Towards generalisable in silico trials:

Exploring and augmenting the drug response data landscape

Chris Sainsbury

Queen Elizabeth University Hospital, Glasgow NHS Research Scotland (NRS) Senior Fellow Hon Senior Research Fellow, University of Birmingham co-founder kvatchii @csainsbury http://glucose.ai





J. Fluid Mech. (2012), vol. 705, pp. 280–305. © Cambridge University Press 2012 doi:10.1017/jfm.2012.220

Rarefaction and blood pressure in systemic and pulmonary arteries

Mette S. Olufsen¹, N. A. Hill²[†], Gareth D. A. Vaughan², Christopher Sainsbury³ and Martin Johnson⁴

¹ Department of Mathematics, North Carolina State University, Raleigh, NC 27502, USA
 ² School of Mathematics and Statistics, University of Glasgow, Glasgow G12 8QW, UK
 ³ BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow G12 8QQ, UK
 ⁴ Scottish Pulmonary Vascular Unit, Golden Jubilee National Hospital, Glasgow G81 4HX, UK

Biomech Model Mechanobiol (2014) 13:1137–1154 DOI 10.1007/s10237-014-0563-y

ORIGINAL PAPER

Numerical simulation of blood flow and pressure drop in the pulmonary arterial and venous circulation

M. Umar Qureshi · Gareth D. A. Vaughan · Christopher Sainsbury · Martin Johnson · Charles S. Peskin · Mette S. Olufsen · N. A. Hill



Fortran 90 ¶

Admission Glucose Number (AGN): A Point of Admission Score Associated With Inpatient Glucose Variability, Hypoglycemia, and Mortality

Journal of Diabetes Science and Technology 2019, Vol. 13(2) 213-220 © 2018 Diabetes Technology Society Article reuse guidelines: sagepub.com/iournals-permissions DOI: 10.1177/1932296818800722 journals.sagepub.com/home/dst (S)SAGE

Jennifer McKechnie, BSc, MBChB¹, Rahat Maitland, MD, FRCP¹, Christopher A. R. Sainsbury, BSc, MD, FRCP, FFCI¹, and Gregory C. Jones, MBChB, FRCP¹

Received: 13 October 2017 Revised: 6 December 2017 Accepted: 17 December 2017

DOI: 10.1111/dom.13193

WILEY

ORIGINAL ARTICLE

Visit-to-visit HbA1c variability and systolic blood pressure (SBP) variability are significantly and additively associated with mortality in individuals with type 1 diabetes: An observational study

Stuart S. Wightman | Christopher A. R. Sainsbury MD | Gregory C. Jones MBChB (19)

DIABETICMedicine

DOI: 10.1111/dme.13621

Short Report: Educational and Psychological Aspects Structured education using Dose Adjustment for Normal Eating (DAFNE) reduces long-term HbA_{1c} and HbA_{1c} variability

G. S. Walker¹, J. Y. Chen¹, H. Hopkinson², C. A. R. Sainsbury¹ and G. C. Jones¹

¹Diabetes Centre, Gartnavel General Hospital and ²Diabetes Centre, New Victoria Hospital, Glasgow, UK

Accepted 9 March 2018

Hypoglycemia and Clinical Outcomes in **Hospitalized Patients With Diabetes: Does** Association With Adverse Outcomes **Remain When Number of Glucose Tests** Performed Is Accounted For?

Iournal of Diabetes Science and Technology 2017, Vol. 11(4) 720-723 © 2017 Diabetes Technology Society Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/1932296816688012 journals.sagepub.com/home/dst (S)SAGE

Gregory C. Jones, MB ChB¹, Joseph G. Timmons, MB ChB¹, Scott G. Cunningham, PhD², Stephen J. Cleland, MD¹, and Christopher A. R. Sainsbury, MD¹



? more ambitious use of data resources available

SCI-Diabetes

>99% population coverage ~120 000 IDs with DM NHS GGC

comprehensive dataset primary/secondary care results: phenotyping information HbA1c, BMI, BP, bloods, prescriptions



Statistics are taken from the Scottish Diabetes Survey 2015.pdf

2015 SCI-Diabetes

- i. Large amounts of routinely collected data (RCD) are **unused** in clinical decision making
- ii. Clinical decisions are largely made on the basis of summary information (averages / snapshots)
- iii. Clinical guidelines are based on randomised controlled trial (RCT) data
- iv. RCTs not representative of the vast majority of the real world population (comorbid, elderly individuals excluded)
- v. RCD (retrospective cohort studies) analyses not highly regarded due to potential hidden biases etc

