Synthesis | Insight

Bringing AI to Clinical Implementation At NUHS

Asst/Prof Kee Yuan Ngiam
MBBS (Lond), MRCS (Glasg), Mmed (Surg), FRCS (Edin)
Group Chief Technology Officer
Consultant Division of Thyroid and Endocrine Surgery
Department of Surgery
Deputy Chief Medical Information Officer
National University Hospital Singapore
Agenda

• What is Clinical AI?
• Clinical AI development process
  - Challenges and processes
• Advanced Clinical AI projects in NUH
• Deploying AI at NUH
  – Common advisory platform
What is Artificial Intelligence?

“The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages”
What about “Augmented Intelligence”

“Augmented intelligence refers to the effective use of information technology in augmenting human intelligence rather than to replace it.”
Healthcare’s 4th Industrial revolution

Parallels in the industrial revolution and the medical informatics revolution

1st Gen
Paper records, Knowledge intensive

2nd Gen
Electronic medical records, analysis intensive

3rd Gen
Analyzed insights, Experience intensive

1st
Mechanization, water and steam power

2nd
Mass production, assembly line, electricity

3rd
Computers and automation

4th
Cyber-physical systems
AI Vision in Healthcare

The unified vision of clinical data analytics is to augment healthcare practitioners in their delivery of safe and effective treatments expediently, and at the lowest cost.

This achieved through novel patient-centric, artificial intelligence-driven, clinical navigators to assist doctors in customizing treatments.
Clinical AI Development Process

**Treatment and Effectors**
Automated pop-up risk assessments and decision points, DICE bedside module

**Augmenting the doctor**
Automated diagnosis
Oncology models

**Improving Safety**
Readmission, falls, DVT prediction

**Pharmacodynamic prediction**
Predicting drug effectiveness and allergies

**Human machine interface**

**Data extraction and refining**

**Clinical problem (re)definition**

**Data analytics and model validation**

**Optimizing care**
Improved queueing theory
Strategy for patient navigation

**Value Driven Outcomes**
Modeling effective interventions to reduce costs

**Preventative and community health**
Modeling and incentivizing healthy behaviors
Layers of Datascience Development

Layer 6 – Clinical validation of model

Layer 5 - Predictive analytics

Layer 4 - Mapping

Layer 3- Automated Biostatistics

Layer 2- extraction and analysis

Layer 1- Verification, error checking
Challenges in AI Development

• Access to multiple data sources
• Messy, missing and noisy data
• Variation in coding standards
• Developing advanced AI tools
• Hardware issues
• Clinical trials
• Deployment in the EMR environment
Access to Multiple Data Sources

Centralized data repositories streamline translational clinical research:

1. Provide relevant data to researchers for research
2. Facilitate storage and maintenance of datasets
3. Promote collaboration
Advantages of Centralized Databases

• Security and Governance
  – Anonymisation, PDPA and data ownership

• Speed and ease of access
  – Tiered data access: retail users vs power users

• Hardware requirements
  – Facilitate storage and maintenance of datasets
  – Core vs Enhanced users (PDC, NSCC supercomputer)

• Expanding research datasets
  – REDCAP, I2B2 - Promote collaboration
# NUHS Datamart Phase 1 and 2

<table>
<thead>
<tr>
<th>Data Source</th>
<th>NNJ Enterprise Data Warehouse</th>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing Data Sources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAP-ISH</td>
<td></td>
<td>Trend Analysis</td>
</tr>
<tr>
<td>eIMR-Order</td>
<td></td>
<td>What-if Analysis</td>
</tr>
<tr>
<td>LIS</td>
<td></td>
<td>Charts &amp; Reports</td>
</tr>
<tr>
<td>New Data Sources</td>
<td></td>
<td>Adhoc Analysis</td>
</tr>
<tr>
<td>CCOE (Lab &amp; Rad Orders)</td>
<td></td>
<td>Interactive Dashboard</td>
</tr>
</tbody>
</table>

