BIG DATA AND THE HEART DISEASE: IMPACT AND POTENTIAL

ADJ ASSOCIATE PROF TAN SWEE YAW

NATIONAL HEART CENTRE SINGAPORE

WHAT IS BIG DATA

- Study and application of datasets that are too large and complex for traditional data analytics to handle

APPLICATIONS

- Internet Boolean Search Engines
- Advertising
- Fintech
- Government
- Medical Research and Service Provision

ROLE OF BIG DATA AND MEDICINE

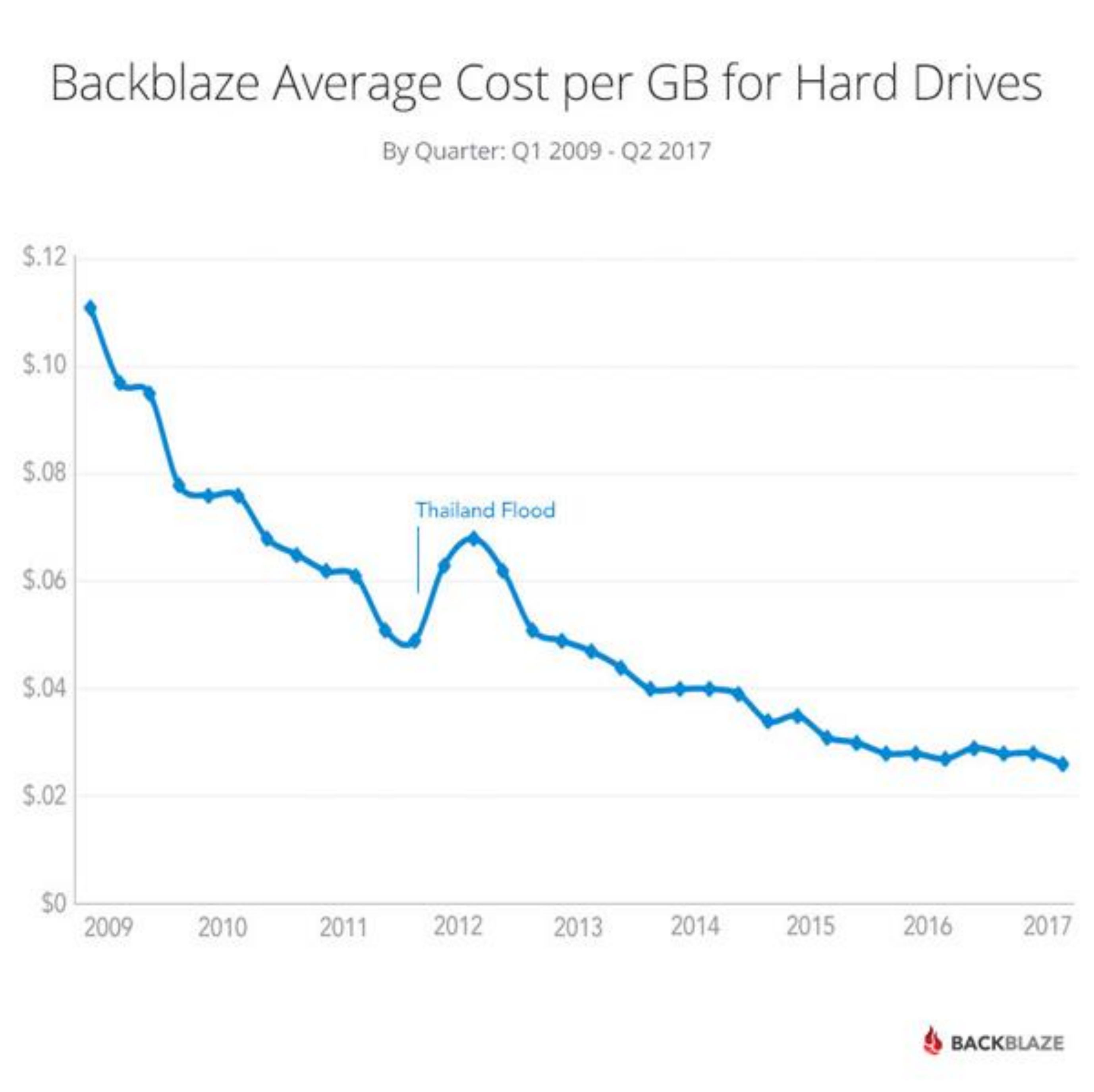
- Learning from large, complex data sets
- build better health profiles and better predictive models
- Individualise treatment
- better diagnose and treat disease.
- improve classification of disease,
- reveal ways to determine the influence of particular physicians on practice patterns,
- predict clinical events.

CHALLENGES

- Accurate data collection and acquisition
- Data Storage
- Data Analysis
- Data Search and Transfer
- Data application for real world uses
- Information Privacy and Security

WHY NOW?

CHEAP STORAGE

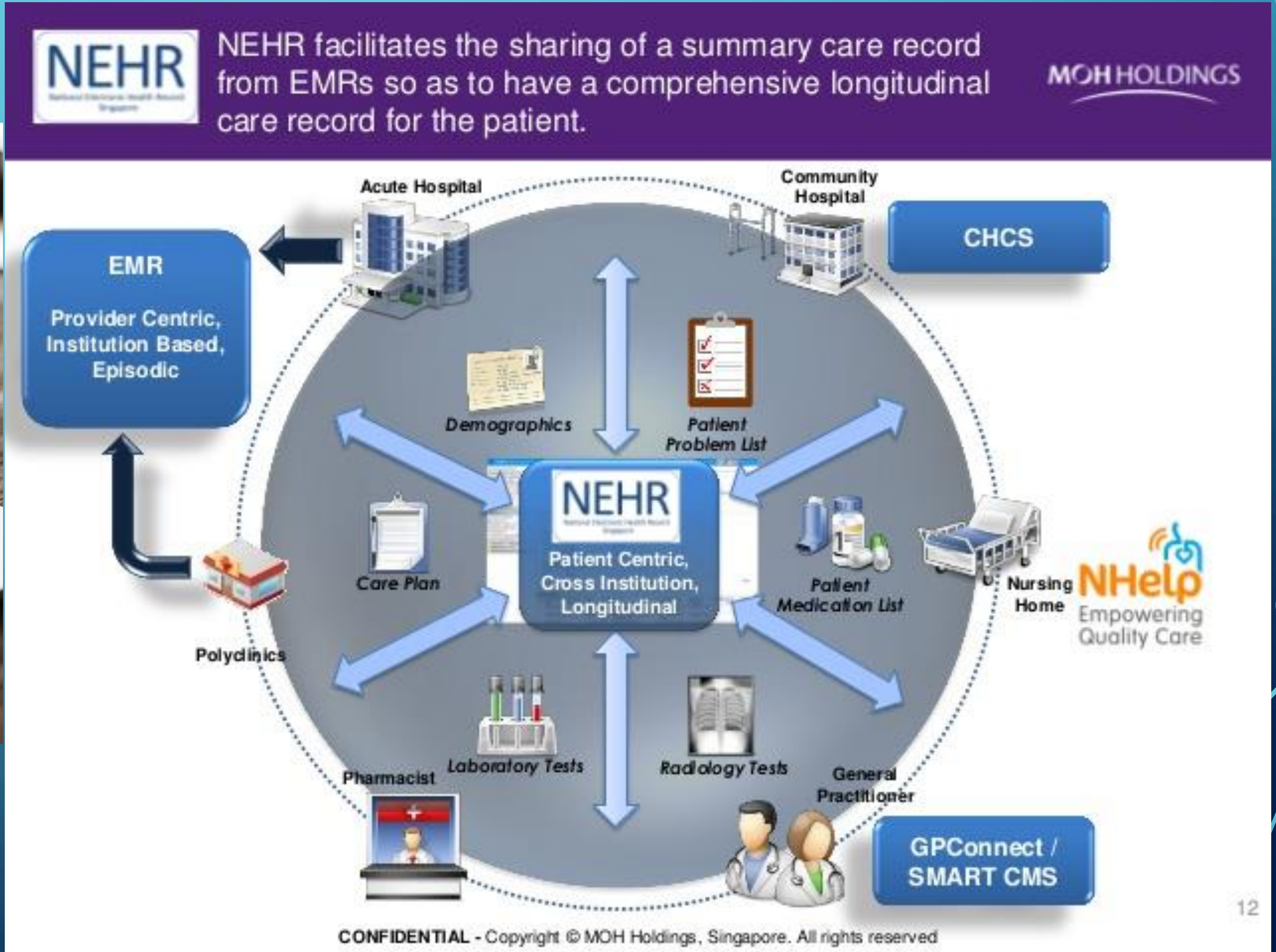


AFFORDABLE DATA GATHERING DEVICES

- Cheap data gathering (Mobile devices, Wearable trackers, Security cameras, Home Sensors) – Data acquisition has increased at double every 40 months

