The use of primary care big data for COVID-19 research: A consensus statement from the COVID-19 Primary Care Database Consortium

Hajira Dambha-Miller[1], Simon J Griffin[2], Duncan Young[3], Peter Watkinson[4] Pui San Tan[5], Ashley K Clift[6], Rupert A Payne[7], Carol Coupland[8], Jemma C Hopewell[9] Jonathan Mant[10], Richard M Martin[11] and Julia Hippisley-Cox[12]

1. NIHR Academic Clinical Lecturer and GP, Division of Primary Care and Population Health, University of Southampton

2. Professor of General Practice and GP, Department of Public Health and Primary Care, University of Cambridge

3. Professor of Intensive Care Medicine, Nuffield Department of Clinical Neurosciences, University of Oxford

4. Associate Professor of Intensive Care Medicine, Consultant in Intensive Care and Acute Medicine, Nuffield Department of Clinical Neurosciences, University of Oxford

5. Data Scientist, Nuffield Department of Primary Care Health Sciences, University of Oxford

6.Intensive Care Doctor, NHS Nightingale Hospital London and Honorary Clinical Research Fellow, Department of Surgery and Cancer, Imperial College London

7. Consultant Senior Lecturer in Primary Health Care and GP, Population Health Sciences, University of Bristol

8. Professor of Medical Statistics in Primary Care, Division of Primary Care, School of Medicine, University of Nottingham

9. Associate Professor of Genetic Epidemiology and Clinical Trials, CTSU, Nuffield Department of Population Health, University of Oxford

10. Professor of Primary Care Research, Department of Public Health and Primary Care, University of Cambridge

11. Professor of Clinical Epidemiology, Population Health Sciences and National Institute for Health Research Bristol Biomedical Research Centre. , University of Bristol.

12. Professor of Clinical Epidemiology and General Practice, Director of QResearch database and GP, Nuffield Department of Primary Care Health Sciences, University of Oxford

Correspondence: Dr Hajira Dambha-Miller, Primary Care and Population Health, University of Southampton, Southampton, SO16 5ST, Email: H.Dambha-Miller@soton.ac.uk

Abstract

The use of big data containing millions of primary care medical records provides an opportunity for rapid research to help inform patient care and policy decisions during the first and subsequent waves of the COVID-19 pandemic. Routinely collected UK primary care data have previously been used for national pandemic surveillance, quantifying associations between exposures and outcomes, identifying high-risk populations and examining the effects of interventions at scale. However, there is no consensus on how to effectively conduct or report these data for COVID-19 research. A COVID-19 primary care database consortium was established in April 2020. Collectively, its researchers have ongoing COVID-19 projects in overlapping datasets with millions of primary care records representing 30% of the UK population, that are variously linked to public health, secondary care and vital status records. This consensus agreement is aimed at facilitating transparency and rigour in methodological approaches,

as well as consistency in defining and reporting cases, exposures, confounders, stratification variables and outcomes in relation to the pharmacoepidemiology of COVID-19. This will facilitate comparison, validation and pooling of research during and after the pandemic.

Aim and scope

Primary care 'big data' refers to routinely collected anonymised GP records that form large and complex longitudinal data, often with hundreds of variables at an individual level. These can often be linked to secondary care records, registry data (e.g. cancer) or the Office for National Statistics which records births and deaths.[1–3] A summary of the databases being utilised by our consortium is shown in table 1 below. It is likely that additional data sources will be forthcoming and we would welcome these notifications. The potential use of primary care and linked data for understanding COVID-19 infection is vast, and includes descriptive epidemiology; testing associations with prescribed drugs, including drugs that influence risk; clinical prediction tools for COVID-19 risk and outcome; the impact on, and effects of' health inequalities; or examining the indirect immediate and long-term effects of the infection, such as delayed clinical diagnoses, domestic abuse or mental health sequelae. Our present focus is on the pharmacoepidemiology of COVID-19, i.e. the potential influence of old and new drug therapies on COVID-19 outcomes. With an increasing number of studies using primary care big data to examine the influence of drugs on COVID-19 outcomes, it is timely to consider approaches to the conduct of these studies. This will facilitate study consistency and rigour, improve transparency and reduce ambiguity in both methods and reporting. As the pandemic progresses alongside the urgency to find solutions, rapid and rigorous research needs to be conducted with emergent findings externally validated. Consensus on definitions of COVID-19, exposures, outcomes, and consistency in considering potential confounders and stratification variables, will enable meaningful comparisons between findings and facilitate the potential for pooling results in meta-analyses. Our collective efforts and agreement on transparency in reporting and methodological approaches may contribute towards improved clinical decisions and in turn, improved population health.