- i. Tackle the problem of summary data use by constructing analyses that use time series information
- ii. Calibrate RCD cohort investigations to RCT trial data understand associations and biases
- iii. Explore the effect size of interventions in populations that reflect the real world (comorbid, elderly etc)

predict the response (effect size within a specified domain) to an **arbitrary intervention** in an **arbitrary population** (or individual)

report effect size in both RCD and RCT-space

- predict optimum therapy choice for an individual / population (decision support)
- predict at-risk populations for adverse events (decision support)
- investigate potential for indication expansion for existing therapies (trial design)
- Investigate sources and size of hidden biases



Type 2 Diabetes - 8y duration No other diagnoses

- HbA1c 67 (high)
- 73 y male N
 - Type 2 Diabetes 3y duration COPD
- CASE

ASE

HbA1c 67 (high)

80 y female \mathbf{m}

- Type 2 Diabetes 20y duration
- ASE COPD, Psoriatic Arthritis, Ca Breast
 - HbA1c 67 (high)

Increasing complexity More like real world Less like evidence / guidelines Harder clinical decision making

Case 1

58 y maleType 2 Diabetes - 8y durationNo other diagnosesHbA1c 67 (high)

clinical context simple actual decision making remains difficult guidelines (in DM at least) ambiguous / open fine for specialists – not so much for generalists / primary care etc

theme 1 drug response prediction in diabetes: (virtual n=1 drug trial) myDiabetesIQ

related projects – deepmind collaboration / similarOme

Innovate UK (Digital Health Technology Catalyst) 1M grant 2018-2021



Funding competition Digital health technology catalyst 2017 round 1

UK businesses can apply for a share of up to £8 million to speed up development of new digital technology healthcare solutions.

Competition opens: Monday 31 July 2017 Competition closes: Wednesday 11 October 2017 12:00pm

Competition: Digital Health Technology Catalyst 2017 Round 1 Project Title: MyDiabetesIQ