Kuva 1. Internet of Things. Lähde: Huffington Post

ELECTRONIC MEDICAL RECORDS



DATA SECURITY

BIG DATA AND HEALTHCARE - PERSONALISED MEDICINE

- Personalized medicine
- Prescriptive analytics
- Clinical risk intervention
- Replicating Clinical Tests
- Electronic health data management,
- Clinical Imaging storage,
- Sensor data,
- Information Security and Data Protection

PERSONALISED PRECISION CARDIOLOGY

- The Shift from 'One Size Fits All'
- The electronic patient health record (EHR) is a source of big data containing information regarding socio-demographics, medical conditions, genetics, and treatments;
- EHR will provide an integral resource for future clinical decision support systems (CDSSs)
- Adherence, screening and readmission

Patient-tailored prioritization for a pediatric care decision support system through machine learning
J. Am. Med. Inform. Assoc., 20 (e2) (2013), pp. e267-e274

Precision Analytics

PRECISION ANALYTICS

- takes prediction one step further by showing providers even variables which have never been identified
- decision support - protocols treatments useful for new clinical staff
- Prescriptive analytics leverages descriptive reports and predictive data analytics to identify the action that would produce the maximum value for the minimum effort,
- allowing users to develop and adhere to optimal clinical pathways.

TRADITIONAL RISK SCORING

FRAMINGHAM RISK SCORE to predict 10 year ABSOLUTE RISK of CHD EVENT

ST ALBANS & HEMEL HEMPSTEAD NHS TRUST : CARDIOLOGY DEPARTMENT

This risk assessment only applies to assessment for **PRIMARY PREVENTION of CHD**, in people who do not have evidence of established vascular disease.
Patients who *already* have evidence of vascular disease usually have a >20% risk of further events of over 10 years, and require vigorous **SECONDARY PREVENTION**.
People with a Family History of premature vascular disease are at higher risk than predicted; Southern Europeans and some Asians may have a lower risk in relation to standard risk factors.

STEP 1: Add scores by sex for Age, Total Cholesterol, HDL-Cholesterol, BP, Diabetes and Smoking. (If HDL unknown, assume 1.1 in Males, 1.4 in Females)

Age	M	F
30-34	- 1	- 9
35-39	0	- 4
40-44	1	0
45-49	2	3
50-54	3	6
55-59	4	7
60-64	5	8
65-69	6	8
70-74	7	8

Total Cholesterol	M	F
< 4.1	- 3	- 2
4.1 – 5.1	0	0
5.2 – 6.2	1	1
6.3 – 7.1	2	1
7.2	3	3

HDL Cholesterol	M	F
< 0.9	2	5
0.9 - 1.16	1	2
1.17 - 1.29	0	1
1.30 - 1.55	0	0
≥1.56	- 2	- 3

Systolic BP	Diastolic BP				
Male	<80	80-84	85-89	90-99	≥100
<120	0	0	1	2	3
120-129	0	0	1	2	3
130-139	1	1	1	2	3
140-159	2	2	2	2	3
≥160	3	3	3	3	3
Female	<80	80-84	85-89	90-99	≥100
<120	- 3	0	0	2	3
120-129	0	0	0	2	3
130-139	0	0	0	2	3
140-159	2	2	2	2	3
≥160	3	3	3	3	3

If Systolic and Diastolic BP fall into different categories, use score from higher category

Diabetes	M	F
No	0	0
Yes	2	4

Smoking	M	F
No	0	0
Yes	2	2

Categorisation of 10 year Risk of CHD Event	
Very Low risk	< 10%
Low risk	< 15%
Moderate risk	15-20%
High risk	> 20%

STEP 2: Use total score to determine Predicted 10 year Absolute Risk of CHD Event (Coronary Death, Myocardial Infarction, Angina) by sex

Total Score	≤-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	3	14	15	16	≥17
10 year Risk: Male		<2%	3%	3%	4%	5%	7%	8%	10%	13%	16%	20%	25%	31%	37%	45%	53%	53%	53%	53%
10 year Risk: Female	<1%	2%	2%	2%	3%	3%	4%	4%	5%	6%	7%	8%	10%	11%	13%	15%	18%	20%	24%	27%

STEP 3: Compare Predicted 10 year Absolute Risk with "Average" and "Ideal" 10 year Risks, to give Relative Risks

Age	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74
"Average" Male	3%	5%	7%	11%	14%	16%	21%	25%	30%
"Ideal" Male	2%	3%	4%	4%	6%	7%	9%	11%	14%
"Average" Female	< 1%	< 1%	2%	5%	8%	12%	12%	13%	14%
"Ideal" Female	< 1%	1%	2%	3%	5%	7%	8%	8%	8%

"Ideal" risk represents
Total Cholesterol = 4.1 - 5.1
HDL = 1.2 (Male), 1.4 (Female)
BP < 120/80
No Diabetes, Non Smoker

People with an absolute risk of ≥20% should be considered for treatment: with a Statin to achieve a Total Cholesterol <5 and/or LDL cholesterol <3.2
with anti-hypertensives to achieve a BP ≤160/90 (ideally ≤140/80)

from Wilson PWF, et al Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837-47

Dr John Bayliss

Am J Cardiol. 2012 Apr 1;109(7):998-1004. doi: 10.1016/j.amjcard.2011.11.028. Epub 2012 Jan 9.

Comparison of the Diamond-Forrester method and Duke Clinical Score to predict obstructive coronary artery disease by computed tomographic angiography.

Wasfy MM[1], Brady TJ, Abbara S, Nasir K, Ghoshhajra BB, Truong QA, Hoffmann U, Di Carli MF, Blankstein R.

Author information

Abstract

We sought to evaluate the ability of the Diamond and Forrester method (DFM) and the Duke Clinical Score (DCS) to predict obstructive coronary artery disease (CAD) on coronary computed tomographic angiography (CCTA) and the effect of these different risk scores on the appropriateness level using the 2010 Appropriate Use Criteria. Consecutive symptomatic patients who underwent CCTA for evaluation of CAD (n = 114) were classified as having a low, intermediate, or high pretest probability using the DFM and DCS. Using the Appropriate Use Criteria, the indications for CCTA were classified according to the pretest probability and previous testing. The CCTA results were classified as revealing obstructive (≥70% stenosis), nonobstructive (<70%), or no CAD. When the patients' risk was classified using the DFM, 18% were low, 65% intermediate, and 17% high risk. When using the DCS, 53% of patients had a reclassification of their risk, most of whom changed from intermediate to either low or high risk (50% low, 19% intermediate, 35% high risk). The net reclassification improvement for the prediction of obstructive CAD was 51% (p = 0.01). Of the 37 patients who were reclassified as low risk, 36 (97%) lacked obstructive CAD. Appropriateness for CCTA was reclassified for 13% of patients when using the DCS instead of the DFM, and the number of appropriate examinations was significantly fewer (68% vs 55%, p <0.001). In conclusion, reclassification of risk using the DCS instead of the DFM resulted in improved prediction of obstructive CAD on CCTA, especially in low-risk patients. More patients were categorized as having a high pretest probability of CAD, resulting in reclassification of their examination indications as uncertain or inappropriate. These results identify the need for improved pretest risk scores for noninvasive tests such as CCTA and suggest that the method of risk assessment could have important implications for patient selection and quality assurance programs.

