(VERSION 1 OF STUDY PROTOCOL: NOTE CONSIDERABLE CHANGES WERE MADE AS WTHE PANDEMIC EVOLVED)

(CASE CONTROL ANALYSES ORIGINALLY PLANNED ARE OMITTED FROM THE FINAL MANUSCRIPT)

COVID-19 Disease in people with Diabetes in Scotland: Incidence, Severity and Risk Stratification

Helen Colhoun

This version compiled 15 April 2020

Contents

Background

There are anecdotal reports of excesses of people with diabetes, hypertension and cardiovascular disease among hospitalised cases with COVID-19 disease. However formal evaluations of the risk associated with diabetes and whether this risk is attributable to the attendant hypertension and CVD or not is unclear. As a result of possible excess risk it has been proposed that people with diabetes should be included among the most vulnerable for whom stringent distancing measures termed shielding or cocooning are recommended. The added complexity of glucose management during concurrent infection and the impact of diabetes associated reductions in renal function may mean that the proportions of those with disease warranting hospitalisation who are entering critical care and dying may be higher with diabetes.

Here we harnessed the potential of electronic health care linkage in Scotland to ascertain testing, hospitalisation, critical care management and death data for the entire population with diabetes in the national diabetes register. This paper provides estimates of disease risks and the course of disease in those with versus without diabetes. An accompanying paper describes the development and performance of a formal risk prediction stratifier for severe COVID-19 disease among those with diabetes.

As of testing for COVID-19 has been entirely based on nucleic acid tests. At the earlier phase of the epidemic community based testing in those returning from high risk countries and contact tracing of the first 100 cases was carried out but since testing strategy essentially triaged out those with milder disease, limiting most testing to those with ” suspected COVID-19 disease warranting hospital admission because of respiratory distress” . In the past week testing of health care workers and other key workers has accelerated.

Methods

Data sources

The SCI-Diabetes register in Scotland has been described in detail previously. In brief all persons who receive a diagnostic code for diabetes in primary or secondary care are automatically incepted into the national database for diabetes. This captures core data from primary care including issued prescriptions and data from all secondary care encounters including all laboratory measures. As there is a unique health care identifier on all National Health Service records (the community health care index number- CHI) these data are linkable to other datasets. Here we linked to data from the Scottish Intensive Care Society and Audit Group critical care register (SICSAG), RAPID24 which is a daily updated hospital admissions register, Scottish Morbidity Record 01 that is a weekly update of admission with more complete disease coding, The National Register of
Scotland (NRS) death data and the Electronic Communication of Surveillance in Scotland (ECOSS) dataset that captures all virology nucleic acid (Rt-PCR) testing for SARS-CoV-2 nationally. General population controls were obtained by incidence sampling of ten persons without without a test positive age, sex and general practice matched to test positives using the CHI database.

Statistical methods Case definitions

We use the following case groupings here:

- all Test confirmed SARS-CoV-2 infection
- severe cases defined as CCU or death within 28 days of positive test
- hospitalised case -any hospitalisation within 28 days of a positive test
- non-hospitalised test positives- no hospitalisation within 28 days of test
- Total COVID-19 deaths ie any deaths in those with a test positive death from any cause by 28 days post test AND any NRS deaths with a code for COVID19 emergency code U07.1All those who have died with covid 19 code regardless of test status

Comparisons with general population Initial tabulations

Take the matched controls and all *Test confirmed SARS-CoV-2 infection* positives for COVID from ECOSS. Tabulate the frequency (% or medians and IQRs as apppropriate) for the following groups:

- controls,
- all Test confirmed SARS-CoV-2 infection
- hospitalised case -any hospitalisation within 28 days of a positive test
- severe cases defined as CCU or death within 28 days of positive test The following vars Age sex ethnicity when available

SIMD

ANY diabetes, type 1 diabetes, type 2 diabetes (when

type data are available)

Prior CVD (five year SMR01 lookback)

Treated hypertension

On ACE i in past three months

On ARB in past three months

Prior admission for pneumonia in SMR01 in past five years

Logistic models to define OR associated with diabetes

Then Construct a logistic regression model of EACH of the following outcomes using cases and controls with the controls restricted to the relevant group of cases

- having a Test confirmed SARS-CoV-2 infection or not
- being a hospitalised case -any hospitalisation within 28 days of a positive test or not
- being a severe case defined as CCU or death within 28 days of positive test or not

Adjusted for age, sex and diabetes

Then repeat but Adjusted for age, sex and diabetes AND all the other covariates above

Look for INTERACTIONS between DM and all other covariates & re run the regression retaining any significant interactions in the model

When diabetes type becomes available repeat all the logisitic models replacing any diabetes with type 1 and then type 2 respectively

Comparisons of outcomes among cases with and without diabetes

Initial tabulations

Include all those with test positive for COVID 19 Group

into those with and without diabetes Group into age

bands (eg <50, 50-70, 70+)

Tabulate BY Diabetes yes / no status (medians and IQR or % as appropriate)

