rEHR: An R package for manipulating and analysing Electronic Health Records data

Rosa Parisi, PhD
University of Manchester

Springate D, Olier I, Reeves D, Kontopantelis E

NIHR Statistics Group, 21st Jan 2019

MANCHESTER 1824
The University of Manchester

Introduction

Centre for Pharmacoepidemiology & Drug Safety, University of Manchester (UoM)

Statistics group, Centre for Primary Care, (UoM)

Dr Springate

Dr Olier

Dr Reeves

Prof Kontopantelis

MANCHESTER 1824
The University of Manchester

Outline

1. Primary care databases (CPRD)
2. rEHR package
3. Loading data and importing files
4. Querying the database
5. Building a cohort
6. Building a code list
7. Summary

Electronic Health Records (EHR)s

- Systematic collection of electronic records containing health information of patients
- Complex databases more and more available to researchers
- However, tools for extraction, data quality, manipulation of EHR databases are less available and often an issue
- Building a dataset ready for analysis is challenging and time consuming

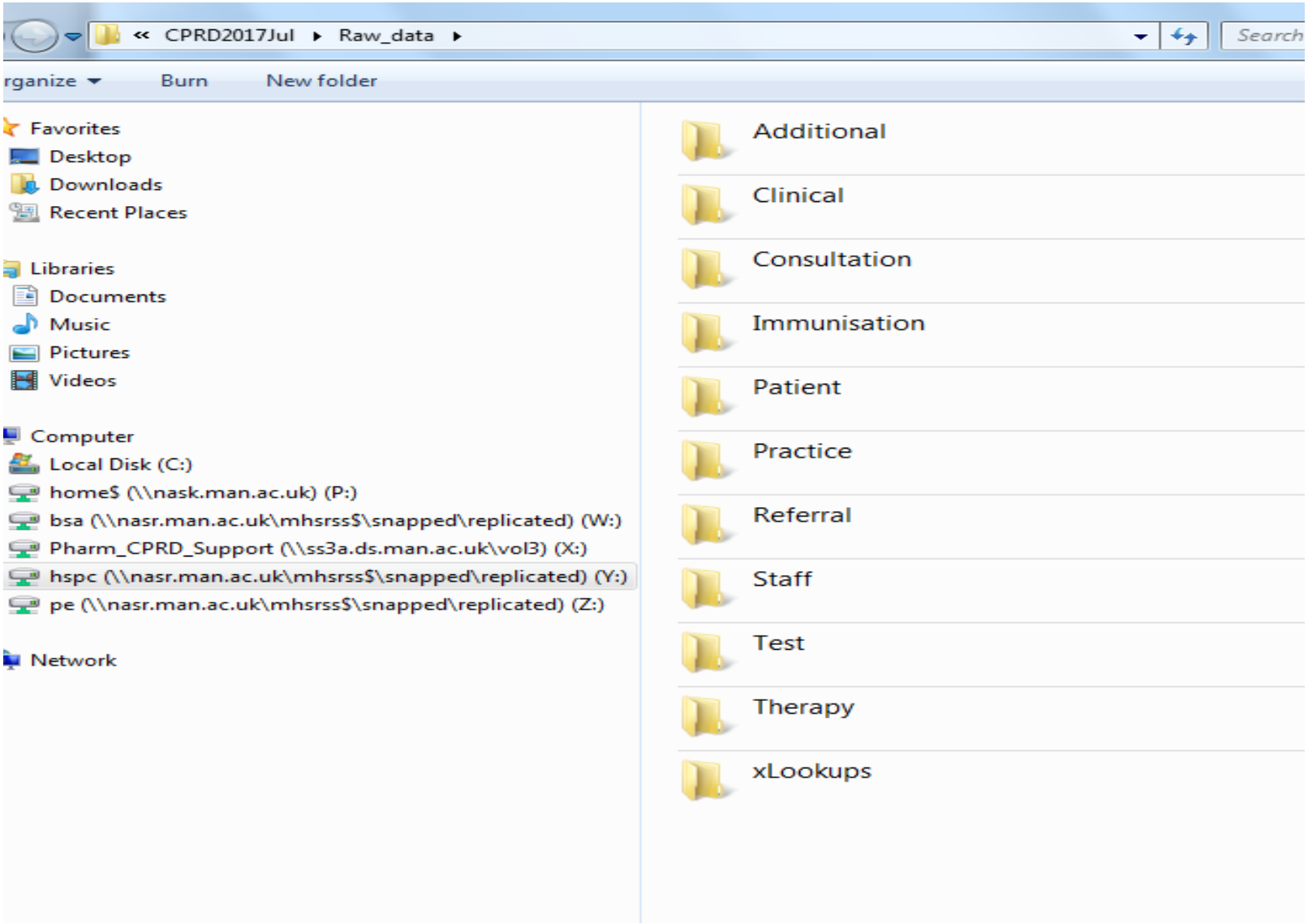
UK Primary Care Databases

- Primary care database store data in complex relational and nested structures
  - Broken down in numerous tables due to the volume of data
  - Text files need to be imported into powerful analysis/database management software
  - All events are entered in codes

Primary Care Database Structure (CPRD)

- Events files:
  - **Demographics**: year of birth, sex
  - **Clinical events**: symptoms, signs and diagnoses
  - **Referrals** to secondary care
  - **Immunisations**
  - **Therapy**: data relating to all prescriptions issued by a GP
  - **Tests**
- Look-up tables
  - **Medical codes**
  - **Product codes**

Clinical Practice Research Datalink



Clinical Practice Research Datalink

hspc (\\nasr.man.ac.uk\mhsrss$\snapped\replicated) (Y:) ▸ CPRD2017Jul ▸ Raw_data ▸ Clinical ▸

Burn New folder

p001_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:20 Size: 63.5 MB
p002_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 117 MB
p003_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 76.9 MB
p004_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 25.8 MB
p005_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 35.1 MB