PREDICTIVE ANALYTICS

- Predictive analytics tell users what is likely to happen by using historical patterns to infer how future events are likely to unfold.
- eg. If a patient fails to fill a prescription for his congestive heart failure, he is likely to have a higher risk of ending up in the emergency department.
- Predictive analytics can identify that patient based on his previous patterns – five late refills in the past year on his pharmacy record, three pounds of weight gain as evidenced by his Bluetooth-connected home scale, and a higher number of calls to his primary care provider in the past two weeks – and calculate the risk that the individual is headed for a costly crisis.

RISK MODEL PREDICTION

- Current ASCVD risk prediction models
- Expert selection of variables, fine-tuning of variable transformations and interactions, and imputing missing values are time-consuming and could bias subsequent analysis, particularly given that missingness in EHR is both high, and may carry meaning.
- Prognostic modelling is important in clinical practice and epidemiology for patient manage- ment and research.

Pooled Cohort Risk Assessment Equations

Predicts 10-year risk for a first atherosclerotic cardiovascular disease (ASCVD) event

ClinCalc.com » Cardiology » Pooled Cohort 10-Year ASCVD Risk Assessment Equations

Risk Factors for ASCVD

Gender	Male / Female	Systolic BP	mmHg
Age	years	Receiving treatment for high blood pressure (if SBP > 120 mmHg)	No / Yes
Race	White or other	Diabetes	No / Yes
Total Cholesterol	mg/dL	Smoker	No / Yes
HDL Cholesterol	mg/dL		

Reset Calculate

US units

Electronic Health Record Management and Data Mining

MACHINE LEARNING AND EHR

- machine learning approaches combined with EHR may make it feasible to produce fine-tuned, individualised prognostic models, which will be particularly valuable in patients with conditions or combinations of conditions which would be very difficult for conventional modelling approaches to capture.

IDENTIFYING NEW RISK VARIABLES

- EHR studies used a median of just 27 variables
- Existing using EHR not many variables
- Issues with missing data

Cardiovascular Risk Factors

JNC 7

Major Risk Factors	Target-Organ Damage
Hypertension*	Heart
Age (>55 years, ♂, >65 years, ♀)	LVH
Diabetes mellitus*	Angina/prior MI
↑LDL/total cholesterol, or ↓HDL*	Prior coronary revascularization
Estimated GFR <60 mL/min	Heart failure
Family history of premature CVD (<55 years, ♂, <65 years, ♀)	Brain
Microalbuminuria (UAE=30-300 mg/d)	Stroke or transient ischemic attack
Obesity* (BMI ≥30 kg/m^2)	CKD
Physical inactivity	Peripheral arterial disease
Tobacco usage	Retinopathy

PLOS | ONE

RESEARCH ARTICLE

Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease

Andrew J. Steele[1]*, Spiros C. Denaxas[2], Anoop D. Shah[2,3], Harry Hemingway[2], Nicholas M. Luscombe[1,4,5]

1 The Francis Crick Institute, London, United Kingdom, **2** Farr Institute of Health Informatics Research, Institute of Health Informatics, University College London, London, United Kingdom, **3** University College London Hospitals NHS Foundation Trust, London, United Kingdom, **4** UCL Genetics Institute, Department of Genetics Evolution and Environment, University College London, London, United Kingdom, **5** Okinawa Institute of Science & Technology Graduate University, Okinawa, Japan

* statto+plosone@andrewsteele.co.uk

Check for updates

OPEN ACCESS

Citation: Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM (2018) Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. PLoS ONE 13(8): e0202344. https://doi.org/10.1371/journal.pone.0202344

Editor: Tiratha Raj Singh, Jaypee University of Information Technology, INDIA

Received: February 5, 2018

Accepted: July 30, 2018

Published: August 31, 2018

Data Availability Statement: While our data does not contain any personal sensitive identifiers, it's deemed as sensitive as it contains sufficient clinical information about patients such as dates of clinical events for there to be a potential risk of patient re-identification. This restriction has been imposed by the data owner (CPRD/MHRA) and the data sharing agreements between UCL and the CPRD/MHRA. Access to data may be requested via the Clinical Practice Research Datalink (CPRD) and applying to the CPRD's Independent Scientific

Abstract

Prognostic modelling is important in clinical practice and epidemiology for patient management and research. Electronic health records (EHR) provide large quantities of data for such models, but conventional epidemiological approaches require significant researcher time to implement. Expert selection of variables, fine-tuning of variable transformations and interactions, and imputing missing values are time-consuming and could bias subsequent analysis, particularly given that missingness in EHR is both high, and may carry meaning. Using a cohort of 80,000 patients from the CALIBER programme, we compared traditional modelling and machine-learning approaches in EHR. First, we used Cox models and random survival forests with and without imputation on 27 expert-selected, preprocessed variables to predict all-cause mortality. We then used Cox models, random forests and elastic net regression on an extended dataset with 586 variables to build prognostic models and identify novel prognostic factors without prior expert input. We observed that data-driven models used on an extended dataset can outperform conventional models for prognosis, without data preprocessing or imputing missing values. An elastic net Cox regression based with 586 unimputed variables with continuous values discretised achieved a C-index of 0.801 (bootstrapped 95% CI 0.799 to 0.802), compared to 0.793 (0.791 to 0.794) for a traditional Cox model comprising 27 expert-selected variables with imputation for missing values. We also found that data-driven models allow identification of novel prognostic variables; that the absence of values for particular variables carries meaning, and can have significant implications for prognosis; and that variables often have a nonlinear association with mortality, which discretised Cox models and random forests can elucidate. This demonstrates that machine-learning approaches applied to raw EHR data can be used to build models for use in research and clinical practice, and identify novel predictive variables and their effects to inform future research.

- AI can sometimes beat doctors use of records in predicting CAD
- Machine learning identify hero unknown variables