## EDW ODS

- Extract Transform Load
- SAP BusinessObjects
- Security & Access
- Audit
- Metadata

## NNJ Enterprise Data Warehouse

- Data Marts
  - PASS
  - CDMD
  - NUHS Data Mart
  - IDM

## Presentation

- Trend Analysis
- What-if Analysis
- Charts & Reports
- Adhoc Analysis
- Interactive Dashboard
- Mobile Devices
NUHS Datamart Phase 1 and 2

- On demand pull from H-Cloud, user friendly OBIEE interface
- Last 3 years only, 1 day lag time from drawing from live data
- Largely structured data, very little text data

<table>
<thead>
<tr>
<th>NUHS Data Sets Available on Datamart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Demographics</td>
</tr>
<tr>
<td>Radiology orders</td>
</tr>
<tr>
<td>Clinical diagnosis (problem list)</td>
</tr>
<tr>
<td>Laboratory orders and results</td>
</tr>
<tr>
<td>Services/procedures</td>
</tr>
<tr>
<td>Prescribed (out-patient) medications</td>
</tr>
<tr>
<td>Visits and encounters</td>
</tr>
<tr>
<td>In-patient medications</td>
</tr>
<tr>
<td>Billing information</td>
</tr>
</tbody>
</table>
Ultra Large Data Repositories (ULDRs)

- NUH-specific in house databases, all data from CCDR fields
- Huge and complex databases (>5TB)
- Large sizes preclude simple extraction methods due to database formats
- Ideal for longitudinal and epidemiological studies (>10k subjects) and power users who need large amounts of data for analysis and building AI tools
- Multiple databases exist with overlapping data and time frames – SDSD data repository, SPH Data repository, etc.
Ultra Large Data Repositories (ULDRs)

• Access to these databases governed centrally via AIO

• Obtain DSRB then apply for access via EPAS system on intranet

• Will need to provide own data analyst or programmer to extract data and analyze data IN NUHS

• Or collaborate with NUS researchers to work on the data

• May database specific issues that need to be addressed by individual users
I2B2 Research Framework

Virtual patient cohort

Genotype-phenotype correlations
Translational research
Rapid, low cost

Matched, de-identified biospecimens from clinical labs

(i2b2 analysis software (phenotype extraction))

CPSS2 EMR

NUHS restricted
Redcap Database

• NUHS is an approved institutional partner
• Endorsed by NUH CEO, GCIO as NUHS’s research database platform
• All departments within NUHS are encouraged to use Redcap as their clinical research capture database software

• Many new features:
  • Automatic population of data points from datamart
  • Insert redcap pdf data into CDOC entries
  • Access redcap from CDOC (single sign on)
Advanced Clinical AI projects in NUH
Development of Clinical AI

Clinical Tools

• Automated Diagnosis of Appendicitis Based on Clinical Notes

• SINGA-DRAGN: Longitudinal Analysis Of A Population’s Electronic Health Records Using Factor Graphs; A Case Study Of A Diabetic Surgical Population In Singapore

• Readmissions Prediction Using Deep Learning

• Disease Progression Modelling For Chronic Kidney Disease

• ICU SIRS prediction system

• Machine Learned Pancreatitis Severity Predictive Model
Development of Clinical AI

Infrastructural

• A Knowledge Base Pipeline For Medical Data Cleaning

• Medical Feature Embedding With Context Understanding

• Application Of Deep Learning In The Right-Siting Of Patients To Primary Care Providers

• Automated Disease Discovery From Lab Tests Using Machine Learning
Automated Diagnosis of Appendicitis Based on Clinical Notes

• One of the commonest causes of abdominal pain is appendicitis.

