J, Doh JH, Ha SJ, Her AY, Kepka C, Kim JY, Kim JW, Kim SW, Kim-W, Lu Y, Kumar A, Heo R, Lee JH, Jung JM, Valeti U, Andreini D, Pontone G, Han D, Villines TC, Lin F, Chang HJ, Min JK, Shaw LJ. Machine-learning insight into the role of imaging and clinical variables for the prediction of obstructive coronary artery disease and revascularization: an exploratory analysis of the CONSERVE study. PLoS ONE. 2020;15(6):e0233791. https://doi.org/10.1371/journal.pone.0233791.

71. Cikes M, Sanchez-Martinez S, Claggett B, Duranceau N, Piella G, Butakoff C, Pouleur AC, Knappe D, Biering-Sørensen T, Kutyifa V, Moss A, Stein K, Solomon SD, Bijnen B. Machine-learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. Eur J Heart Fail. 2019;21(1):74–85. https://doi.org/10.1002/ejhf.1333.

72. Amabile Venkatesh B, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Gauhar E, Bluemke DA, Lima JC. Cardiovascular event prediction by machine learning: the multi-ethnic study of atherosclerosis. Circ Res. 2017;121(9):1092–101. https://doi.org/10.1161/CIRCRESAHA.117.311312.

73. Ayers R, Sandholm T, Gosev I, Prasad S, Kilic A. Using machine learning to improve survival prediction after heart transplantation. J Card Surg. 2021;36(11):4113–20. https://doi.org/10.1111/jocs.15917.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

Learn more: biomedcentral.com/submissions