p006_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 42.3 MB
p007_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 6.74 MB
p008_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 4.90 MB
p009_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 38.2 MB
p010_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 21.9 MB
p011_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 38.6 MB
p012_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:21 Size: 26.4 MB
p013_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:22 Size: 108 MB
p014_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:22 Size: 15.1 MB
p015_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:22 Size: 28.8 MB
p016_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:22 Size: 18.1 MB
p017_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/07/2017 10:22 Size: 65.4 MB

Clinical Practice Research Datalink

Filter

	patid	eventdate	sysdate	constype	consid	medcode	staffid	textid	episode	enttype	adid
1	1001	19860618	19970309	2	103	1	0	0	0	1	1
2	1001	19880920	19970309	2	100	60	0	0	0	4	2
3	1001	19930324	19970309	2	11	4447	0	0	0	5	3
4	1001	19930330	19970309	2	101	3	0	0	0	14	4
5	1001	19860101	19970307	4	3	204	0	0	0	15	0
6	1001	19890213	19970309	2	104	1	0	0	0	1	5
7	1001	19870211	19970309	2	108	11	0	7H+QgYGzTNOkwcZnh/f67/VhPWvMHtiIJI9SGMSjoNg=	0	3	6
8	1001	19930324	19970309	2	11	60	0	0	0	4	7
9	1001	19930720	19970309	2	102	3	0	0	0	14	8
10	1001	19930324	19970309	2	11	1	9001	0	0	130	0
11	1001	19890926	19970309	2	105	1	0	0	0	1	9
12	1001	19911021	20081105	2	109	12910	0	0	0	3	10
13	1001	19900118	19970309	2	94	1	0	0	0	1	11
14	1001	19920302	19970307	6	1	61	9001	WcoK5zdbjlooQ3ek79ptJfRxoMuZiASn+FzgwXG6Nq4=	0	2	0
15	1001	19900329	19970309	2	20	1	0	0	0	1	12
16	1001	19910522	19970307	6	2	61	9001	WcoK5zdbjlooQ3ek79ptJZYO3j4zFuBt/3hTIdeSolE=	0	2	0
17	1001	19900806	19970309	2	23	1	0	0	0	1	13
18	1001	19910228	19970309	2	106	1	0	0	0	1	14
19	1001	19920608	19970309	2	7	1	0	0	0	1	15
20	1001	19930324	19970309	2	11	1	0	0	0	1	16
21	1001	19930922	19970309	2	34	1	0	0	0	1	17
22	1001	19931020	19970309	2	35	1	0	0	0	1	18
23	1001	19931117	19970309	2	36	1	0	0	0	1	19
24	1001	19940303	19970309	2	38	1	0	0	0	1	20
25	1001	19960404	19970309	2	107	1	0	0	0	1	21
26	1001	19930519	19970309	5	111	13204	1001	0	0	70	22