• Most of the relevant and useful information (e.g. signs and symptoms) are in the form of free text notes entered by medical personnel

• Brahmachari et al (2013) showed that the sensitivity of Alvarado Score is 66%

• Mainly to guide the use of CT scans in uncertain cases of appendicitis; exponential increase (2.9% to 82.4% in 22 years)* in CT utilization rate without improvements in outcomes^ 

• 5000 cases of appendicitis with 180,000 cases of abdominal pain controls over 10 years used to train machine
Objective

• To create a **scalable, free-text based, automated tool** to predict the diagnosis of appendicitis

• Mainly to guide the use of CT scans in uncertain cases of appendicitis; exponential increase **(2.9% to 82.4% in 22 years)** in CT utilization rate without improvements in outcomes

• Use of **NLP program** and **deep learning algorithm** to dissect emergency department doctor’s free text note

• 3000 cases of appendicitis with 200,000 cases of abdominal pain controls over 10 years used to train machine

* Trends in the Use of Medical Imaging to Diagnose Appendicitis at an Academic Medical Center. Journal of the American College of Radiology: 30 April 2016.

Workflow of the System

Patient seen at ED with history taken and examination performed

- Physician prompting tool - e.g. “Does patient have RIF Tenderness?”
- Inadequate information

Appendicitis Diagnosis System

Probability score

- Definite appendicitis: operation
- Probable appendicitis: for CT scan
- Not appendicitis: investigate other causes of abdominal pain

Final diagnosis

Use data to continuously train the model
Clinical Predictive Analytics for SIRS in ICU
4-Tier Risk Stratification

Highest probability for SIRS → HCP follow up is highly recommended

Prediction will be made when more data is available
Unified Platform for Messaging and Notification

Message 1, 2, ..., ICU team

Message and Action Taken Notifications

Health Care Provider List
- Ng Staff Nurse Kimberly
- Wong Dr. Yi Ling
- MICU Team

Notifications
- KH Management | Jun 07, 2016 at 01:25
  New message on patient: Patie...
- EM Nurse A3 | Jun 07, 2016 at 01:15
  New message on patient: P...
- AO Medical Doctor 1 | Jun 07, 2016 ...
A Deep Learning Approach for Readmission Prediction

Luo Zhaojing
School of Computing, NUS
Example: Readmission Prediction
CNN-based Framework
Disease Progression Modelling for Chronic Kidney Disease – A Deep Learning Approach

Kaiping Zheng
June 27th, 2017
Disease Progression Modelling

- Patients: \( x_D \), \( x_2 \), \( x_1 \)
- Education: \( g_S \), \( g_2 \), \( g_1 \)
- Gender, Age

Severity vs. Time

- \( y_1 \), \( y_2 \), \( y_3 \), \( y_4 \)
- \( t_1 \), \( t_2 \), \( t_3 \), \( t_4 \), \( t_5 \), \( t_6 \), \( t_7 \), \( t_8 \), \( t_9 \), \( t_{10} \)

Medical Features:
- DM
- HbA1C
- Glucose
- Insulin
- CKD
- Dialysis

CutPoint: \( t_{\psi} \)
Recurrent Neural Network

- Recurrent Neural Networks (RNN) are designed for processing sequential data and capturing dynamic behavior in data.

\[
\begin{align*}
\text{Vanilla RNN} & \\
\text{Long Short-Term Memory (LSTM)} & \\
\text{Gated Recurrent Unit (GRU)} &
\end{align*}
\]
Irregularity – Case Study

Mild yet deteriorating
- In the beginning, GFR indicates only moderately reduced kidney function. However, GFR decreases slowly over time before the 52nd week.
- After the 52nd week, our model predicts that the patient will suffer from a large drop in GFR, indicating the deterioration of kidney functioning.
- Our model would suggest healthcare workers to provide more aggressive interventions to Patient2 in advance.