GLUCOSE-LOWERING MEDICATION IN TYPE 2 DIABETES: OVERALL APPROACH

TO AVOID **CLINICAL INERTIA** FIRST-LINE THERAPY IS METFORMIN AND COMPREHENSIVE LIFESTYLE (INCLUDING WEIGHT MANAGEMENT AND PHYSICAL ACTIVITY) REASSESS AND IF HbA, ABOVE TARGET PROCEED AS BELOW MODIFY TREATMENT REGULARLY (3-6 MONTHS) NO ESTABLISHED ASCVD OR CKD WITHOUT ESTABLISHED ASCVD OR CKD **ASCVD PREDOMINATES** HE OR CKD PREDOMINATES COMPELLING NEED TO MINIMISE WEIGHT EITHER/ GAIN OR PROMOTE WEIGHT LOSS **COMPELLING NEED TO MINIMISE HYPOGLYCAEMIA** COST IS A MAJOR ISSUE9-10 OR PREFERABLY SGLT2i with evidence of reducing EITHER HF and/or CKD progression in SGLT2i with OR GLP-1 RA with CVOTs if eGFR adequate³ GLP-1 RA proven CVD SGLT2i² DPP-4i GLP-1 RA TZD SU6 TZD¹⁰ good efficacy SGLT2i² ----- OR ----with proven benefit1, for weight loss⁸ CVD benefit¹ if eGFR If SGLT2i not tolerated or J. 1 1 J J. \mathbf{J} adequate² contraindicated or if eGFR less If HbA, If HbA, If HbA, If HbA, than adequate² add GLP-1 RA If HbA, above target If HbA, above target above target above target above target above target with proven CVD benefit¹ J J J. \mathbf{v} 4 GLP-1 RA SGLT2i² SGLT2i² SGLT2i² OR OR GLP-1 RA with If HbA, above target If HbA, above target OR OR DPP-4i DPP-4i SGLT2i² good efficacy TZD¹⁰ SU⁶ TZD TZD OR OR for weight loss⁸ If further intensification is required or TZD Avoid TZD in the setting of HF GLP-1 RA patient is now unable to tolerate 1 \mathbf{J} \mathbf{v} 1 \mathbf{J} \mathbf{J} Choose agents demonstrating CV safety: GLP-1 RA and/or SGLT2i, choose Consider adding the other class If HbA, above target If HbA, above target If HbA, above target agents demonstrating CV safety: with proven CVD benefit¹ 1 4 1 1 Consider adding the other class DPP-4i (not saxagliptin) in the setting (GLP-1 RA or SGLT2i) with proven Continue with addition of other agents as outlined above Insulin therapy basal insulin with If triple therapy required or SGLT2i of HF (if not on GLP-1 RA) **CVD** benefit lowest acquisition cost Basal insulin⁴ and/or GLP-1 RA not tolerated or DPP-4i if not on GLP-1 RA OR SU⁶ contraindicated use regimen with Basal insulin⁴ If HbA, above target Consider DPP-4i OR SGLT2i with lowest risk of weight gain TZD⁵ lowest acquisition cost¹⁰ PREFERABLY SU⁶ Consider the addition of SU⁶ OR basal insulin: DPP-4i (if not on GLP-1 RA) based on weight neutrality Choose later generation SU with lower risk of hypoglycaemia Consider basal insulin with lower risk of hypoglycaemia⁷ If DPP-4i not tolerated or 1. Proven CVD benefit means it has label indication of reducing CVD events. For GLP-1 RA strongest 5. Low dose may be better tolerated though less well studied for CVD effects contraindicated or patient already on evidence for liraglutide > semaglutide > exenatide extended release. For SGLT2i evidence 6. Choose later generation SU with lower risk of hypoglycaemia GLP-1 RA, cautious addition of: modestly stronger for empagliflozin > canagliflozin. 7. Degludec / glargine U300 < glargine U100 / detemir < NPH insulin • SU⁶ • TZD⁵ • Basal insulin 2. Be aware that SGLT2i vary by region and individual agent with regard to indicated level of eGFR 8. Semaglutide > liraglutide > dulaglutide > exenatide > lixisenatide for initiation and continued use 9. If no specific comorbidities (i.e. no established CVD, low risk of hypoglycaemia and lower 3. Both empagliflozin and canagliflozin have shown reduction in HF and reduction in CKD priority to avoid weight gain or no weight-related comorbidities) progression in CVOTs 10. Consider country- and region-specific cost of drugs. In some countries TZDs relatively more 4. Degludec or U100 glargine have demonstrated CVD safety expensive and DPP-4i relatively cheaper Fig. 2 Glucose-lowering medication in type 2 diabetes: overall approach