Table 1: Summary of database characteristics (for more information see relevant websites)

	QResearch	RCGP Research & Surveillance Network Centre	Clinical Practice Research Datalink	UK Biobank
Established	2003	1957	1989	2006
Number of GP Practices	1500 (increasing to 2519 from April 2020)	700	1841	Partial cohort coverage
Number of current patient records as of 01.01.2020	10.6 million (rising to 21 million from April 2020)	5 million	14 million	0.5 million
Coverage	England	England	All of UK	England, Wales and Scotland
Age groups	All	All	All	40-69 years at recruitment
Clinical system	EMIS	EMIS Web, INPS Vision, TPP System One	EMIS Web, INPS Vision	Bespoke system
Birth registration	Yes	Yes	Yes	No
Death registration	Yes	Yes	Yes	Yes
Sociodemographic data	Yes	Yes	Yes	Yes

Ethnicity	Yes	Yes	Yes	Yes
Genome-wide genotyping data	No	No	No	Yes
Geographical location	Yes	Yes	Yes	Yes
Laboratory tests including COVID-19 results	Yes	Yes	Yes	Yes
Anthropometric data	Yes	Yes	Yes	Yes
Clinical signs and symptoms	Yes	Yes	Yes	Yes
Drugs prescribed	Yes	Yes	Yes	Yes
Radiology reports	Yes	Yes	Yes	No
Hospital referral	Yes	Yes	Yes	Yes
Hospital diagnosis	Yes	Yes	Yes	Yes
GP attendances	Yes	Yes	Yes	Partial
Hospital attendance	Yes	Yes	Yes	Yes
Additional key linkages to other datasets (full lists available from each database on request)				
Hospital Episode Statistics (HES)	Yes	Yes	Yes	No
HES outpatient data	Yes	Yes	Yes	No
HES accident and emergency data	Yes	Yes	Yes	No
HES diagnostic imaging dataset	Yes	Yes	Yes	No
Death registration data from the Office for National Statistics	Yes	Yes	Yes	Yes
Intensive Care dataset: ICNARC Case Mix Programme	Yes	No	Pending	No
Further information, including data access	https://www.qresearch.org/	https://www.qresearch.org/	https://www.cprd.com/	https://www.ukbiobank.ac.uk/about-biobank-uk/

Development of the consensus statement

This statement was developed by our primary care database consortium which includes UK experts in big data, epidemiologists, researchers in intensive care, researchers in primary care, statisticians, patient and public representatives, editors for journals in the field and front-line clinical staff. Universities represented are Bristol, Cambridge, Manchester, Nottingham, Oxford and Southampton. After initial discussions on the need for such a statement in our respective projects, we met weekly to refine ideas and reach agreement on item inclusion. The recommendations are entirely those of the consortium with sponsors and funders having no role in their development or reporting.

Agreed items for inclusion

1. Protocols

Wherever possible, protocols and analysis plans for COVID-19 research using primary care big data will be made widely available as quickly as possible through open-access journals or publicly available institutional repositories ahead of commencing the data analysis. This will facilitate transparency and subsequent scrutiny of findings, and *a priori* analysis plans should reduce false positive findings. Given the urgency of the research during the current pandemic, this will encourage greater efficiency and less duplication of efforts, unless conducted for validation purposes.