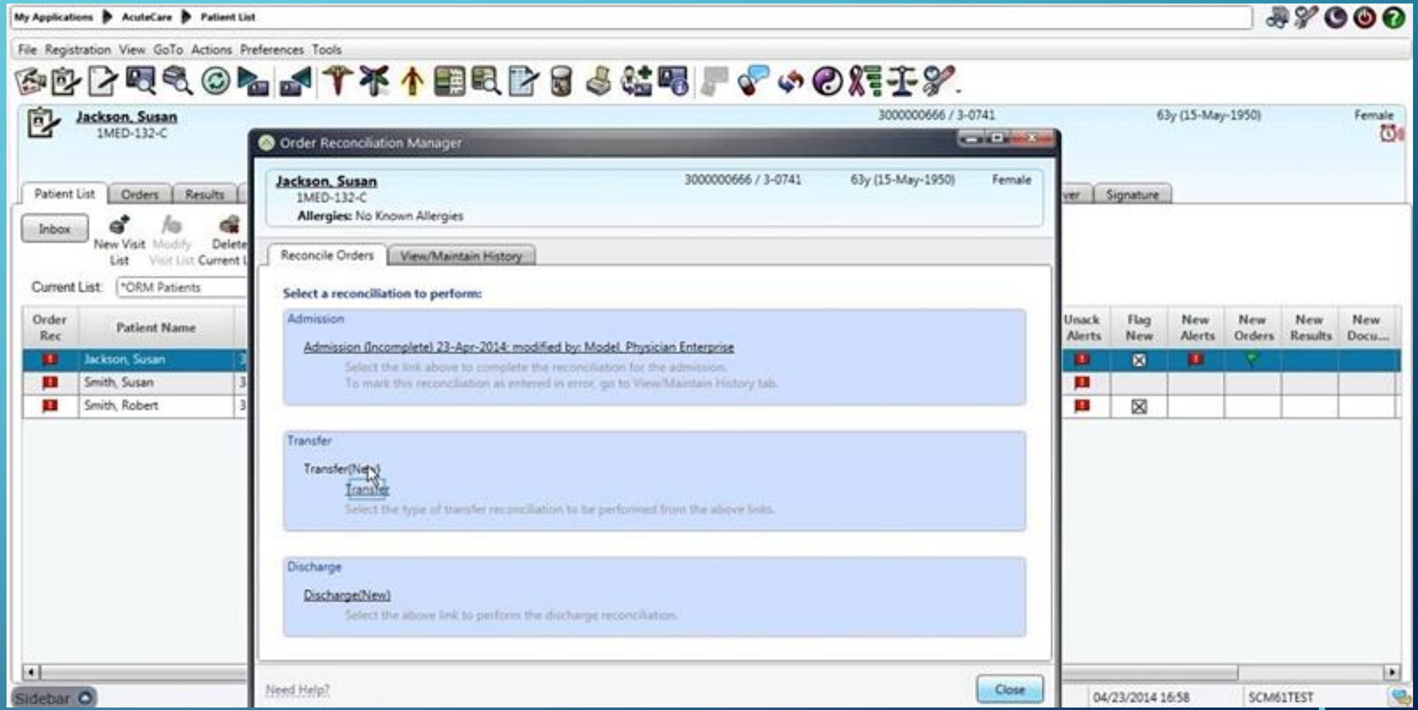
Clinical Risk Intervention

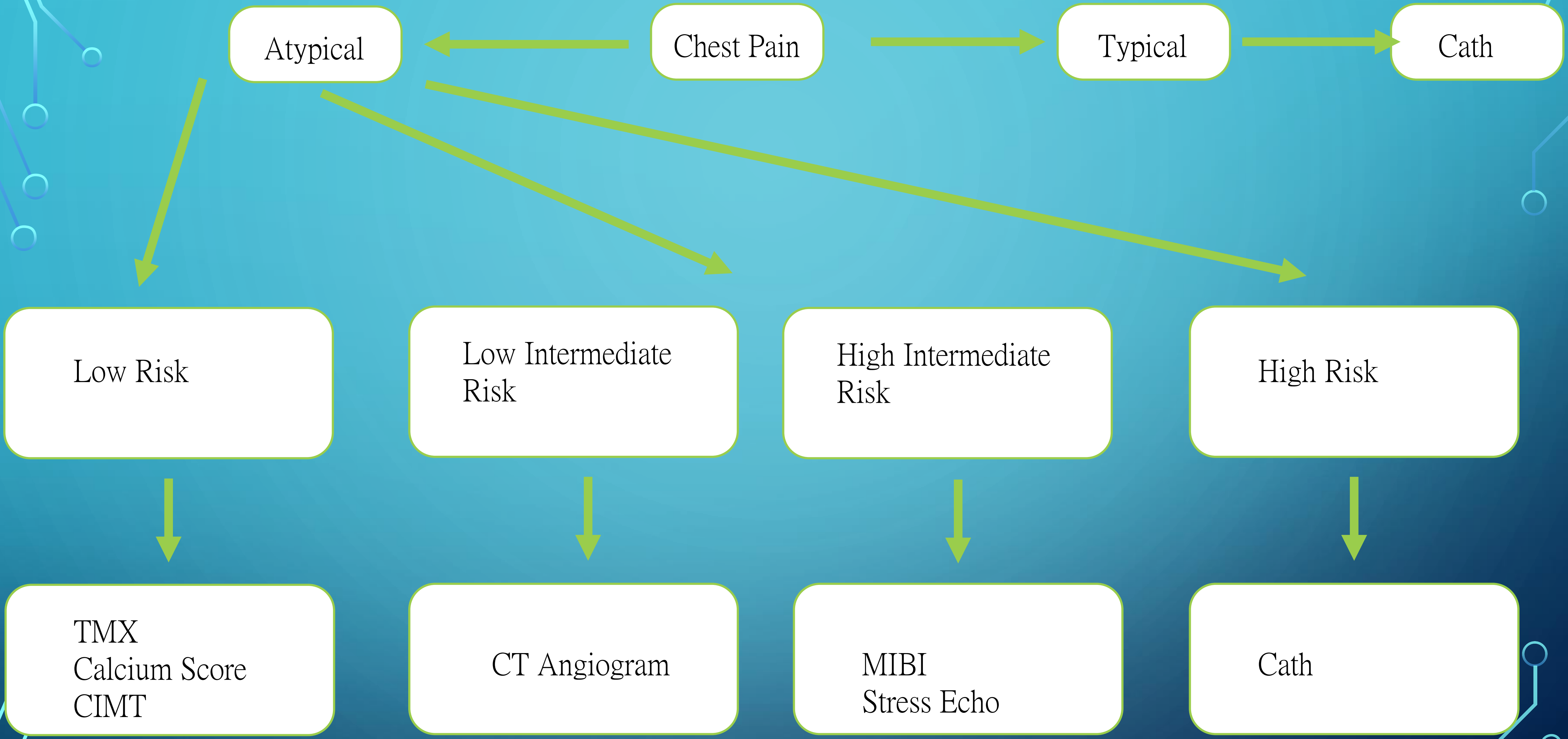
CDSS CLINICAL DECISION SUPPORT SOFTWARE

- Health Records have been successfully mined for post-market surveillance of medications and improved pharmacovigilance.
- CDS can be closely integrated with CPOE (Computer Provider Order Entry)
- CPOE for pharmacy orders has been around for several years in an effort to leverage technology to reduce medication error

CPOE

- Helps in CDSS
- Protocolsize certain prescriptions
- Tracks prescriber data
- Tracks dispensation of medications , eg warfarin, antibiotics, opiod and sedative use.




Chest Pain
Atypical
Typical
Cath
Low Risk
Low Intermediate Risk
High Intermediate Risk
High Risk
TMX
Calcium Score
CIMT
CT Angiogram
MIBI
Stress Echo
Cath

PHARMACO-VIGILANCE

- Using Big data to help programme software to detect and help resolve prescription errors
- Detect where most common errors occur
- Help formulate Decision support software to prevent that from occuring

COUNTERFEIT DETECTION

- ‘track and trace’ legislation to combat illegal drugs in the pharmaceutical supply chain.
- Artificial Intelligence (AI) will allow companies to use their ‘track and trace’ programs to deliver new levels of transparency and visibility to their supply chains.

WAFARIN DOSING

- Improves safety
- Training
- CDSS

Research Article

Revisiting Warfarin Dosing Using Machine Learning Techniques

Ashkan Sharabiani,[1] Adam Bress,[2] Elnaz Douzali,[1] and Houshang Darabi[3]

[1] *Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Room 4209, SEL West Building, 950 South Halsted Street, Chicago, IL 60607, USA*
[2] *Department of Pharmacotherapy, University of Utah, 30 South 2000 East, Room 4929, Salt Lake City, UT 84112, USA*
[3] *Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Room 2055, ERF Building, 842 W Taylor Street, Chicago, IL 60607, USA*

Correspondence should be addressed to Houshang Darabi; hdarabi@uic.edu

Received 13 February 2015; Revised 11 May 2015; Accepted 21 May 2015

Academic Editor: Chuangyin Dang

Determining the appropriate dosage of warfarin is an important yet challenging task. Several prediction models have been proposed to estimate a therapeutic dose for patients. The models are either clinical models which contain clinical and demographic variables or pharmacogenetic models which additionally contain the genetic variables. In this paper, a new methodology for warfarin dosing is proposed. The patients are initially classified into two classes. The first class contains patients who require doses of >30 mg/wk and the second class contains patients who require doses of ≤30 mg/wk. This phase is performed using relevance vector machines. In the second phase, the optimal dose for each patient is predicted by two clinical regression models that are customized for each class of patients. The prediction accuracy of the model was 11.6 in terms of root mean squared error (RMSE) and 8.4 in terms of mean absolute error (MAE). This was 15% and 5% lower than IWPC and Gage models (which are the most widely used models in practice), respectively, in terms of RMSE. In addition, the proposed model was compared with fixed-dose approach of 35 mg/wk, and the model proposed by Sharabiani et al. and its outperformance were proved in terms of both MAE and RMSE.