Diapetologia

what is the next best drug(s) for my patient? virtual n = 1 drug trial

eg 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 – numerical data



time

managing time series data – prescription data

2017-06-30, Lantus 100units/ml solution for injection 3ml pre-filled SoloStar pen (Sanofi), 6.1.1.2 2017-09-15, NovoRapid FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.1 2017-05-08, Levemir FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.2 2017-06-30, Humalog KwikPen 100units/ml solution for injection 3ml pre-filled pen (Eli Lilly and Company Ltd), 6.1.1.1 2016-12-05,Gliclazide 80mg tablets,6.1.2.1 2017-09-15, Levemir FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.2 2017-05-08, NovoRapid FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.1 2017-04-04, NovoRapid FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.1 2017-04-04, Levemir FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.2 2017-08-01, NovoRapid FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.1 2017-06-30,Gliclazide 80mg tablets,6.1.2.1 2017-03-06, NovoRapid FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.1 2017-03-06, Levemir FlexPen 100units/ml solution for injection 3ml pre-filled pen (Novo Nordisk Ltd), 6.1.1.2 2017-06-30, Metformin 850mg tablets, 6.1.2.2 2016-12-15, GlucoGel 40% gel original (BBI Healthcare Ltd), 6.1.4.0 2017-01-30, NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd), 6.1.1.1 2016-11-04, NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd), 6.1.1.1 2017-01-30, Lantus 100 units/ml solution for injection 3ml pre-filled SoloStar pen (Sanofi), 6.1.1.2 2016-11-22, Metformin 850mg tablets, 6.1.2.2 2017-04-24, Metformin 850mg tablets, 6.1.2.2 2017-04-24,Gliclazide 80mg tablets,6.1.2.1 2016-11-04, Lantus 100 units/ml solution for injection 3ml pre-filled SoloStar pen (Sanofi), 6.1.1.2 2016-11-22, Gliclazide 80mg tablets, 6.1.2.1 2016-12-16, GlucaGen Hypokit 1mg powder and solvent for solution for injection (Novo Nordisk Ltd), 6.1.4.0 2017-03-07, Gliclazide 80mg tablets, 6.1.2.1 2017-05-31, Lantus 100units/ml solution for injection 3ml pre-filled SoloStar pen (Sanofi), 6.1.1.2 2017-01-06,Gliclazide 80mg tablets,6.1.2.1 2017-04-18, NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd), 6.1.1.1 2017-06-30, Gliclazide 80mg tablets, 6.1.2.1 2017-03-07, Gliclazide 80mg tablets, 6.1.2.1 2017-03-06, NovoRapid FlexPen 100units/ml solution for injection 3ml ..., 6.1.1.1 2017-03-06,Levemir FlexPen 100units/ml solution for injection 3ml pr...,6.1.1.2 2016-11-04, NovoRapid Penfill 100units/ml solution for injection 3ml ..., 6.1.1.1 2017-01-25,Lantus 100units/ml solution for injection 3ml pre-filled ...,6.1.1.2 2017-08-17, Gliclazide 80mg tablets, 6.1.2.1 2017-05-24, Gliclazide 80mg tablets, 6.1.2.1 2017-04-25, Gliclazide 80mg tablets, 6.1.2.1 2017-03-28, Gliclazide 80mg tablets, 6.1.2.1 2017-02-27, Gliclazide 80mg tablets, 6.1.2.1 2017-02-02, Gliclazide 80mg tablets, 6.1.2.1 2017-01-10,Gliclazide 80mg tablets,6.1.2.1 2016-11-16,Gliclazide 80mg tablets,6.1.2.1 2017-07-19, Gliclazide 80mg tablets, 6.1.2.1 2017-02-24, Humulin M3 KwikPen 100units/ml suspension for injection 3ml pre-filled pen (Eli Lilly and Company Ltd), 6.1.1.51 2017-01-04, Metformin 500mg tablets, 6.1.2.2 2017-01-04, Humulin M3 KwikPen 100units/ml suspension for injection 3ml pre-filled pen (Eli Lilly and Company Ltd), 6.1.1.51 2016-11-14, Metformin 500mg tablets, 6.1.2.2 2016-12-05, Humulin M3 KwikPen 100units/ml suspension for injection 3ml pre-filled pen (Eli Lilly and Company Ltd), 6.1.1.51 2017-02-28, Metformin 500mg tablets, 6.1.2.2 2017-01-30, Metformin 500mg tablets, 6.1.2.2 2017-04-24, Metformin 500mg tablets, 6.1.2.2 2017-04-24, Metformin 500mg tablets, 6.1.2.2 2017-04-10, Humulin M3 KwikPen 100units/ml suspension for injection 3..., 6.1.1.51 2017-03-27, Metformin 500mg tablets, 6.1.2.2 2017-02-28, Metformin 500mg tablets, 6.1.2.2 2017-02-24, Humulin M3 KwikPen 100units/ml suspension for injection 3...,6.1.1.51 2017-01-30, Metformin 500mg tablets, 6.1.2.2 2017-01-04, Metformin 500mg tablets, 6.1.2.2 2017-01-04, Humulin M3 KwikPen 100units/ml suspension for injection 3..., 6.1.1.51 2016-12-08, Metformin 500mg tablets, 6.1.2.2 2016-12-05, Humulin M3 KwikPen 100units/ml suspension for injection 3..., 6.1.1.51 2017-03-27, Metformin 500mg tablets, 6.1.2.2 2017-03-27, FreeStyle Optium testing strips (Abbott Laboratories Ltd), 6.1.6.0