2. Defining COVID-19 infection

We considered the World Health Organization and Public Health England definitions of COVID-19 infection (as of 13th March 2020). [4,5] These definitions have several limitations in terms of operationalising them within database of electronic medical records. For example, clinical symptoms such as nasal discharge or sneezing may not necessarily be coded within GP records. Virology and serology tests that are being used to define cases are still being validated with varying reports on sensitivity and specificity.[6] UK national testing has been sparse, and has varied between the hospital and community setting at different stages of the pandemic. Definitions are likely to change and develop over time as more standardisation of GP coding is introduced, testing rates are widened from largely secondary to community settings and more accurate tests of infection and immunity are developed. Acknowledging these limitations, we have agreed on the following definitions at this time;

COVID-19 confirmed case	Positive result on RT-PCR assay of nasal or pharyngeal swab specimen ever OR Serology positive ever
COVID-19 suspected case	From 20th January 2020 (date that first confirmed case outside mainland China was reported to the World Health Organization) AND 1.Requiring admission to hospital AND one or more of the following: Clinical evidence of pneumonia OR Radiological evidence of pneumonia OR Acute respiratory distress syndrome OR Influenza like illness (fever ≥37.8°C and at least one of the following respiratory symptoms, which must be of acute onset: persistent cough (with or without sputum), hoarseness, nasal discharge or congestion, shortness of breath, sore throat, wheezing, sneezing

	OR 2.Not admitted to hospital AND one or more of the following: New continuous cough AND/OR High temperature
COVID-19 suspected with a negative test	Negative result (never positive) on RT-PCR assay of nasal or pharyngeal swab specimen OR Negative (never positive) serology

3. Drug exposures

Where prescribed drugs are being examined in COVID-19 infection, the following reporting principles should be applied:

Report generic drug name
List distinct classes of drugs (e.g. Angiotensin Receptor Blockers and Angiotensin-Converting Enzyme-Inhibitors as separate drug classes) AND Provide individual drug names
Where combination preparations have been prescribed, consider the component ingredients as separate for the purposes of the analysis
Provide clear definitions of drug exposure including: i. Exposure time – describe relevant dates of prescription for drug being investigated in relation to COVID-19 case-definition date; i.e. time duration before/during/after infection. For research questions with specific aim of altering outcomes, drug exposure during/after infection will be most informative ii. Repeat prescriptions for long term medication – list number of prescriptions within a defined time period iii. Dosage – describe how different drug dosage regimens are being treated in analysis

A list of UK prescribed drugs that require urgent characterisation for their potential in treating or altering outcomes in COVID-19 is given in the supplementary material 1. This list was compiled from the limited existing literature on this subject, our ongoing systematic review and anecdotal evidence from front line clinical staff who are treating patients with COVID-19 infections. [7–11] It is not an exhaustive list and will be updated as more data become available. Drug names within classes have been extracted from the British National Formulary.

4. Confounding variables

A list of variables that we recommend reporting and considering for inclusion as confounders within statistical models can be found in the supplementary material 2. We have not provided restrictive recommendations on how these variables should be categorised as this is dependent on the data available and the individual discretion of the researcher. We suggest consideration of these variables when determining confounders alongside a clear description of how the categorisation of the variable was determined and the rationale for its inclusion.

5. Stratification variables

In response to emergent findings on particular subgroups within the population being disproportionally affected by COVID-19 infection, [12] we propose examining outcomes stratified by age, sex, ethnic group and domicile (own home vs care/nursing home). As more data are published, further stratification variables should be considered.

6. Outcome reporting

We endorse the use of an appropriate reporting guideline such as the STROBE, RECORD or TRIPOD checklist. [13–15] We have considered outcomes from both ICU and primary care data as some linked datasets are being used within our consortium. Wherever possible, the following data should be reported:

Primary care outcomes	Short term	Lower Respiratory Tract Infection (Pneumonia)
		Emergency admission
		ICU admission
	Long term	All-cause mortality and cause-specific mortality including COVID-related mortality
ICU outcomes	Short term	Vital status at ICU discharge (alive/dead)
		Vital status at acute hospital discharge (alive/dead)
		Days of advanced respiratory support (artificial ventilation)
		Days of advanced cardiovascular support (inotropes, pressors or mechanical cardiovascular support)
		Days of renal support (use of renal replacement therapy)
		Days of ICU care (reported as ICU admission to discharge)
		Days of acute hospital care after ICU discharge (for repeat ICU admissions in the same acute hospital admission the total days not on ICU should be used)
	Long term	Vital status (alive/dead) at 30 and 90 days after ICU admission
		All-cause and COVID-related specific mortality at 6 and 12 months after ICU admission