This work has been submitted to CIKM 2017 (ACM International Conference on Information and Knowledge Management)
Singa-DRAGN: **Diabetic Readmissions** Graphical Network

Arjun P. Athreya  
CompGen Fellow and PhD Candidate

Asst/Prof. Kee Yuan Ngiam, NUHS  
Prof. Ravi Iyer, UIUC
Characteristics

10 year time period

Multiple readmissions

Multiple diagnoses per admission

Some readmission within 30 days

95% probability of readmission within 30 days after the first 30 day readmission
A knowledge base pipeline for cleaning medical data

Moaz Reyad
27 June 2017
ICD Cleaning Pipeline

Disease Description
  Stop word removal
  Splitting words
  Expand Abbreviations
  Query Knowledge Base

Result found
  no
  SPARQL

Result found
  yes
  Text Similarity
  ICD Code(s)

Wikipeida Infoboxes

RDF Knowledge Base

Lucene

ICD index
Medical Feature Embedding with Context Understanding

Xiangrui Cai*, Jinyang Gao§, Kee Yuan Nigam†, Beng Chin Ooi§, Kian-Lee Tan§, Xiaojie Yuan*

*Nankai University
§National University of Singapore
†National University Hospital

June 27, 2017
Raw Input (one-hot)

<table>
<thead>
<tr>
<th>Condition</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>pneumonia</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>bronchitis</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>obesity</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Neural word embedding

Long context scope

Short context scope

Context scopes of medical features

Experiments

- Clustering
- Nearest neighbor search
- Mortality prediction
Machine learned Pancreatitis Severity Predictive model

Muhd Muzamir
School of Computing, NUS

<table>
<thead>
<tr>
<th>Based on 47 features</th>
<th>F1_Score/ Harmonic Mean of PPV</th>
<th>Recall/ Sensitivity</th>
<th>Precision /PPV</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logit</td>
<td>0.675</td>
<td>0.619</td>
<td>0.757</td>
<td>0.658</td>
<td>0.633</td>
</tr>
<tr>
<td>SVM</td>
<td>0.607</td>
<td>0.515</td>
<td>0.752</td>
<td>0.721</td>
<td>0.593</td>
</tr>
<tr>
<td>NeuralNet</td>
<td>0.755</td>
<td>0.803</td>
<td>0.722</td>
<td>0.469</td>
<td>0.682</td>
</tr>
<tr>
<td>Glasgow &gt;2</td>
<td>0.524</td>
<td>0.397</td>
<td>0.809</td>
<td>0.843</td>
<td>0.560</td>
</tr>
<tr>
<td>Ranson &gt;2</td>
<td>0.733</td>
<td>0.873</td>
<td>0.638</td>
<td>0.160</td>
<td>0.610</td>
</tr>
</tbody>
</table>
My Health Emergency Resource Organizer SG - myHEROsg

An app with intuitive UI and powerful functionalities

NUS Computing Students
Nguyen Van Hoang, Wu Zefeng

NUS Medical Students
Kenneth Leong, Tan Jian Wei, Yuan Jing
Background

A&E units flooded with non-emergency cases

Such cases make up more than half of A&E patients in four public hospitals

Melissa Pang

...Continued

But the worst is at Kebon Tekko Hospital (KTPH) in Yishun. Last year, of the 400 A&E patients it saw daily, about 70 per cent were non-emergency or what the hospitals call P3 (Priority 3) cases.

These are patients with mild ailments, such as a spasm or stomach ache, that can be treated easily by a general practitioner (GP).

The high volume of P3 cases means that doctors are busy and have limited time to see more serious cases.

Of the 400 patients that KTPH sees daily, about 70% were non-emergency cases.

Ministry of Health (MOH) figures show that only fewer than 10 per cent of cases need to be hospitalised.

NUH's emergency medicine, head, Associate Professor Malcolm Maidan, pinpointed the ageing population as a reason, saying more patients with chronic conditions head to the A&E for their minor ailments.

But, the head of Tan Tock Seng Hospital's emergency department, Dr Tan Seow Yuen, noted that an elderly person with the same ailment as someone who is younger and healthier would require more attention.

His hospital, mainly for people living in central and northern Singapore, tends to have patients who are 10 years older than those in other parts of the island.

Overall, public hospital emergency attendance is on the rise. Every year for the past five years, it has been going up by 5.4 per cent, which is an additional 36,000 visits, said MOH spokesman.