Disesae Mapping and Resouce Allocation

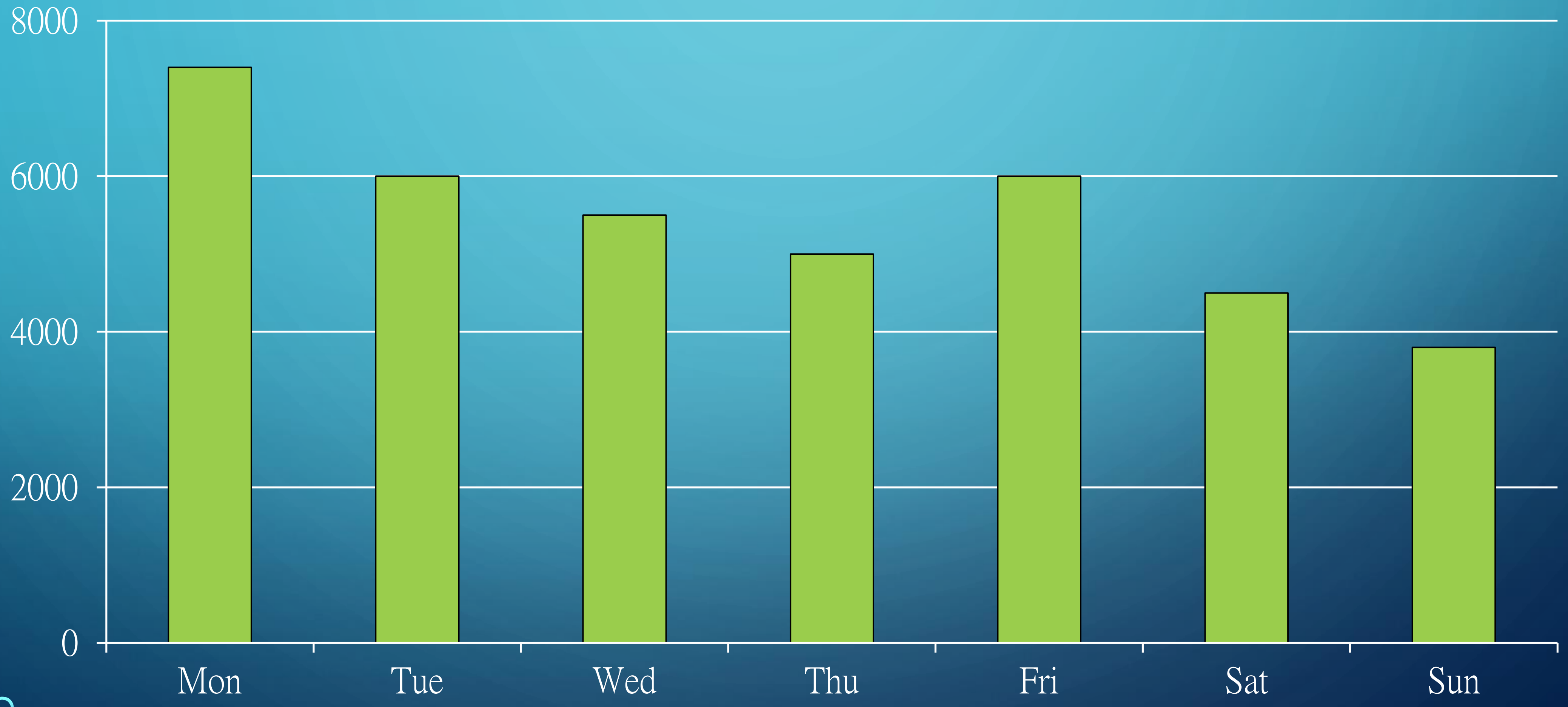
DISEASE MAPPING

- In the wake of serious shortages of doctors and nurses, international public health organizations have been forced to make tough decisions about which groups are in the most dire need of treatment.
- Ebola crisis inspired the use of various platforms, apps, and services to collect data and communication and provide real-time information about recent outbreak situations and developments.
- a mobile phone based data-collection system that was able to collect real-time data about teacher and student attendance to help ensure hygiene equipment had been delivered to schools in Sierra Leone.

SEE TRENDS FOR DEPLOYMENT OF STAFF

- ability to predict and track major outbreaks in order to improve public healthcare resources and the dissemination of healthcare messages to victims using social media.
- Predictions of serious healthcare emergencies such as exacerbations of asthma can be better predicted in models that combine social media analysis with environmental data.
- prediction models of the volume of daily emergency-department visits for acute asthma (volume defined as low, moderate, or high) using Twitter activity, Google searches, and air-quality data

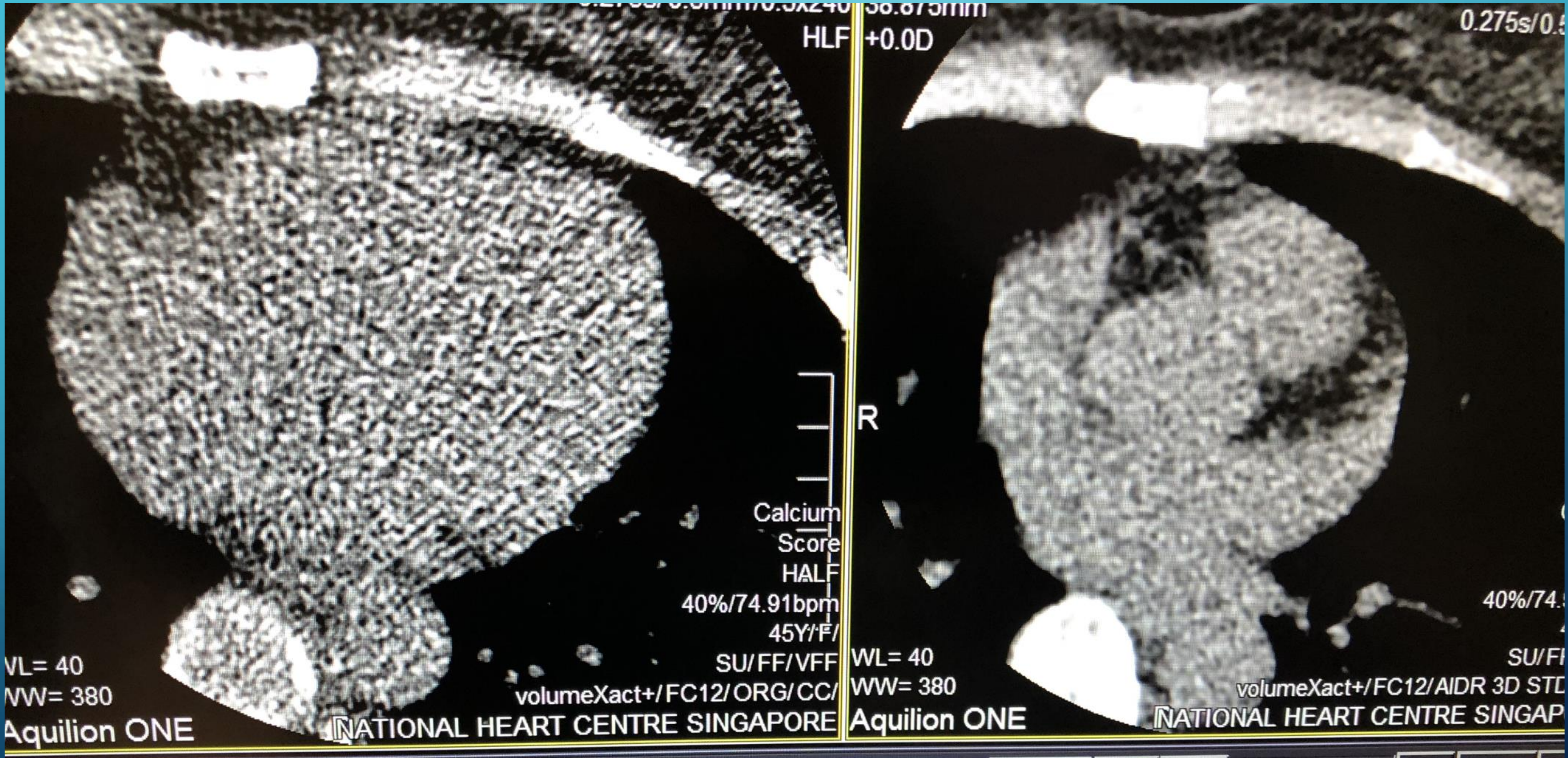
ACCIDENT AND EMERGENCY ADMISSIONS



BIG DATA AND CARDIAC IMAGING

- AI to improve image quality
- AIDR (Adaptive Iterative Dose Reduction)
- Interpolation of images - much like how TVs convert DVD images to 4K

BIG DATA AND CARDIAC IMAGING



Original

AI Enhanced

BIG DATA COMPARED TO ESTABLISHED ALGORITHMS

RESEARCH ARTICLE

Can machine-learning improve cardiovascular risk prediction using routine clinical data?