7. Analytical methods

Individual study analyses are likely to vary from project to project. We agree on the ensuring analytical methods are transparent and reported in full, with the following guiding principles;

State an *a priori* hypothesis
Report descriptive characteristics including age, sex, ethnic group, measures of deprivation (such as the IMD or Townsend deprivation score), comorbidities and medication use
State sample size considerations, power calculations , and multiple testing considerations

Consider clustering by ICU or general practices or physician and employ appropriate methods e.g. robust standard errors
Check assumptions for any models (e.g. proportional hazards assumption)
Report how missing data were managed (e.g. multiple imputation method to replace missing data)
Report both unadjusted and adjusted models
Report any methods used to examine subgroups and interactions
Report on causal analysis methods, e.g. instrumental variables analysis [16]
Report any sensitivity analyses
Report on steps taken to mitigate time-window bias (in case-control study designs) or immortal time bias (in cohort study designs) [17]

Key challenges and shortcomings in the use of primary care big data for COVID-19 research

Across the respective primary care datasets, we acknowledge the potential limitations of using big data for COVID-19 research. All data are collected from routine clinical care. They are dependent on accurate coding by individual clinicians, and this does not guarantee consistency or accuracy of codes. Data on exposures and confounders will have been entered before the pandemic, and there might be a delay in outcome data reaching GP records. Uptake of newly introduced codes that are specific to COVID-19 may not be universal. Historically, however, UK primary care records have always been of a high level in terms of accuracy, and completeness of clinical diagnosis and medication prescribing. [18,19] The use of non-randomised observational data to make causal inferences still requires careful interpretation and appropriate analyses.[20] Other considerations relate to the case definition of COVID-19. Our definitions have been informed by those proposed by Public Health England and the WHO. We will utilise positive RT-PCR or serology testing as a definitive way of confirming cases. UK testing for COVID-19 has been limited and to date, we are not aware of any established serology or virology test with high sensitivity or specificity. Moreover, recent modelling suggests that there might be a substantial proportion of asymptomatic COVID-19 cases. [21] These individuals will not have presented to the health services and will not be identified within our datasets.

It is plausible that big data will over-represents disease severity and the factors contributing towards it because these less severe and asymptomatic cases are not recorded. This will be less relevant in subgroup designs or analyses assessing the risk of adverse outcomes in those presenting to hospital or ICU. All observational studies nested within these databases will be subject to the usual risk of statistical error (type 1 or type 2), bias and confounding. These must be considered in terms of size and direction. It is likely that many of the biases will be non-differential and minimised to some extent by the large sample sizes afforded by the data. Moreover, these primary care data lack selection and recall biases and often include multiple linkages to enable best-attainable ascertainment of outcome and exposure data. Large sample sizes will increase precision but could also lead to false positives. In the early stages, the number of people with outcomes recorded in GP records will be small but this is rising rapidly and the timing of analyses will therefore be important. If conducted too early the sample size will be inadequate but if too late, opportunities for findings to influence policy will be missed.

Competing Interests

All authors have completed the Competing Interest form.

PST has consulted for AstraZeneca Limited. JH reports that the CTSU receives research grants from the pharmaceutical industry that are governed by University of Oxford contracts that protect its independence, and has a staff policy of not taking personal payments from industry. From 01.02.2019 JHC is Professor of Clinical Epidemiology and General Practice at the University of Oxford. She is founder and director of QResearch database which is not-for-profit organisation with EMIS (leading commercial supplier of IT for 55% of general practices in the UK). JHC is co-owner of ClinRisk Ltd and was a paid director until June 2019. ClinRisk Ltd develops open and closed source software to ensure the reliable and updatable implementation of clinical risk equations within clinical computer systems to help improve patient care. No other authors have any competing interests to declare. The authors declare that no support from any organisation and no financial relationships have influenced the submitted work.