,2017-06-21,Metformin 500mg tablets,6.1.2.2 ,2017-06-21,Gliclazide 80mg tablets,6.1.2.1 ,2017-03-24,Gliclazide 80mg tablets,6.1.2.1 ,2017-03-24,Metformin 500mg tablets,6.1.2.2 ,2017-02-21,Gliclazide 80mg tablets,6.1.2.1 ,2017-06-14,Pioglitazone 30mg tablets,6.1.2.3 ,2017-04-28,Vildagliptin 50mg / Metformin 850mg tablets,6.1.2.3 ,2017-04-07, Metformin 500mg tablets, 6.1.2.2 ,2017-04-07,NovoMix 30 FlexPen 100units/ml suspension for injection 3...,6.1.1.51 ,2017-03-09,Sitagliptin 100mg tablets,6.1.2.3 ,2017-04-24,Gliclazide 80mg tablets,6.1.2.1 ,2017-01-30,NovoRapid Penfill 100units/ml solution for injection 3ml ...,6.1.1.1 ,2017-01-06,Gliclazide 80mg tablets,6.1.2.1 ,2016-12-16,GlucaGen Hypokit 1mg powder and solvent for solution for ...,6.1.4.0 ,2016-12-15,GlucoGel 40% gel original (BBI Healthcare Ltd),6.1.4.0 ,2016-11-22,Gliclazide 80mg tablets,6.1.2.1 ,2017-01-06,Metformin 850mg tablets,6.1.2.2 2017-04-28, Pioglitazone 30mg tablets, 6.1.2.3 ,2017-02-28,Glipizide 5mg tablets,6.1.2.1 ,2017-09-27,Glipizide 5mg tablets,6.1.2.1 ,2017-01-09,GlucoGel 40% gel original (BBI Healthcare Ltd),6.1.4.0 ,2017-03-07,Metformin 850mg tablets,6.1.2.2 ,2016-11-14, FreeStyle Optium testing strips (Abbott Laboratories Ltd), 6.1.6.0 ,2016-12-08,Metformin 500mg tablets,6.1.2.2 ,2017-04-10,Humulin M3 KwikPen 100units/ml suspension for injection 3ml pre-filled pen (Eli Lilly and Company Ltd),6.1.1.51 ,2017-06-21,Gliclazide 80mg tablets,6.1.2.1 ,2016-11-14,Metformin 500mg tablets,6.1.2.2 ,2017-05-22,Gliclazide 80mg tablets,6.1.2.1 ,2017-04-24,Gliclazide 80mg tablets,6.1.2.1 ,2017-04-24,Metformin 500mg tablets,6.1.2.2 ,2017-01-26,Metformin 500mg tablets,6.1.2.2 ,2017-01-26,Gliclazide 80mg tablets,6.1.2.1 ,2016-12-21,Metformin 500mg tablets,6.1.2.2 ,2016-12-21,Gliclazide 80mg tablets,6.1.2.1 ,2016-12-19,Metformin 500mg tablets,6.1.2.2 ,2017-10-24,Pioglitazone 30mg tablets,6.1.2.3 ,2017-10-24,Vildagliptin 50mg / Metformin 850mg tablets,6.1.2.3 ,2017-08-14, Vildagliptin 50mg / Metformin 850mg tablets, 6.1.2.3 ,2017-08-14,Pioglitazone 30mg tablets,6.1.2.3 ,2017-03-03,Vildagliptin 50mg / Metformin 850mg tablets,6.1.2.3 ,2017-03-03,Pioglitazone 30mg tablets,6.1.2.3 ,2017-01-11,Pioglitazone 30mg tablets,6.1.2.3 ,2016-11-12, Vildagliptin 50mg / Metformin 850mg tablets, 6.1.2.3 ,2017-06-14,Vildagliptin 50mg / Metformin 850mg tablets,6.1.2.3 ,2016-11-08,Pioglitazone 30mg tablets,6.1.2.3 ,2016-12-30,Vildagliptin 50mg / Metformin 850mg tablets,6.1.2.3 ,2017-05-17, NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd),6.1.1.1 ,2017-05-17, Levemir Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd),6.1.1.2 ,2017-05-17,NovoRapid Penfill 100units/ml solution for injection 3ml ...,6.1.1.1 ,2017-05-17,Levemir Penfill 100units/ml solution for injection 3ml ca...,6.1.1.2 ,2017-03-08,NovoRapid Penfill 100units/ml solution for injection 3ml ...,6.1.1.1 ,2016-12-22, NovoRapid Penfill 100units/ml solution for injection 3ml ...,6.1.1.1 ,2016-12-22, Levemir Penfill 100units/ml solution for injection 3ml ca...,6.1.1.2 ,2017-03-08, NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd), 6.1.1.1 ,2017-03-08,Levemir Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd),6.1.1.2 ,2016-12-22,NovoRapid Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd),6.1.1.1 ,2016-12-22,Levemir Penfill 100units/ml solution for injection 3ml cartridges (Novo Nordisk Ltd),6.1.1.2 ,2017-06-15,NovoMix 30 FlexPen 100units/ml suspension for injection 3...,6.1.1.51 ,2017-06-15,Metformin 500mg tablets,6.1.2.2 ,2017-01-25,Sitagliptin 100mg tablets,6.1.2.3 ,2017-01-24,Glipizide 5mg tablets,6.1.2.1 ,2016-12-13,Glipizide 5mg tablets,6.1.2.1 2016 11 02 Clipizido Emg tablete 6 1 2 1