rEHR

rEHR

- Tool which simplifies and accelerates the process of extracting ready-for-analysis datasets from EHRs databases
- To increase transparency and reproducibility of research
- Developed using primary care data from the UK (CPRD)
- Software R

rEHR

- It uses Structured Query Language (SQL) and SQLite to store data
  - SQLite files are stored efficiently and are relatively small compared to text files
  - SQL language has been optimised for very rapid and efficient queries of SQLite files
- It uses packages which optimise data manipulation (*dyplr*)
- Multicore processing (e.g. Linux)

rEHR functions

- Loading package and importing files into a SQL database
- Querying the database
- Building longitudinal data and calculation of prevalence and incidence
- Building a cohort dataset ready for survival analysis, matching
- Accessory functions

Loading rEHR and Importing files

Loading rEHR

- rEHR is available from CRAN* and Github
- From CRAN

```
if(!"rEHR"%in%rownames(installed.packages()))
install.packages("rEHR")
library(rEHR)
```

- Development version from Github

```
library(devtools)
install_github("rOpenHealth/rEHR")
library(rEHR)
```

* currently archived on CRAN

MANCHESTER 1824
The University of Manchester

rEHR: Importing files

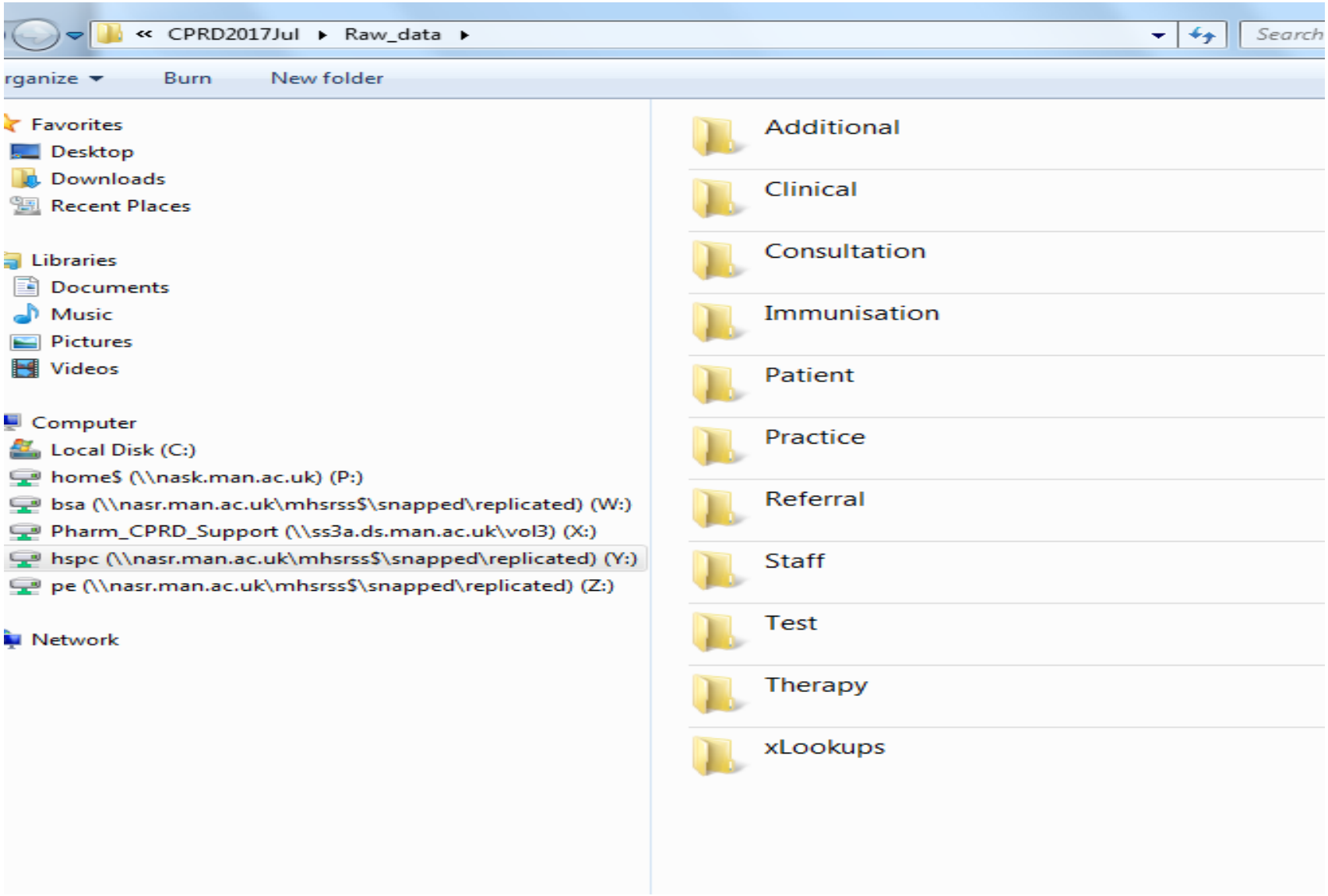
- Create database connection

```
dbname<-"Y:/rEHR_test//CPRDJul2017.sqlite"
db<-database(dbname)
```

- Import multiple files into the database