Funding

RMM is supported in part by the National Institute for Health Research Bristol Biomedical Research Centre and by a Cancer Research UK (C18281/A19169) programme grant (the Integrative Cancer Epidemiology Programme). PW is supported in part by the National Institute for Health Research Oxford Biomedical Research Centre. JCH acknowledges personal support from the British Heart Foundation (FS/14/55/30806) and Cancer Research UK (C5255/A18085) through the Cancer Research UK Oxford Centre. JHC also receives support from the NHS and National Institute for Health Research. SJG is supported by an MRC Epidemiology Unit programme: MC_UU_12015/4. The University of Cambridge has received salary support in respect of SJG from the NHS in the East of England through the Clinical Academic Reserve. HDM is an National Institute for Health Research funded Academic Clinical Lecturer. JM is an National Institute for Health Research Senior Investigator. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the UK National Institute for Health Research (NIHR) or the Department of Health and Social Care.

Contributors

HDM drafted the manuscript and all authors contributed towards revising it.

Transparency declaration

This manuscript is an honest, accurate, and transparent account; no important aspects have been omitted.

Ethical approval

Not applicable

Data sharing

No other data is available for this consensus statement

References

1 Williams T, Van Staa T, Puri S, *et al.* Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. Therapeutic Advances in Drug Safety. 2012;**3**:89–99. doi:10.1177/2042098611435911

2 Herrett E, Gallagher AM, Bhaskaran K, *et al.* Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology* 2015;**44**:827–36. doi:10.1093/ije/dyv098

3 De Lusignan S, Correa A, Smith GE, *et al.* RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. British Journal of General Practice. 2017;**67**:440–1. doi:10.3399/bjgp17X692645

4 Global Surveillance for human infection with coronavirus disease (COVID-19). https://www.who.int/publications-detail/global-surveillance-for-human-infection-with-novel-coronavirus-(2019-ncov) (accessed 13 Apr 2020).

5 COVID-19: investigation and initial clinical management of possible cases - GOV.UK. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases (accessed 13 Apr 2020).

6 Hoffman T, Nissen K, Krambrich J, *et al.* Evaluation of a COVID-19 IgM and IgG rapid test; an efficient tool for assessment of past exposure to SARS-CoV-2. *Infection Ecology & Epidemiology* 2020;**10**:1754538. doi:10.1080/20008686.2020.1754538

7 Patel AB, Verma A. COVID-19 and Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers. *JAMA* Published Online First: 24 March 2020. doi:10.1001/jama.2020.4812

8 Gbinigie K, Frie K. Should chloroquine and hydroxychloroquine be used to treat COVID-19? A rapid review. *BJGP open* Published Online First: 7 April 2020. doi:10.3399/bjgpopen20X101069

9 Little P. Non-steroidal anti-inflammatory drugs and covid-19. *BMJ* 2020;**368**:m1185. doi:10.1136/bmj.m1185

10 Cortegiani A, Ingoglia G, Ippolito M, *et al.* A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19. *Journal of Critical Care* Published Online First: 10 March 2020. doi:10.1016/j.jcrc.2020.03.005

11 Dong L, Hu S, Gao J. Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discoveries & Therapeutics* 2020;**14**:58–60. doi:10.5582/ddt.2020.01012

12 Rimmer A. Covid-19: Disproportionate impact on ethnic minority healthcare workers will be explored by government. *BMJ (Clinical research ed)* 2020;**369**:m1562. doi:10.1136/bmj.m1562

13 Von Elm E, Altman DG, Egger M, *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. Annals of Internal Medicine. 2007;**147**:573–7. doi:10.7326/0003-4819-147-8-200710160-00010

14 Collins GS, Reitsma JB, Altman DG, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine* 2015;**162**:55–63. doi:10.7326/M14-0697

15 Benchimol EI, Smeeth L, Guttmann A, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015;**12**. doi:10.1371/journal.pmed.1001885

16 Rassen JA, Brookhart MA, Glynn RJ, *et al.* Instrumental variables II: instrumental variable application-in 25 variations, the physician prescribing preference generally was strong and reduced covariate imbalance. *Journal of Clinical Epidemiology* 2009;**62**:1233–41. doi:10.1016/j.jclinepi.2008.12.006

17 Immortal time bias in observational studies of drug effects. ;**16**:241–9.

18 Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. doi:10.1136/bmj.302.6779.766

19 Majeed A. Sources, uses, strengths and limitations of data collected in primary care in England. *Health statistics quarterly / Office for National Statistics* 2004;:5–14.