- n of Deaths with a COVID code without a prior test positive
- n of test positives
- among test positives
  - age band and sex
  - ECOSS vars : HEALTAUT, SPECORIG,PATCAT ,HOSPITALISED and DATEREC (need to group DATEREC by week)
  - n and & hospitalised
  - duration between hospitalisation and discharge in survivors
  - n and % dying within 28 days
  - n and % meeting severe definition ie CCU or death within 28 days
  - n and % dying ever died or not by checking against NRS deaths (since some may not meet 28 day definition but may have died since)
  - n and % entering CCU (ie an entry in SICSAG)
  - n and % of those in CCU ever on ventilator in SICSAG using “Patient on ventilator” variable
  - median iqr days to ventilator from admission among those ventilated
  - % with acute renal failure in CCU (from SICSAG)
  - n% of those entering CCU who have died

Repeat above tabulation for age band stratum

Repeat all the tabulations and regressions separately for type 1 diabetes and type 2 diabetes when available

Case Logistic regression of outcomes among cases by DM status

Construct conditional logistic model of

Being hospitalised given test positive status by age sex diabetes

Being a severe case given test positive status by age sex diabetes

Linear Regression model of time to ventilation by age sex diabetes

Case fatality rate for COVID1- in diabetes

Inclusion criteria: Those with test confirmed SARS-CoV-2 infection as defined above

Total mortality is death regardless of underlying cause

Date of death is ascertained from NRS deaths

Those without a death notification are censored at the date of last availability of NRS deaths data Person time is contributed from the i) date of first positive test to date of death regardless of time to death or underlying cause being assigned as due to COVID-19 or ii) censoring date whichever applies Main Outputs required :

- Generate survival curves overall and by diabetes
- Report log rank test for difference in survival by diabetes status
- Tabulate the counts of deaths and the median survival times by diabetes
- Fit a Cox regression or Poisson model to yield the HR and CIs for diabetes adjusted for age Sex, Ethnicity, SIMD, Prior CVD, Prior Antihypertensive Drugs, Prior pneumonia, residence in care home if available from ECOSS,
- (repeat by type of diabetes when type available)
- For discussion. How to take into consideration different testing policy time periods.

Associations of risk factors with disease and its outcomes and risk prediction model AMONG those with diabetes

Inclusion criteria / population under study

All those alive with type 1 or type 2 diabetes in SCI-diabetes as of date of 1st test positive case (regardless of diabetes in Scotland) less 14 days=

Date in: as above

Date out: data of death or last date of extraction of SCI-DM

Initial tabulations

Tabulate by type of diabetes and overall and by broad age band (

medians and IQR or % as appropriate)

- n % with Deaths with a COVID code without a prior test positive
- n % of test positives
- among test positives – sex
  - ECOSS vars : HEALTAUT, SPECORIG,PATCAT ,HOSPITALISED and DATEREC (need to group DATEREC by week)
  - n and & % hospitalised
  - duration between hospitalisation and discharge in survivors
  - n and % dying within 28 days
  - n and % meeting severe definition ie CCU or death within 28 days

- n and % dying ever died or not by checking against NRS deaths (since some may not meet 28 day definition but may have died since)
- n and % entering CCU (ie an entry in SICSAG)
- n and % of those in CCU ever on ventilator in SICSAG using “Patient on ventilator” variable
- median iqr days to ventilator from admission among those ventilated
- % with acute renal failure in CCU (from SICSAG)
- n% of those entering CCU who have died Repeat above tabulation for age band stratum

Tabulate DM risk factors in those with various COVID status

For type 2 DM and type 1 DM separately and all combined

Columns

No Covid, test positive, hospitalised with covid, severe covid, any death from covid, Any covid (ie test positive or death with COVID code) Summarise

Age

Sex,

Ethnicity

SIMD

Duration of DM

Hba1c

On pump

On Libre

Prior DKA admission in past five years

Prior hypo admission in past five years

Number of hypo admissions in past five years

Number of DKA admissions in past five years Number of DM

drugs used

Then binary for each class of DM drugs

Prior admission with respiratory disease in past five years

Prior CVD

Prior stroke

Prior ischaemic heart disease

Prior heart failure

Prior Hypertension

Prior any retinopathy

RRT

Current or ever Smoking

SBP DBP eGFR

Albuminuria status

BMI

On CVD drugs

Other specific drug classes as per REACT 1 study - also acyclovir

For any DM patients under 50 in CCU we will describe the above covariates by type of diabetes

We will also list for all death by type of diabetes the underlying cause of death on death certs by dm and what are the two most common additional causes on the death cert

We will also list for all admissions by type of diabetes the three most common causes for admission cited in SMR01

Simple Logistic model of severe disease

run logistic for severe disease yes/no adjusted for age sex dm duration for each type of dm then all combined then run "univariate" logistic model for each of the other covariates above adjusted for age sex dm duration for each type of dm then all combined

Constructing a predictive model for disease severity and for death from COVID in those with DM

We will build a cross-validated model predicting severe COVID-19 initially combining type 1 and type 2 diabetes.

Person time / date in / date out will be contributed as above Person time will be broken into two week intervals

These intervals will be time updated with whether they are pre or post March23rd (introduction of lockdown).

We will use

Poisson regression with K fold cross validation (suggest 10 fold)

We will consider the above covariates in the model and type of diabetes- variable selection will be by forward selection with selection termination when adding a new variable does not reduce the deviance (evaluated on the fit of the model to the training fold) by more than 4 natural log units.

The prediction performance will be quantified as the AUC and as the weight of evidence (Lambda) ==