🦆 python[™] 🔣 Keras

schematic of neural network classifier



nn output vs logistic regression HbA1c TS vs HbA1c median / CV

predicting mortality at 3y

8 year runin analysis : outcome: all-cause mortality at 3y



nn output AUROC 0.6

logistic regression AUROC 0.51





T1 multistep n=26860 target = gradient of line

probability of HbA1c reduction >10mmol/mol at 1y



drug combinations (400)

probability of HbA1c reduction >10mmol/mol at 1y n = 1450 (initial HbA1c >60mmol/mol)

Case 2



clinical context slightly more complex no good evidence to guide management forced to extrapolate from multiple (siloed) disease area guidelines

but – many examples in the real world. can we use the examples within RCD to explore treatment options, and relate this to current evidence

theme 2 RCD / RCT calibration Extend in RCD to multimorbidity and map to RCT space A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook^{*}

Brett R. Gordon Kellogg School of Management Northwestern University Florian Zettelmeyer Kellogg School of Management Northwestern University and NBER

Neha Bhargava Facebook Dan Chapsky Facebook

September 23, 2018

Abstract

Measuring the causal effects of digital advertising remains challenging despite the availability of granular data. Unobservable factors make exposure endogenous, and advertising's effect on outcomes tends to be small. In principle, these concerns could be addressed using randomized controlled trials (RCTs). In practice, few online ad campaigns rely on RCTs, and instead use observational methods to estimate ad effects. We assess empirically whether the variation in data typically available in the advertising industry enables observational methods to recover the causal effects of online advertising. Using data from 15 US advertising experiments at Facebook comprising 500 million user-experiment observations and 1.6 billion ad impressions, we contrast the experimental results to those obtained from multiple observational models. The observational methods often fail to produce the same effects as the randomized experiments. even after conditioning on extensive demographic and behavioral variables. In our setting, advances in causal inference methods do not allow us to isolate the exogenous variation needed to estimate the treatment effects. We also characterize the incremental explanatory power our data would require to enable observational methods to successfully measure advertising effects. Our findings suggest that commonly used observational approaches based on the data usually available in the industry often fail to accurately measure the true effect of advertising.

> Gordon, Brett R., Florian Zettelmeyer, Neha Bhargava, and Dan Chapsky. n.d. "A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3033144.

Case 3

80 y femaleType 2 Diabetes - 20y durationCOPD, Psoriatic Arthritis, Ca BreastHbA1c 67 (high)

clinical context much more complex no good evidence to guide management forced to extrapolate from multiple (siloed) disease area guidelines

this time – not so many examples in the real world. ? can we use generative techniques to develop multiple unobserved examples of the rare class

theme 3 generative techniques to augment RCD

related projects – synthetic datasets, rare class identification

requirements for synthetic data generator

- generate multi-dimensional time series data
- reflect distributions of individual parameters
- include interactions/associations between parameters over time
- allow training models on synthetic data that will perform well on real data

Potential applications of synthetic data generator

synthetic data generation

- upsampling rare classes
- potential to help with another problem data sharing

rare class identification - difficult as small sample size to work from (but may be confident in label accuracy)



Differentially Private Generative Adversarial Network

Liyang Xie¹, Kaixiang Lin¹, Shu Wang², Fei Wang³, Jiayu Zhou¹ ¹Computer Science and Engineering, Michigan State University ²Department of Computer Science, Rutgers University ³Department of Healthcare Policy and Research, Weill Cornell Medical School {xieliyan,linkaixi}@msu.edu,sw498@cs.rutgers.edu,few2001@med.cornell.edu,jiayuz@msu.edu

ABSTRACT

Generative Adversarial Network (GAN) and its varian cently attracted intensive research interests due to th theoretical foundation and excellent empirical performa erative models. These tools provide a promising direct studies where data availability is limited. One commo GANs is that the density of the learned generative d could concentrate on the training data points, meanin can easily remember training samples due to the high i plexity of deep networks. This becomes a major con GANs are applied to private or sensitive data such as p ical records, and the concentration of distribution m critical patient information. To address this issue, in thi propose a differentially private GAN (DPGAN) model, in achieve differential privacy in GANs by adding carefull noise to gradients during the learning procedure. We r orous proof for the privacy guarantee, as well as comp empirical evidence to support our analysis, where we de that our method can generate high quality data points a able privacy level.