20 Hernán M, Epidemiology JR-, 2006 undefined. Instruments for causal inference: an epidemiologist's dream? *JSTOR*

21 Lourenco J, Paton R, Ghafari M, *et al.* Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. *medRxiv* 2020;:2020.03.24.20042291. doi:10.1101/2020.03.24.20042291

Supplement 1:

Angiotensin Converting Enzyme-Inhibitors including:

Captopril
Enalapril maleate
Enalapril with hydrochlorothiazide
Fosinopril sodium
Imidapril hydrochloride
Lisinopril
Lisinopril with hydrochlorothiazide
Perindopril arginine
Perindopril arginine with indapamide
Perindopril erbumine
Quinapril
Quinapril with hydrochlorothiazide
Ramipril
Ramipril with felodipine
Trandolapril

Angiotensin Receptor Blocker and combinations including:

Amlodipine with valsartan
Azilsartan medoxomil
Candesartan cilexetil
Eprosartan
Irbesartan
Irbesartan with hydrochlorothiazide
Losartan potassium
Losartan with hydrochlorothiazide
Olmesartan medoxomil
Olmesartan with amlodipine
Olmesartan with amlodipine and hydrochlorothiazide
Olmesartan with hydrochlorothiazide
Sacubitril with valsartan

Telmisartan
Telmisartan with hydrochlorothiazide
Valsartan
Valsartan with hydrochlorothiazide

Non-steroidal Anti-inflammatory Drugs including:

Aceclofenac
Aspirin
Aspirin with codeine
Aspirin with metoclopramide
Benzydamine hydrochloride
Bromfenac
Celecoxib
Dexibuprofen
Dexketoprofen
Diclofenac
Diclofenac potassium
Diclofenac sodium
Diclofenac sodium with misoprostol
Etodolac
Etoricoxib
Felbinac
Flurbiprofen
Ibuprofen
Indometacin
Ketoprofen
Ketorolac trometamol
Mefenamic acid
Meloxicam
Nabumetone
Naproxen
Naproxen with esomeprazole
Nepafenac
Parecoxib
Piroxicam
Sulindac
Tenoxicam
Tiaprofenic acid

Tolfenamic acid
Tramadol with dexketoprofen

Chloroquines including:

Chloroquine
Hydroxychloroquine sulfate
Quinine

Thiazolidinediones and their combinations including:

Pioglitazone with metformin
Pioglitazone

Metformin and its combinations including:

Alogliptin with metformin
Canagliflozin with metformin
Dapagliflozin with metformin
Empagliflozin with metformin
Linagliptin with metformin
Metformin hydrochloride
Saxagliptin with metformin
Sitagliptin with metformin
Vildagliptin with metformin

Sulphonylureas including:

Glibenclamide
Gliclazide
Glimepiride
Glipizide
Tolbutamide

Insulin including:

Insulin aspart
Insulin glulisine
Insulin lispro
Insulin degludec with liraglutide
Insulin detemir
Insulin glargine
Insulin glargine with lixisenatide

Protamine zinc insulin
Biphasic isophane insulin
Biphasic insulin aspart
Biphasic insulin lispro
Insulin degludec
Isophane insulin

Sodium glucose co-transporter 2 inhibitors and combinations including:

Canagliflozin
Canagliflozin with metformin
Dapagliflozin
Dapagliflozin with metformin
Empagliflozin
Empagliflozin with linagliptin
Empagliflozin with metformin
Ertugliflozin
Saxagliptin with dapagliflozin

Gliptans and its combinations including:

Alogliptin
Alogliptin with metformin
Empagliflozin with linagliptin
Linagliptin
Linagliptin with metformin
Saxagliptin
Saxagliptin with dapagliflozin
Saxagliptin with metformin
Sitagliptin
Sitagliptin with metformin
Vildagliptin
Vildagliptin with metformin

