MyDiabetesIQ
machine learning for decision support & outcome prediction in diabetes

Chris Sainsbury
Consultant Diabetologist, Glasgow

@csainsbury
SCI-Diabetes

c 300k patients with DM
c 500k historical patients

longitudinal data 1990 onwards:
biochemistry
phenotypic data
complications
mortality
prescribing data (encashment)
‘if we have one of the best datasets in the world, we should be able to build high quality algorithms to understand relationships within diabetes data…’
initial goals

i. suggest best next therapy / combination of therapies to achieve goals in multiple domains (hba1c reduction / blood pressure / mortality etc)

ii. predict complications (LLA, CV events)

iii. predict acute complications (hypoglycaemia etc)

iv. predict diabetes type at diagnosis, identify MODY etc

InnovateUK (Digital Health Technology Catalyst) 1M grant 2018-2021
what is the next best drug(s) for my patient?

virtual n = 1 drug trial

what drug should I prescribe to give this patient the best chance of having an HbA1c <60mmol/mol, with a reduction in blood pressure and BMI in 1 year?

taking into account their individual history of:

- HbA1c / BMI / blood pressure
- previously prescribed combinations of drug therapies
- how previous drugs have impacted on HbA1c / BMI / blood pressure
- sex
- age
- ethnicity
managing time series data - 1

hba1c

sbp

bmi

time
managing time series data - 2
drug combinations as words - for natural language processing approach

Drug Sentence: MF, GLP1_MF, GLP1_MF_SGLT2, GLP1_MF
→
Embedding, eg: 1, 2, 3, 2
→
input into RNN / LSTM
Recurrent Neural Network (LSTM)

extracts information from sequence of input
multi-dimensional RNN will learn interactions between input sequences over time
widely used – eg 30% google energy consumption running RNNs
SCI diabetes data input → R → visualisations etc → Keras → Python

environments used:
schematic of RNN/LSTM based classifier
HbA1c
SBP
BMI
Drug Combinations

Data Input Period

Prediction

Variable Metric

outcome measure of interest

virtual n = 1 trial approach

- add insulin
- add GLP1
- add SGLT2

predict response
simple problem: which of 4 combinations most likely to reduce hba1c by 10mmol/mol?
more complex problem: which of 16 combinations most likely to reduce hba1c to <60mmol/mol without causing weight gain?
The best combination to achieve HbA1c <60mmol/mol, with reduction in BMI is: dpp4_mf, dpp4_mf_sglt2, dpp4_mf_su, glp1_mf, glp1_mf_sglt2, glp1_mf_su, mf, mf_sglt2, mf_sglt2_su, mf_su. The best combination to achieve HbA1c <60mmol/mol, with reduction in SBP is: dpp4_glp1_mf, dpp4_mf, dpp4_mf_sglt2, dpp4_mf_su, dpp4_mf_tzd, glp1_mf, glp1_mf_sglt2, glp1_mf_su, mf, mf_sglt2, mf_sglt2_su, mf_su.
aims

simple system end 2018 (hybrid ML / rules based)

CE marking

development within SCI diabetes dataset, but with validation in external datasets