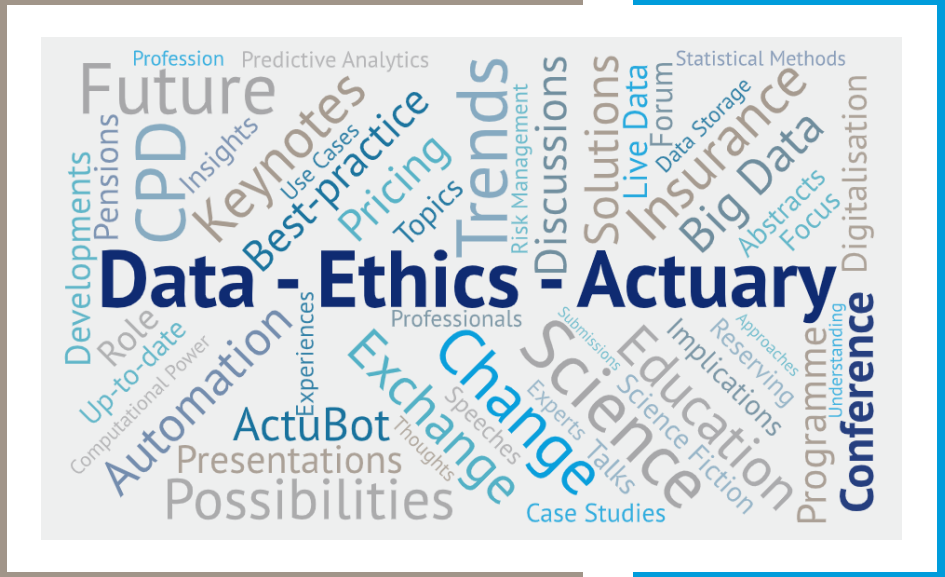


EHR Data Analytics
-Opportunities & Challenges

EAA e-Conference on
Data Science & Data Ethics

29 June 2021

Atreyee Bhattacharyya

IFoA AI & Automation in Life & Healthcare Working Party

WHAT IS EHR?

01

european actuarial academy

WHAT IS EHR

- EHR is an integrated digitised view of patients' medical and health records from various providers
- Captures health condition of patients across time
- Should allow secure exchange of information between any authorised providers
- Should enable optimal management of patients' health at an individual as well as at a population level

EHR DATA ITEMS

- Administrative and billing data
- Patient demographics
- Clinical notes and diagnoses
- Vital signs- height, weight
- Treatments-therapies, medication
- Details of immunisation
- Diagnostic imaging reports/radiology images
- Lab and test results
- Allergies
- Legal permissions

BENEFITS OF EHR?

02

BENEFITS OF EHR

Improves quality of care

- Quick access to patients‘ records
- Makes patient-physician communication better
- Population studies on EHR data give insight on the best care specific to the health characteristics of the patient
- Improves preventative care
- Automatically checks for reactions to medication and allergies

BENEFITS OF EHR

Better co-ordination of patient care and cost savings

- Helps avoid ordering duplicate tests
- Enables more efficient patient visits
- Allows better resource planning
- Achieves cost savings for physicians and hospitals as well as for patients and insurance companies

BENEFITS OF EHR

Higher patient satisfaction

Results from a survey[1] on patient-accessible electronic health records in Norway show that most respondents think that EHR data allows

- enhanced knowledge of their health condition
- easier control over their health status
- better self-care
- greater empowerment
- easier communication with health care provider

- The biggest benefit is experienced by patients with complex, long-term or chronic conditions

Advantages of EHR data analytics

- Predictive modelling using EHR data can be used to estimate clinical outcomes such as mortality, re-admission probability and help identify lives with higher risks

- Population studies based on EHR data can help in predicting disease progression by condition and hence can help with disease management

- EHR data combined with the emergence of genomic data is opening up the potential for precision healthcare and personalised medicine

APPLICATION OF AI/ML ON EHR DATA

03

APPLICATION OF AI/ML ON EHR DATA

- The adoption of EHR systems grew massively over the last few years
- In the US the percent of non-Federal acute care hospitals with the adoption of at least a Basic EHR system increased from 9.4 to 83.8% over the 7 years between 2008 and 2015 [2]
- The broad adoption of EHR presents a huge opportunity for the application of AI/ML techniques

APPLICATION OF AI/ML ON EHR DATA

- Meaningful analysis on clinical notes can be performed with the help of NLP techniques
- NLP techniques have the ability to analyse unstructured data including grammatical structure and help obtain meaningful information from the notes
- Deep neural networks such as CNN are used in radiology image classification
- In the next section, we discuss some research studies where AI/ML techniques have been used on EHR data

RESEARCH STUDY 1

- A study[3] was conducted on the MIMIC-III dataset (study period 2001 to 2012) to integrate heterogeneous data types across EHRs
- The data contained both static and longitudinal structured data and sequential unstructured data
- Static structured data included demographics such as, gender, marital status etc
- Temporal data included
  - Structured: Vital signs, pulse, systolic & diastolic blood pressure and many clinical test results
  - Unstructured: Sequential unstructured clinical notes, that could be used in predictive modelling using NLP