```
import_CPRD_data(db, data_dir = cprd_path,
                 filetypes = c("Patient","Clinical", "Referral",
                               "Test", "Practice","Immunisation",
                               "Consultation", "Therapy"),
                 dateformat = "%Y%m%d",
                 yob_origin = 1800,
                 regex = "p*",
                 recursive = TRUE)
```

rEHR: Importing files



rEHR: Importing files

- Create database connection

```
dbname<-"Y:/rEHR_test//CPRDJul2017.sqlite"
db<-database(dbname)
```

- Import multiple files into the database

```
import_CPRD_data(db, data_dir = cprd_path,
                 filetypes = c("Patient","Clinical", "Referral",
                               "Test", "Practice","Immunisation",
                               "Consultation", "Therapy"),
                 dateformat = "%Y%m%d",
                 yob_origin = 1800,
                 regex = "p*",
                 recursive = TRUE)
```

rEHR: Importing files

« CPRD2017Jul ▸ Raw_data ▸ Clinical ▸

Search Clinical

Organize ▾ Burn New folder

Favorites
- Desktop
- Downloads
- Recent Places

Libraries
- Documents
- Music
- Pictures
- Videos

Computer
- Local Disk (C:)
- home$ (\\nask.man.ac.uk) (P:)
- bsa (\\nasr.man.ac.uk\mhsrss$\snapped\replicated) (W:)
- Pharm_CPRD_Support (\\ss3a.ds.man.ac.uk\vol3) (X:)
- hspc (\\nasr.man.ac.uk\mhsrss$\snapped\replicated) (Y:)
- pe (\\nasr.man.ac.uk\mhsrss$\snapped\replicated) (Z:)