Stephen F. Weng[1,2☯]*, **Jenna Reps**[3,4☯], **Joe Kai**[1,2‡], **Jonathan M. Garibaldi**[3,4‡], **Nadeem Qureshi**[1,2‡]

1 NIHR School for Primary Care Research, University of Nottingham, Nottingham, United Kingdom, **2** Division of Primary Care, School of Medicine, University of Nottingham, Nottingham, United Kingdom, **3** Advanced Data Analysis Centre, University of Nottingham, Nottingham, United Kingdom, **4** School of Computer Science, University of Nottingham, Nottingham, United Kingdom

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* stephen.weng@nottingham.ac.uk

Citation: Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE 12(4): e0174944. https://doi.org/10.1371/journal.pone.0174944

Editor: Bin Liu, Harbin Institute of Technology Shenzhen Graduate School, CHINA

Received: December 14, 2016

Accepted: March 18, 2017

Published: April 4, 2017

Data Availability Statement: This dataset contains patient level health records with intellectual property rights held by The Crown copyright, which is subject to UK information governance laws. The authors will make their data available upon specific requests subject to the requestor obtaining ethical and research approvals from the Clinical Practice Research Datalink Independent Scientific Advisory Committee (https://www.cprd.com/intro.asp) at the UK Medicines and Health Products Regulatory Agency.

Abstract

Background

Current approaches to predict cardiovascular risk fail to identify many people who would benefit from preventive treatment, while others receive unnecessary intervention. Machine-learning offers opportunity to improve accuracy by exploiting complex interactions between risk factors. We assessed whether machine-learning can improve cardiovascular risk prediction.

Methods

Prospective cohort study using routine clinical data of 378,256 patients from UK family practices, free from cardiovascular disease at outset. Four machine-learning algorithms (random forest, logistic regression, gradient boosting machines, neural networks) were compared to an established algorithm (American College of Cardiology guidelines) to predict first cardiovascular event over 10-years. Predictive accuracy was assessed by area under the 'receiver operating curve' (AUC); and sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) to predict 7.5% cardiovascular risk (threshold for initiating statins).

Findings

24,970 incident cardiovascular events (6.6%) occurred. Compared to the established risk prediction algorithm (AUC 0.728, 95% CI 0.723–0.735), machine-learning algorithms improved prediction: random forest +1.7% (AUC 0.745, 95% CI 0.739–0.750), logistic regression +3.2% (AUC 0.760, 95% CI 0.755–0.766), gradient boosting +3.3% (AUC 0.761, 95% CI 0.755–0.766), neural networks +3.6% (AUC 0.764, 95% CI 0.759–0.769). The highest achieving (neural networks) algorithm predicted 4,998/7,404 cases (sensitivity 67.5%, PPV 18.4%) and 53,458/75,585 non-cases (specificity 70.7%, NPV 95.7%), correctly predicting 355 (+7.6%) more patients who developed cardiovascular disease compared to the established algorithm.

PLOS ONE

RESEARCH ARTICLE

Can machine-learning improve cardiovascular risk prediction using routine clinical data?

Stephen F. Weng[1,2☯*], **Jenna Reps**[3,4☯], **Joe Kai**[1,2‡], **Jonathan M. Garibaldi**[3,4‡], **Nadeem Qureshi**[1,2‡]

1 NIHR School for Primary Care Research, University of Nottingham, Nottingham, United Kingdom, **2** Division of Primary Care, School of Medicine, University of Nottingham, Nottingham, United Kingdom, **3** Advanced Data Analysis Centre, University of Nottingham, Nottingham, United Kingdom, **4** School of Computer Science, University of Nottingham, Nottingham, United Kingdom

☯ These authors contributed equally to this work.
‡ These authors also contributed equally to this work.
* stephen.weng@nottingham.ac.uk

Citation: Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE 12(4): e0174944. https://doi.org/10.1371/journal.pone.0174944

Editor: Bin Liu, Harbin Institute of Technology Shenzhen Graduate School, CHINA

Received: December 14, 2016

Accepted: March 18, 2017

Published: April 4, 2017

Data Availability Statement: This dataset contains patient level health records with intellectual property rights held by The Crown copyright, which is subject to UK information governance laws. The authors will make their data available upon specific requests subject to the requestor obtaining ethical and research approvals from the Clinical Practice Research Datalink Independent Scientific Advisory Committee (https://www.cprd.com/intro.asp) at the UK Medicines and Health Products Regulatory Agency.

Abstract

Background

Current approaches to predict cardiovascular risk fail to identify many people who would benefit from preventive treatment, while others receive unnecessary intervention. Machine-learning offers opportunity to improve accuracy by exploiting complex interactions between risk factors. We assessed whether machine-learning can improve cardiovascular risk prediction.

Methods

Prospective cohort study using routine clinical data of 378,256 patients from UK family practices, free from cardiovascular disease at outset. Four machine-learning algorithms (random forest, logistic regression, gradient boosting machines, neural networks) were compared to an established algorithm (American College of Cardiology guidelines) to predict first cardiovascular event over 10-years. Predictive accuracy was assessed by area under the 'receiver operating curve' (AUC); and sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) to predict 7.5% cardiovascular risk (threshold for initiating statins).

Findings

24,970 incident cardiovascular events (6.6%) occurred. Compared to the established risk prediction algorithm (AUC 0.728, 95% CI 0.723–0.735), machine-learning algorithms improved prediction: random forest +1.7% (AUC 0.745, 95% CI 0.739–0.750), logistic regression +3.2% (AUC 0.760, 95% CI 0.755–0.766), gradient boosting +3.3% (AUC 0.761, 95% CI 0.755–0.766), neural networks +3.6% (AUC 0.764, 95% CI 0.759–0.769). The highest achieving (neural networks) algorithm predicted 4,998/7,404 cases (sensitivity 67.5%, PPV 18.4%) and 53,458/75,585 non-cases (specificity 70.7%, NPV 95.7%), correctly predicting 355 (+7.6%) more patients who developed cardiovascular disease compared to the established algorithm.

Wearbles and Sensor Data

WEARABLES AND SENSOR DATA

- smartwatch to monitor your heart rate , heart rate recovery , ECG
- Blood pressure
- Blood glucose
- SPO2
- Steps and Calorie burn

WEARABLES

- Through the use of these wearables, patients may be alerted to take their medication or moderate their activity levels, due to an irregular heartbeat or a change in blood-oxygen levels.
- This remotely-collected data will automatically be integrated with other streams of data and EHRs to help alert doctors and family members, in the event of an emergency
- Physician Visits just 10 mins, Wearables give longitude data

WEARABLES

- baseline may predict a disease state or sliding into a disease state.
- able to intervene sooner to prevent you from that kind of slide.
- And while the wearable devices today are in this more recreational-grade state, they' re changing incredibly rapidly into research grade and ultimately clinical grade.