RESEARCH STUDY 1

- Structured data and unstructured text from clinical notes were combined directly through multi-modal deep neural networks for learning patient representation

- This was then used to predict patient outcomes such as mortality, long length of stay and re-admission

- The models were based on fusion CNN and fusion LSTM which can have general broader application without requiring domain knowledge

- The results showed that by combining unstructured clinical notes with structured data, the proposed models outperformed models that use either unstructured notes or structured data

RESEARCH STUDY 2

- Rajkomar et al. performed a study[4] that incorporated the entire EHR from University of California from 2012 to 2016, and the University of Chicago Medicine from 2009 to 2016

- Both dataset contained structured data including patient demographics, provider orders, diagnoses, procedures, medications, laboratory values, vital signs, and flowsheet data and one dataset additionally included free-text medical notes

- Three models were built-one based on recurrent neural networks, one on an attention-based TANN, and one on a neural network with boosted time-based decision stumps

- The results from these models were combined using ensembling

RESEARCH STUDY 2

- For prediction of inpatient mortality, the AUROC of this model was significantly higher than that of a traditional predictive model, which was a 28-factor logistic regression model
- The models outperformed the existing traditional models for predicting re-admission and length of stay as well
- The dataset had tens of thousands of potential predictor variables for each patient, including clinical notes
- The models were able to identify which predictors to include for a particular prediction without hand-selection of variables by an expert

RESEARCH STUDY 3

- Lauritsen et al. developed xAI-EWS – an explainable AI model with early warning score system which predicts acute critical illness using electronic health records[5]

- The model was composed of a temporal convolutional network (TCN) prediction module and a deep Taylor decomposition (DTD) explanation module, tailored to temporal explanations

- xAI-EWS was able to explain how the prediction outcomes are driven by specific input variables

- This allowed the clinicians to understand the reasoning behind the predictions

POTENTIAL USE CASES FOR ACTUARIES

04

LIFE & HEALTH INSURANCE UNDERWRITING

- Quicker and easier process than obtaining attending physicians' statement (APS), resulting in better conversion rates and lower underwriting costs
- Effective especially in current times of pandemic since this avoids face to face personal interactions
- Gives a longitudinal view and hence enables more accurate assessment of risk
- Automated underwriting through a rules engine should allow automatic acceptance and decline for the super healthy/super unhealthy lives
- Data analytics and predictive modelling can be used to produce risk score, on which underwriting loadings could be based

PRICING: NO/SPARSE PAST CLAIMS DATA

- EHR data can accurately predict hospitalisation and hence the expected cost of future claims for health insurance
- This is particularly important where past claims data is sparse or not available e.g. for new business
- Case study: Researchers analysed EHR data for patients at Atrius Health, a large multi-specialty group in Massachusetts, from June 2013 to November 2015[6]

PRICING: NO/SPARSE PAST CLAIMS DATA

- A risk score was derived based on patients' demographics , past utilisation, medical diagnoses and medications
- This score was used to predict hospitalisation in the next 6 months
- Predicted hospitalisation in the next 6 months based on the models produced a high degree of accuracy
- Models using EHR-only and claims-only data showed similar predictive power

CLAIMS MANAGEMENT

- Life and health insurers can benefit by using insights from population studies based on EHR data
- Longitudinal population studies based on EHR data can improve the accuracy of the existing understanding of progression of particular health conditions
- This can help better predict the expected cost of future claims for both health and life insurers, thus resulting in better claims management, more accurate pricing and reserving
- Understanding on disease progression should help develop disease management tools, which could reduce claims cost and increase retention

CHALLENGES & CONSIDERATIONS

05

CHALLENGES & CONSIDERATIONS

- A consultation survey[7] by the European Commission highlighted that the major obstacles to exchanging health data and advancing digital health and care in Europe are:
  - Risks of privacy breaches
  - Heterogeneity and lack of standardisation of EHRs and
  - Current lack of infrastructure
- The respondents encouraged the following actions:
  - developing standards for data quality and reliability
  - standardising electronic health records
  - proposing health related cybersecurity standards
  - supporting interoperability with open exchange formats

european actuarial academy

CHALLENGES & CONSIDERATIONS

- There is currently no 'one format' for EHR data, which affects the ease of use for actuaries

- Due to the sensitive nature of the data, EHRs are governed by strict government privacy laws such as the Health Insurance Portability and Accountability Act (HIPAA) in the US and data protection act in the UK– which has important implications for actuarial access

- Cost to insurers of piloting and implementing an EHR system is significant

- Insurers can have access to EHR data through vendors who obtain this data for insurers by logging into a patient portal and providing electronic authorisation or through aggregators that allow acquisition of applicant-authorised EHR from multiple vendors

CHALLENGES & CONSIDERATIONS

- But the amount of information available to insurers on these portals is often less than the full EHR
- Current hit rates on EHRs for insurance applicants are still low although this is expected to increase
- Wide access to EHR data would be available for actuarial/underwriting purposes, however this is still expected take a few years
- In summary, EHR has great potential to revolutionise insurance underwriting and healthcare analytics
- However, the challenges around ease of access, interoperability, low hit rates etc. need to be addressed before the full benefit can be realised