Network

p001_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 63.5 MB
p002_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 117 MB
p003_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 76.9 MB
p004_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 25.8 MB
p005_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 35.1 MB
p006_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 42.3 MB
p007_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 6.74 MB
p008_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 4.90 MB
p009_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 38.2 MB
p010_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 21.9 MB
p011_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 38.6 MB
p012_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 26.4 MB
p013_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 108 MB
p014_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 15.1 MB
p015_Clinicaltxt.zip	Type: zip Archive	Date modified: 19/0 Size: 28.8 MB
p016_Clinicaltxt.zip		Date modified: 19/0

718 items Offline status: Online
Offline availability: Not available

rEHR: Importing files

- View the database

```
head(db)
#> head(db)
#type           name      tbl_name
#1 table       Patient       Patient
#2 table      Clinical      Clinical
#3 table      Referral      Referral
#4 table          Test          Test
#5 table Immunisation Immunisation
#6 table Consultation Consultation
#7 table       Therapy       Therapy
#8 table      Practice      Practice
```

- View a table

```
head(db, table = "Patient")
#> head(db, table = "Patient")
#patid  vmid gender  yob mob marital famnum chsreg chsdate prescr capsup ses  frd  crd regstat reggap internal  tod
#1  1001  2457      2 1930   0       2   1063      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  2  2001  2456      1 1930   0       6   1063      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  3  3001 19417      1 1939   0       2   9788      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  4  4001 15663      1 1937   0       2   7827      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  5  5001 15662      2 1936   0       2   7827      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  6  6001 15118      2 1942   0       3   7563      2      NA      0      4   0 <NA> <NA>       0      0        0 <NA>
#  toreason deathdate accept practid
#1       27      <NA>      1       1
#2        3      <NA>      1       1
#3        2      <NA>      1       1
#4        0      <NA>      1       1
#5        0      <NA>      1       1
#6        2      <NA>      1       1
```

rEHR: Importing clinical codes lists

- Import a list of Read codes

```
diabetes_codes<-clinical_codes[clinical_codes$list == "Diabetes",]
```

```
diabetes_codes
# A tibble: 117 x 4
medcode readcode                                                    desc     list
<int>    <chr>                                                     <chr>    <chr>
  1      251  C10..00                                            Diabetes mellitus Diabetes
2      252  C109J00                     Insulin treated Type 2 diabetes mellitus Diabetes
3      253  C109K00   Hyperosmolar non-ketotic state in type 2 diabetes mellitus Diabetes
4      254  C10C.00                          Diabetes mellitus autosomal dominant Diabetes
5      255  C10D.00                   Diabetes mellitus autosomal dominant type 2 Diabetes
6      256  C10E.00                                       Type 1 diabetes mellitus Diabetes
7      257  C10E.11                                       Type I diabetes mellitus Diabetes
8      258  C10E.12                           Insulin dependent diabetes mellitus Diabetes
9      259  C10E000             Type 1 diabetes mellitus with renal complications Diabetes
10     260  C10E012 Insulin-dependent diabetes mellitus with renal complications Diabetes
# ... with 107 more rows
```

Querying the database

rEHR: Selecting events

- Identify all individuals with a specific diagnosis

```
diabetes_patients<-select_events(db, tab = "Clinical",
                columns = c("patid", "eventdate", "medcode"),
                where = "medcode %in% .(diabetes_codes$medcode) &
                eventdate < '2006-01-01' & eventdate >= '2005-01-01'")
```

- Alternatively, you can use SQL query

rEHR: Selecting first or last events

- Selecting first events

```
first_DM<-first_events(db, tab="Clinical",
                       columns=c("patid", "eventdate", "medcode"),
                  where="medcode %in% .(diabetes_codes$medcode)")
```

```
head(first_DM)
```

```
##   patid    eventdate   medcode
## 1  1004   2007-12-25       351
## 2  1005   2004-08-31       351
## 3  1008   2002-03-02       351
## 4  1010   2014-04-11       351
## 5  1012   2012-05-28       351
## 6  1015   2008-08-16       351
```