Last year, the six restructured public hospitals treated more than 700,000 A&E cases.

CGH is seeing more P3 cases as well. In 2006, the daily average was 25. Last year, it was 43.

Patients pay a flat fee of about $95 for A&E consultations, which include medicine, basic tests and X-Ray services.

While a gradient of fees has been discussed, doctors like Associate Professor Mohan argue against it.

“Certain conditions, like heart attacks, need to be treated quickly. Price differentials can cause delays in getting medical help,” he said.

2Pang, M (2013, March 31) A&E units flooded with non-emergency cases. The Straits Times
Free Text Symptom Checker

I am coughing a lot and I feel feverish

You can see your nearest general practitioner for your medical condition.
Map View for Clinics and EDs and Appointments Tracker

INCOMING

Novena Medical Center Singapore
10 Sinaran Dr, Singapore 307506
Contact: 63976862
Time: 2017-06-20 22午後10:01
Price: $200.0

Healthway West Coast Clinic
727 Clementi West Street 2, #01-258, 120727
Contact: 67745901
Time: 2017-06-20 22午後10:01
Price: $18.0

International Medical Clinic
1 Orchard Boulevard, #14-06, 248649
Contact: 67334440
Time: 2017-06-20 22午後10:01
Price: $200.0

Dover Clinic & Surgery
28 Dover Cres, Singapore 130028
Automated Disease Discovery from Lab Tests Using Machine Learning

WANG DANDING
UBICOMP LAB
SCHOOL OF COMPUTING, NUS
Hyperparathyroidism

- An excess of parathyroid hormone
- Overactivity of one or more of 4 parathyroid glands
- Types of hyperparathyroidism
- Primary hyperparathyroidism:
  - Enlargement of parathyroid glands
  - → overproduce hormone
  - → high levels of calcium in the blood
  - → a variety of health problems
  - the most common treatment: surgery
Machine Learning Models

- Preliminary results:
  - Model comparison: trained on normalized data

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-nearest neighbor</td>
<td>66.4%</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>70.9%</td>
</tr>
<tr>
<td>Linear SVM</td>
<td>67.9%</td>
</tr>
<tr>
<td>RBF SVM</td>
<td>73.4%</td>
</tr>
<tr>
<td>Gaussian process</td>
<td>73.9%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision tree</td>
<td>83.5%</td>
</tr>
<tr>
<td>Neural Network</td>
<td>72.4%</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>83.6%</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>60.2%</td>
</tr>
<tr>
<td>QDA</td>
<td>52.2%</td>
</tr>
</tbody>
</table>
Bayesian Rule List

Pattern 1

Pattern 2

Pattern 3

Pattern 4

Positive Class

Area not covered by any of the positive patterns are classified as negative

Area covered by any of the positive patterns are classified as positive
Schematic for AI Implementation at NUH

- Demographic information
- ED notes
- Dispensed medication
- Visits and encounters
- Labtest results
- Radiology reports
- Procedures
- Discharge summaries
- Vital signs
- Inpatient medications
- Inpatient notes
- Outpatient notes

Production AI Modules
- Diagnosis module
- Readmissions module
- Complications module
- Disease progression mod
- VDO module
- Future Extensions

Deep machine learning
- Reinforced learning

Predicted clinical WARNING

CDOC

REDCAP

Pre-processing filter matrix

H-Cloud
WARNING

88.6% Chance of readmission

Ranked Factors:
1. Uncontrolled diabetes H/C 16
2. > 6 medications
3. 72.3% chance of post-op wound infection
4. Past readmissions due to social factors

Acknowledgment
Conclusion

- Clinical Healthcare AI can augment medical practitioners
- Significant challenges to data access and processing largely overcome
- Next challenge is to develop and implement machine learning protocols in a clinical setting
- Researchers should take advantages of the scalability and power of computing to augment clinician’s abilities
- Ultimately, the AI stool should function autonomously in the background to help physicians improve their quality of care and reduce costs
Thank you for your attention