REFERENCES

[1] https://www.jmir.org/2020/2/e16144/

[2] https://dashboard.healthit.gov/evaluations/data-briefs/non-federal-acute-care-hospital-ehr-adoption-2008-2015.phphttps://www.jmir.org/2020/2/e16144/

[3] https://bmcmedinformdecismak.biomedcentral.com/track/pdf/10.1186/s12911-020-01297-6.pdf

[4] https://www.nature.com/articles/s41746-018-0029-1

[5] https://arxiv.org/pdf/1912.01266.pdf

[6] https://www.ajmc.com/view/predicting-hospitalizations-from-electronic-health-record-data

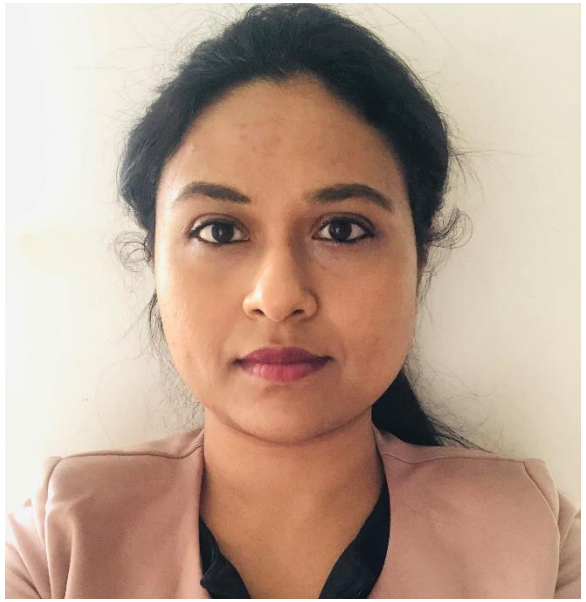
[7] https://ec.europa.eu/health/sites/default/files/ehealth/docs/2018_consultation_dsm_en.pdf

ABOUT ME

Atreyee Bhattacharyya is the chairperson of the Institute and Faculty of Actuaries' AI and Automation in life and healthcare working party and is a fellow actuary of the IFoA.

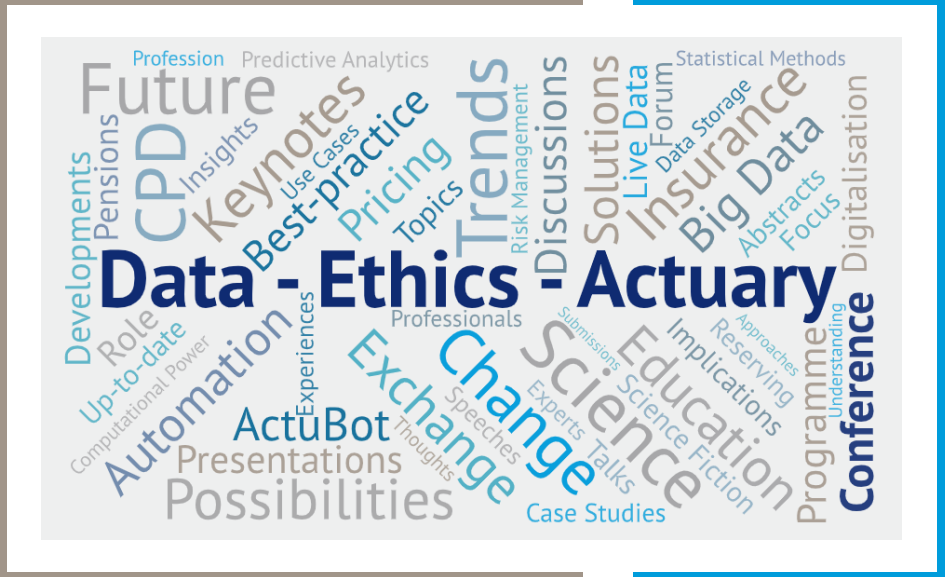
Atreyee currently works at Reinsurance Group of America (RGA) as a Health Actuary within Global Health. Prior to joining RGA, Atreyee worked for Bupa where she managed the SME pricing team. Before her time in Bupa, Atreyee worked for VitalityHealth in the technical pricing team where she worked closely with the data science team. Her role involved extensive modelling using both traditional actuarial methods and machine learning techniques.

Atreyee was a speaker in the first EAA conference in 2020. She has also presented in the Digital Actuary Virtual Summit, 2020 and Discovery Actuarial Conference in the years 2016 and 2017.



Atreyee Bhattacharyya

Chair, AI and Automation in life and healthcare working party, IFoA



THANK YOU

EAA e-Conference on
Data Science & Data Ethics

29 June 2021

Contact

Atreyee Bhattacharyya

IFoA AI & Automation in Life & Healthcare Working Party

Contact email:
atreyee.bhattacharyya@rgare.com