rEHR: Selecting first or last events

- Selecting last events

```
last_DM<-last_events(db, tab="Clinical",
                     columns=c("patid", "eventdate", "medcode"),
                 where="medcode %in% .(diabetes_codes$medcode)")

head(last_DM)
```

```
##   patid    eventdate  medcode
## 1  1004   2007-12-25      351
## 2  1005   2009-03-09      351
## 3  1008   2002-03-02      351
## 4  1010   2014-04-11      351
## 5  1012   2013-02-14      351
## 6  1015   2013-08-17      273
```

rEHR: Querying longitudinal data

- Extracting longitudinal data:
  - eg calculate the incidence/prevalence of a disease over time
- *select_by_year* function:
  - It can use parallel processing on multi-cores machine (e.g. Linux)

rEHR: Querying longitudinal data

- Selecting events by year

```
registered_patients <- select_by_year(db = db,
                                      tables = "Patient",
                                      columns = c("patid", "practid", "gender",
                                                  "yob", "crd", "tod", "deathdate"),
                                      where = "crd < STARTDATE",
                                      year_range = c(2008:2012),
                                      year_fn = standard_years)
```

```
> table(registered_patients$year)

2008 2009 2010 2011 2012
189  195  201  206  214
```

MANCHESTER 1824
The University of Manchester

Incidence and prevalence data

- Incidence data

```
incident_cases <- select_by_year(db = db,
                                 tables = c("Clinical", "Referral"),
                                 columns = c("patid", "eventdate", "medcode"),
                                 where = "medcode %in% .(diabetes_codes$medcode) &
                                              eventdate <= ENDDATE",
                                 year_range = c(2008:2012),
                                 year_fn = standard_years,
                                 selector_fn = first_events)
```

- Remove duplicates

```
> incident_cases %>%
      group_by(patid, year) %>%
      arrange(eventdate) %>%
      distinct() %>%
      ungroup -> incident_cases
```

Incidence and prevalence data

- Calculate prevalence and incidence data

```
prevalence_dat <- left_join(registered_patients, incident_cases)
prevalence_dat <- prev_terms(prevalence_dat)
totals <- prev_totals(prevalence_dat)
```

```
> totals$prevalence$year_counts
# A tibble: 5 x 4
   year numerator denominator prevalence
  <int>     <int>       <dbl>      <dbl>
1  2008        32    175.6715   18.21582
2  2009        37    181.3717   20.40010
3  2010        43    185.1335   23.22649
4  2011        53    188.4079   28.13045
5  2012        59    195.5811   30.16651
```

```
> totals$incidence$year_counts
# A tibble: 5 x 4
   year numerator denominator incidence
  <int>     <int>       <dbl>     <dbl>
1  2008         5    143.9014  3.474600
2  2009         4    144.4983  2.768199
3  2010         4    142.2806  2.811345
4  2011         8    135.5893  5.900170
5  2012         6    137.4675  4.364668
```

Building a cohort

Building a cohort

- Use *build_cohort*

```
cohort<-build_cohort(prevalence_dat, cohort_type="incid",
                     cohort_start = "2006-01-01", cohort_end = "2012-12-31"
                     diagnosis_start = "eventdate")
```

- Add a column for death

```
cohort$death <- with(cohort,
                     ifelse(!is.null(deathdate) &
                              (deathdate > as.Date("2006-01-01") &
                                 deathdate < as.Date("2012-12-31")),
                            1, 0))
cohort$death[is.na(cohort$death)] <- 0
```

- Run survival analysis

```
library(survival)
surv_obj <- with(cohort, Surv(start, end, death))
coxph(surv_obj ~ gender + case, data = cohort)
```

MANCHESTER 1824
The University of Manchester

Matching

Matching

- rEHR has three matching functions:
  - Incidence Density Matching
  - Exact matching
  - Matching on index date

Incidence Density Matching

- Build a longitudinal cohort

```
cohort2 <- build_cohort(prevalence_dat, cohort_type = "incid",
                        cohort_start = "2006-01-01", cohort_end = "2012-12-31",
                        diagnosis_start = "eventdate")
```