CCS CONCEPTS

• Computing methodologies → Neural networks; • systems organization → Neural networks; • Securit vacy → Privacy-preserving protocols;

KEYWORDS

Deep Learning; Differential Privacy; Generative model

arXiv:1802.06739

REAL-VALUED (MEDICAL) TIME SERIES GENERA-TION WITH RECURRENT CONDITIONAL GANS

Stephanie L. Hyland*

ETH Zurich, Switzerland Tri-Institutional Training Program in Computational Biology and Medicine, Weill stephanie.hyland@inf.ethz.ch

Cristóbal Esteban*

ETH Zurich, Switzerland cristobal.esteban@inf.ethz.ch

Gunnar Rätsch ETH Zurich, Switzerland raetsch@inf.ethz.c

ABSTRACT

Generative Adversarial Networks (GANs) have shown remarkable success framework for training models to produce realistic-looking data. In this wor propose a Recurrent GAN (RGAN) and Recurrent Conditional GAN (RCGA produce realistic real-valued multi-dimensional time series, with an emphas their application to medical data. RGANs make use of recurrent neural net (RNNs) in the generator and the discriminator. In the case of RCGANs, be these RNNs are conditioned on auxiliary information. We demonstrate our m in a set of toy datasets, where we show visually and quantitatively (using sa likelihood and maximum mean discrepancy) that they can successfully get realistic time-series. We also describe novel evaluation methods for GANs, we generate a synthetic labelled training dataset, and evaluate on a real te the performance of a model trained on the synthetic data, and vice-versa illustrate with these metrics that RCGANs can generate time-series data u for supervised training, with only minor degradation in performance on rea data. This is demonstrated on digit classification from 'serialised' MNIST a training an early warning system on a medical dataset of 17,000 patients fro intensive care unit. We further discuss and analyse the privacy concerns that arise when using RCGANs to generate realistic synthetic medical time series and demonstrate results from differentially private training of the RCGAN.

arXiv:1706.02633

Privacy-preserving generative deep neural networks support clinical data sharing

Authors: Brett K. Beaulieu-Jones¹, Zhiwei Steven Wu², Chris Williams³, Casey S. Greene³*

¹Genomics and Computational Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

²Computer and Information Science, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

³Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA.

*To whom correspondence should be addressed: csgreene@upenn.edu

One Sentence Summary: Deep neural networks can generate shareable biomedical data to allow reanalysis while preserving the privacy of study participants.

Abstract: Though it is widely recognized that data sharing enables faster scientific progress, the sensible need to protect participant privacy hampers this practice in medicine. We train deep neural networks that generate synthetic subjects closely resembling study participants. Using the SPRINT trial as an example, we show that machine-learning models built from simulated participants generalize to the original dataset. We incorporate differential privacy, which offers strong guarantees on the likelihood that a subject could be identified as a member of the trial. Investigators who have compiled a dataset can use our method to provide a freely accessible public version that enables other scientists to perform discovery-oriented analyses. Generated data can be released alongside analytical code to enable fully reproducible workflows, even when privacy is a concern. By addressing data sharing challenges, deep neural networks can facilitate the rigorous and reproducible investigation of clinical datasets.

https://doi.org/10.1101/159756



differential privacy inhibits data leakage

- quality vs security tradeoff
- privacy budgets
- computational limits

glucose.ai group, NHS GGC

Dr Greg Jones Dr David Carty Dr Jim Boyle Dr Debbie Morrison Dr Sean Harbison

myDiabetesIQ Dr Debbie Wake

Dr Scott Cunningham Dr Felix Agakov Dr Nicky Conway Dr Sam Philip

HDR UK Midlands Prof Alastair Denniston

kvatchii

Prof Sandosh Padmanabhan Ajay Sharma Dr Shyam Visweswaran

redstar Andrew Conkie

thinkingAl group, University of Birmingham Dr Krish Nirantharakumar

Krishna Gokhale Dr Malcolm Price Dr Chris Yau Dr Tom Taverner Prof Peter Tino