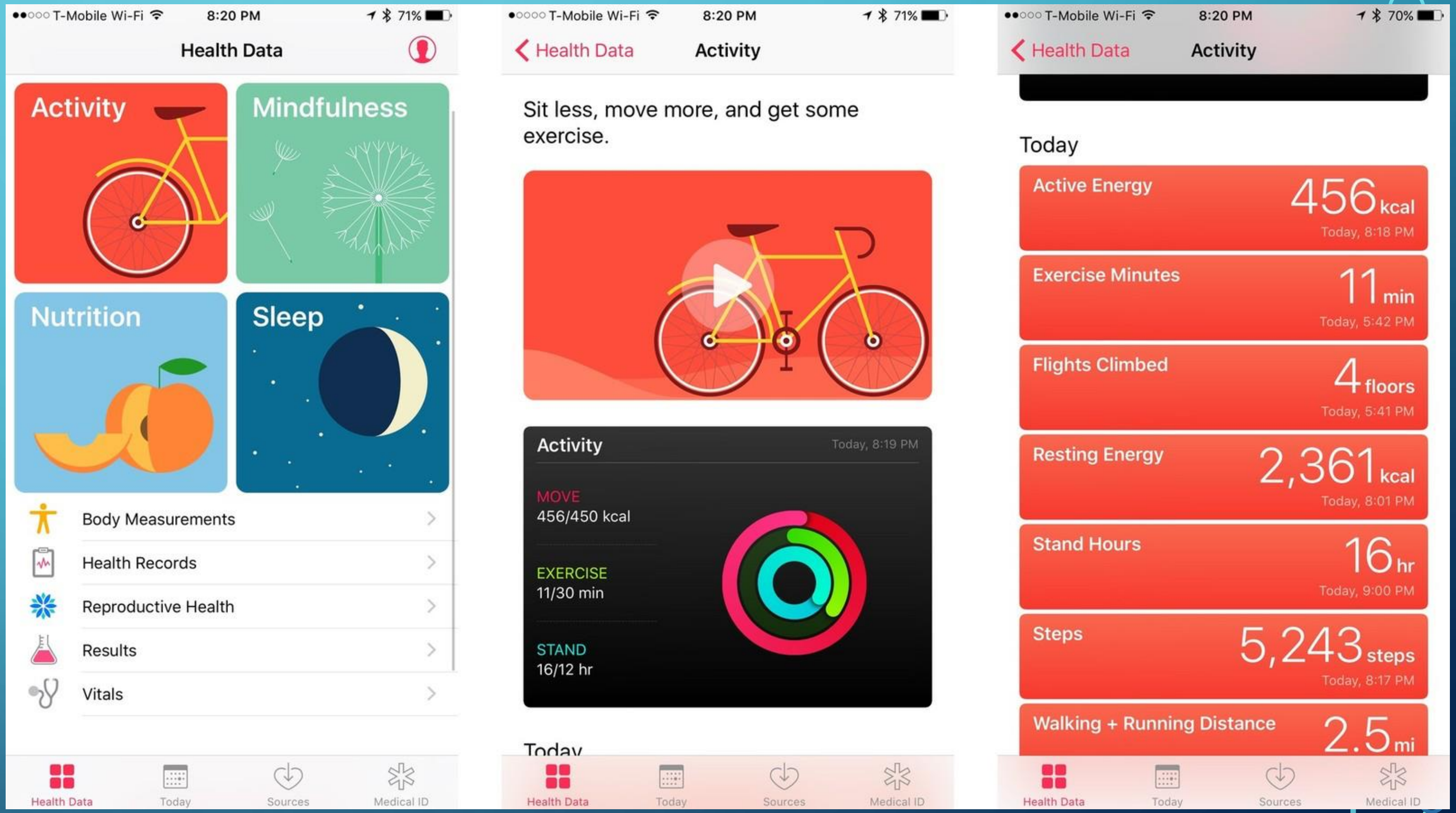
Apple Watch's heart rate sensors alert man to undiagnosed atrial fibrillation

By Roger Fingas
Monday, May 28, 2018, 05:59 am PT (08:59 am ET)

An Apple Watch owner in England was recently rescued by a watchOS alert that his heart rate had suddenly spiked, even though he otherwise felt fine.

WEARABLES

- data collection will be passive,
- looking at dashboards about themselves
- they' re able to see it every day and understand what it means.



Data Security, Integrity and Protection

DATA SECURITY CHALLENGES

- increase in patient numbers,
- harder for hospitals and clinics to process and store information.
- Secure information sharing methods, are crucial
- This is where blockchain comes in useful, as one of its main advantages is data integrity. When information is recorded and encrypted, it becomes impossible to change or remove.

BLOCK CHAIN

- technology that creates immutable and distributable data records which are shared peer to peer between networked database systems.
- records digital events in a way that does not allow for the data to be changed or recognized until it reaches the recipient.
- data is theoretically secured and protected from data breach threats.

BLOCK CHAIN

- secure recording and sharing of information is anchoring data to the public blockchain.
- Secure verification of medical data and audits
- Improve the integrity of clinical research resultsand ensure regulatory compliance

Block Chain Serial Graphic cards Array

WHY BLOCK CHAIN IS SECURE

- decentralized register, recording every transaction made through the system.
- Every device that is a part of the system stores a copy of this block.
- Before making a transaction, the system confirms whether a blockchain version coincides with another in the network.
- Therefore, each blockchain user can identify the owner of a particular data block at any time.
- blockchain is not only a secure way to send money, but a fully protected data sharing method that widens its potential use in healthcare.

Is Clinical Medicine dead?

NO!

DON’ T FOLLOW COMPUTERS BLIND

- AI and Big Data are tools tat help physicians identify and decide how to treat patients
- It cannot supplant the reassurance that a doctor gives the patient.

A Calcium Score of Zero has a High Negative Predictive Value for Excluding Severe Coronary Artery Stenosis in Symptomatic Patients in an Asian Population

TAY YL[1], CHUA SJ[2] , ALLEN JC[1], THANJU A.[1], TAN SY[2]
[1]Duke-NUS Graduate Medical School [2]National Heart Centre, Singapore

Author Disclosures:
Chua SJ – National PI for ALTITUDE Study & TRILOGY Trial
All Others – No Disclosures

Introduction

Coronary Artery Calcium (CAC) scoring may have potential as a gatekeeper to further testing with Coronary Computed Tomography Angiography (CCTA) or other tests in patients presenting with chest pain and suspected Coronary Artery Disease (CAD). Apart from a substudy of the CONFIRM Registry, other studies evaluating CAC for this role had limited sample sizes with conflicting results. Moreover, none of these studies were performed in an Asian population. The aim of our study was to assess the Negative Predictive Value (NPV) of CAC scoring for CAD as defined by CCTA in a large symptomatic Asian population.

Coronary Artery Calcification

Coronary artery calcification has been extensively studied and is one of the strongest predictors of future coronary events and mortality. The advantages of CAC are that it is relatively low cost and efficient, involves low radiation doses, requires a lower level of expertise to interpret and does not require the administration of intravenous contrast.

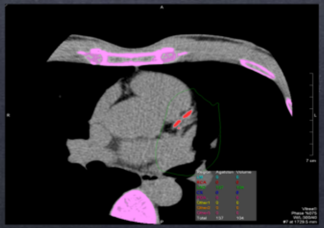


Figure 1. A patient with coronary calcification in the left anterior descending artery. Such patients have been shown to be at risk from coronary events.

Methods

This was a single-center, observational study of all patients referred to our institution for CCTA from March 2007 to September 2012. All patients underwent CAC prior to CCTA on either a 64-Slice or 320-Slice CT using a standard protocol. CAC scores were interpreted using a dedicated work station via the Agatston Schema. Patients with no symptoms, prior infarct, known significant CAD, previous revascularization or uninterpretable scans were excluded from the study. Pre-test risk for severe CAD was calculated for patients presenting with chest pain using the Duke Clinical Score. 95% Confidence Intervals (CI) were calculated using the Clopper-Peason Exact method.