- Use get_matches for IDM

```
IDM_controls <- get_matches(cases = filter(cohort2, case == 1),
                            control_pool = filter(cohort2, case == 0),
                            match_vars = c("gender", "region"),
                            n_controls = 4, cores = 1,
                            method = "incidence_density", diagnosis_date  = "eventdate")
```

Exact Matching

- Selecting a comparison cohort without replacement

```
exact_controls3 <- get_matches(cases = filter(cohort2, case == 1),
                               control_pool = filter(cohort2, case == 0),
                               match_vars = c("gender", "region"),
                               n_controls = 4, cores = 2,
                               method = "exact", diagnosis_date  = "eventdate")
```

Matching on index date

- You can use the consultation files

```
consultation_dir <- "tempdir()"
flat_files(db, out_dir = consultation_dir, file_type = "csv")
```

- Use *match_on_index*

```
index_controls <- match_on_index(cases = filter(cohort2, case == 1),
                                 control_pool = filter(cohort2, case == 0),
                                 index_var = "eventdate",
                                 match_vars = c("gender", "region"),
                                 index_diff_limit = 90,
                                 consult_path = consultation_dir,
                                 n_controls = 4,
                                 import_fn = function(x) convert_dates(read.csv(x)))
```

Accessory functions

Accessory functions

- Time-varying covariates
  - *cut_tv*
- Unit conversion (HbA1C)
  - *cprd_uniform_hba1c_values()*
- Exporting data to Stata format
  - *to_stata()*
- Building clinical codes list

Building Clinical Codes List

- Based on a methodology previously described[1]
- Construct a clinical code list using *MedicalDefinition()*
  - *terms() :* clinical search terms
  - *codes():* clinical codes
  - *tests():* test search terms
  - *drugs():* drug search terms
  - *drugcodes():* drug product codes
- Run the search against look-up tables provided with EHRs using *build_definition_lists()*

[1]Olier I, Springate DA, Ashcroft DM, et al. (2016) Modelling Conditions and Health Care Processes. in Electronic Health Records: An Application to Severe Mental Illness with the Clinical Practice Research Datalink. PLOS ONE 11(2): e0146715

MANCHESTER 1824
The University of Manchester

Building Clinical Codes Lists

```
def <- MedicalDefinition(
  terms = list(
    "peripheral vascular disease", "peripheral gangrene", "-wrong answer",
    "intermittent claudication", "thromboangiitis obliterans",
    "thromboangiitis obliterans", "diabetic peripheral angiopathy",
    c("diabetes", "peripheral angiopathy"),  # single AND expression
    c("buerger",  "disease presenile_gangrene"),
    "-excepted", # exclusion
    codes = list("G73"),
    tests = NULL,
    drugs = list("insulin", "diabet", "aspirin")))
```

```
medical_table <- read.delim("Lookups/medical.txt", fileEncoding = "latin1", stringsAsFactors = FALSE)
drug_table <- read.delim("Lookups/product.txt", fileEncoding = "latin1", stringsAsFactors = FALSE)
```

```
draft_lists <- build_definition_lists(def, medical_table = medical_table, drug_table = drug_table)
```

Clinicalcode.org

- Repository holding lists of clinical/drugs codes used in EHRs databases
- It aims to improve transparency and reproducibility of research by sharing codes/drugs lists used in published studies
- It currently contains 84,346 clinical codes deposited over 499 code lists

ClinicalCodes.org

An online clinical codes repository to improve validity and reproducibility of medical database research

All publications with clinical code lists:

Type	Title	Journal	Year	Authors
Research article	Landmark Models for Optimizing the Use of Repeated Measurements of Risk Factors in Electronic Health Records to Predict Future Disease Risk	American Journal of Epidemiology	2018	Ellie Paige, Jessica Barrett, David Stevens, Ruth H. Keogh, Michael J. Sweeting, Irwin Nazareth, Irene Petersen, and Angela M. Wood
Research article	Computing Care Quality Improvement Tactics from Health Records: Closing the Gap Between Audit and Action	AMIA Symposium	2014	Benjamin Brown, Richard Williams, Matthew Sperrin, Timothy Frank, John Ainsworth, Iain Buchan
Research article	Antibacterial Drugs and the Risk of Community-Associated Methicillin-Resistant Staphylococcus aureus in Children	Archives of pediatrics and adolescent medicine	2011	Verena Schneider-Lindner, Caroline Quach, James A. Hanley, Samy Suissa
Research article	Smoking-related mortality in patients with early rheumatoid arthritis – a retrospective cohort study using the Clinical Practice Research Datalink	Arthritis Care and Research	2016	Rebecca M Joseph, Mohammad Movahedi, William G Dixon, Deborah PM Symmons
Research article	Suicide risk in primary care patients diagnosed with personality disorder: a nested case control study	BMC Family Practice	2015	Michael Doyle; David While; Pearl L.H. Mok; Kirsten Windfuhr; Darren M. Ashcroft; Evangelos Kontopantelis; Carolyn Chew-Graham; Louis Appleby; Jenny Shaw; Roger T. Webb
Research article	Adaptation and validation of the Charlson Index for Read/OXMIS coded databases	BMC Family Practice	2010	Nada F Khan, Rafael Perera, Stephen Harper, Peter W Rose

Quick links:

Contact

Upload

Login

Twitter

MANCHESTER 1824

The University of Manchester

ClinicalCodes is a project at the University of Manchester Institute of Population Health

NHS

National Institute for Health Research

The ClinicalCodes project is funded by the National Institute for Health Research (NIHR) School for Primary Care Research (SPCR)

Summary

- Working with EHRs data requires computational and statistical expertise
- rEHR package greatly simplifies and accelerates the extraction and processing of coded data from EHR databases
- rEHR is many times faster than equivalent code:
  - SQL
  - Packages which can optimise data manipulation (*dyplr*)
  - Multicore functionality

Limitations

- Currently tested only with CPRD data
  - However application to other EHR databases is possible
- Not always flexible
  - The user needs to follow the steps described in the package
- It heavily depends on other packages such as *dyplr*, therefore the package needs to be updated often

Future work

- Future version of rEHR:
- Implementation of "resample"[2] algorithm for representative sampling of practices
- Algorithm for determining smoking status (as tvc)
- Interfaces to other EHR systems in particular THIN, QResearch and Research One
- Uniform units functions for other clinical measurements such as blood pressure, cholesterol and serum creatinine

[2]Kontopantelis (2013). A Greedy Algorithm for Representative Sampling: resample in Stata. *Journal of Statistical Software*

MANCHESTER 1824
The University of Manchester

rEHR

PLOS | ONE

RESEARCH ARTICLE

rEHR: An R package for manipulating and analysing Electronic Health Record data

David A. Springate[1,2], Rosa Parisi[3], Ivan Olier[4], David Reeves[1,2], Evangelos Kontopantelis[1,5¤]*

1 NIHR School for Primary Care Research, University of Manchester, Manchester, United Kingdom, **2** Centre for Biostatistics, Faculty of Biology, Medicine & Health, University of Manchester, Manchester, United Kingdom, **3** Centre for Pharmacoepidemiology & Drug Safety, Faculty of Biology, Medicine & Health, University of Manchester, Manchester, United Kingdom, **4** Informatics Research Centre, School of Computing Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, United Kingdom, **5** The Farr Institute for Health Informatics Research, Faculty of Biology, Medicine & Health, University of Manchester, Manchester, United Kingdom

¤ Current address: Vaughan House, Portsmouth Street, M13 9GB, Manchester, United Kingdom
* e.kontopantelis@manchester.ac.uk

THANK YOU!

Email: rosa.parisi@manchester.ac.uk

MANCHESTER
1824
The University of Manchester