Results

Demographics	
Mean Age	55.7 ± 10.9
Mean BMI	25.6 ± 4.3
Male	715 (58.3%)
Female	512 (41.7%)
Symptoms	**Frequency**
Chest Pain	968 (78.9%)
Dyspnea	173 (14.1%)
Others	86 (7.0%)
Risk Factors	**Frequency**
Current Smoker	126 (10.7%)
Hypertension	617 (50.3%)
Dyslipidemia	853 (69.5%)
Diabetes Mellitus	198 (16.1%)
FHx of premature CAD	236 (19.2%)

Degree of Stenosis	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
≥70%	96.4 [91.0-99.0]	46.8 [43.9-49.8]	**99.2** [98.1-99.8]	15.1 [12.6-18.0]
≥50%	96.1 [93.5-97.7]	57.8 [54.4-61.0]	**97.5** [95.8-98.7]	46.3 [42.5-50.1]

Risk Category	*Chest pain patients, ≥70% stenosis*			
	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
High (n=96)	92.9 [76.5-99.1]	15.5 [8.0-26.0]	**84.6** [54.6-98.1]	30.2 [20.8-41.1]
Intermed (n=415)	100.0 [90.5-100.0]	38.1 [33.2-43.1]	**100.0** [97.5-100.0]	13.3 [9.5-17.9]
Low (n=461)	94.7 [74.0-99.9]	67.1 [62.4-71.5]	**99.7** [98.1-100.0]	11.3 [6.8-17.3]

Of 1227 symptomatic patients who underwent CAC scoring and CCTA, 527 patients had a CAC score of zero. Four of 527 patients (0.8%) had severe stenosis (≥70% stenosis) while 13 of 527 patients (2.5%) had moderate to severe stenosis (≥50% stenosis) on CCTA. The NPV of CAC score of zero for excluding ≥70% stenosis was 99.2% and for excluding ≥50% stenosis was 97.5%. The NPV of CAC score of zero for excluding ≥70% stenosis in chest pain patients with high, intermediate and low risk for severe CAD were 84.6%, 100% and 99.7% respectively. The NPV for excluding ≥50% stenosis were 76.9% [46.2-95.0], 96.0% [91.4-98.5] and 99.0% [97.0-99.8].

Conclusion

In a symptomatic Asian population referred for CCTA, a CAC of zero had a high NPV for excluding severe coronary artery stenosis in patients with an intermediate to low probability of CAD, thus may have potential as a gatekeeper for further testing in this population. In patients with high pretest likelihood of CAD, CAC of zero does not reliably exclude significant CAD.

References:
Villines et al. Prevalence and Severity of Coronary Artery Disease and Adverse Events Among Symptomatic Patients With Coronary Artery Calcification Scores of Zero Undergoing Coronary Computed Tomography Angiography. J.Am.Coll.Cardiol. 2011; Vol. 58, No. 24, 2533–40

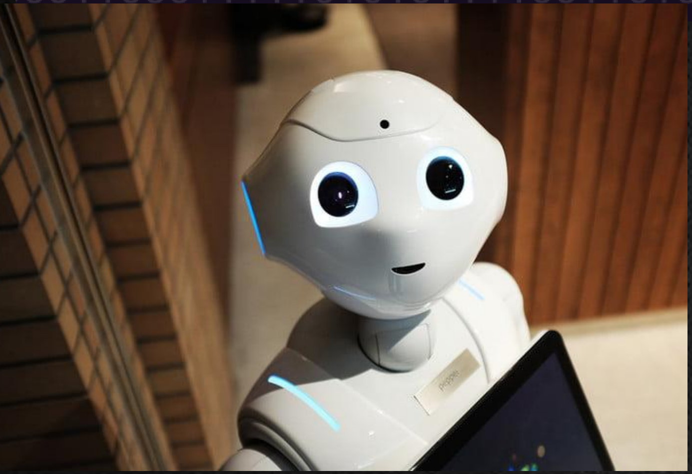
Van Werkhoven et al. Impact of Clinical Presentation and Pretest Likelihood on the Relation Between Calcium Acore and Computed Tomographic Coronary Angiography. Am J Cardiol 2010;106:1675-1679

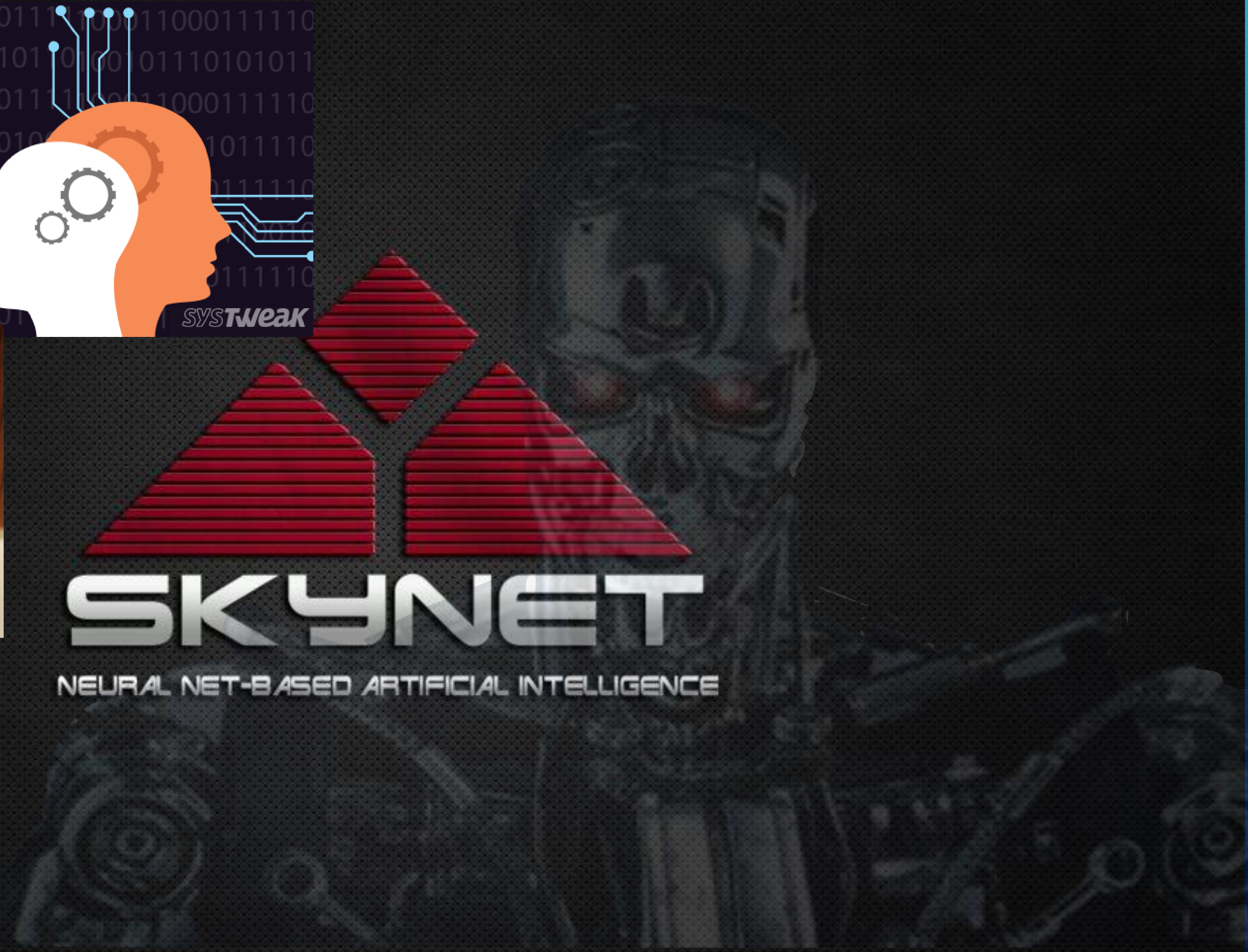
Acknowledgements
The authors would like to thank Max Wu and Kenneth Chew for their invaluable assistance in this project.

SUMMARY

- Big data and AI can be leveraged to help health care in many ways automating mundane processes
- Its an tool
- Cannot fully rely


The Rise of
Artificial Intelligence
SYSTWEAK




SKYNET
NEURAL NET-BASED ARTIFICIAL INTELLIGENCE