Glucagon-like peptide-1 receptor agonists and combinations including:

Dulaglutide
Exenatide
Insulin glargine with lixisenatide
Liraglutide

Lixisenatide
Semaglutide

Meglitinides including:

Nateglinide
Repaglinide

Protease inhibitors including:

Atazanavir
Atazanavir with cobicistat
Darunavir
Darunavir with cobicistat
Darunavir with cobicistat, emtricitabine and tenofovir alafenamide
Fosamprenavir
Lopinavir with ritonavir
Ombitasvir with paritaprevir and ritonavir
Ritonavir
Saquinavir
Tipranavir

TNF inhibitors including:

Adalimumab
Certolizumab pegol
Etanercept
Golimumab
Infliximab

Oral steroids including:

Deflazacort
Dexamethasone
Fludrocortisone acetate
Hydrocortisone
Prednisolone

Inhaled and nasal steroids including:

Beclometasone dipropionate
Beclometasone with formoterol
Beclometasone with formoterol and glycopyrronium
Betamethasone
Budesonide
Budesonide with formoterol
Ciclesonide
Fluticasone with azelastine
Fluticasone with formoterol
Fluticasone with salmeterol
Fluticasone with umeclidinium and vilanterol
Fluticasone with vilanterol
Mometasone furoate
Triamcinolone acetonide

Other immune modulators including:

Allopurinol
Azathioprine
Ciclosporin
Cyclophosphamide
Mercaptopurine
Methotrexate
Mycophenolate mofetil
Pimecrolimus
Sirolimus
Tacrolimus

Monoclonal antibodies including:

Atezolizumab
Avelumab
Benralizumab
Bevacizumab
Brentuximab vedotin
Canakinumab
Cemiplimab
Cetuximab
Daratumumab
Dinutuximab beta
Durvalumab

Elotuzumab
Gemtuzumab ozogamicin
Inotuzumab ozogamicin
Ipilimumab
Mepolizumab
Necitumumab
Nivolumab
Obinutuzumab
Omalizumab
Palivizumab
Panitumumab
Pembrolizumab
Pertuzumab
Ramucirumab
Reslizumab
Rituximab
Siltuximab
Trastuzumab
Trastuzumab emtansine
Basiliximab

Cytokine-neutralising monoclonal antibodies including:

Anakinra
Brodalumab
Dupilumab
Guselkumab
Ixekizumab
Risankizumab
Sarilumab
Secukinumab
Tildrakizumab
Tocilizumab
Ustekinumab

Anti-lymphocyte monoclonal antibiotics:

Alemtuzumab
Basiliximab
Belimumab
Natalizumab

Ocrelizumab
Vedolizumab

Janus kinase inhibitors including:

Baricitinib
Tofacitinib
Upadacitinib

Proton pump inhibitors including:

Esomeprazole
Lansoprazole
Naproxen with esomeprazole
Omeprazole
Pantoprazole
Rabeprazole sodium

Statins including:

Atorvastatin
Fluvastatin
Rosuvastatin
Simvastatin
Simvastatin with ezetimibe
Simvastatin with fenofibrate
Pravastatin sodium

Anti-platelets including:

Aspirin
Cilostazol
Clopidogrel
Dipyridamole
Dipyridamole with aspirin
Prasugrel
Selexipag
Ticagrelor

Oral anticoagulants including:

Acenocoumarol

Apixaban
Edoxaban
Phenindione
Rivaroxaban
Warfarin sodium

Calcium channel blockers including:

Amlodipine
Amlodipine with valsartan
Atenolol with nifedipine
Diltiazem hydrochloride
Felodipine
Lacidipine
Lercanidipine hydrochloride
Nicardipine hydrochloride
Nifedipine
Olmesartan with amlodipine
Olmesartan with amlodipine and hydrochlorothiazide
Ramipril with felodipine
Verapamil hydrochloride

Supplement 2:

Sociodemographic variables:

Age
Sex
Ethnicity
Index of Multiple Deprivation
Geographical Location
Smoking status
Alcohol intake
Body mass index

Co-morbidities:

COPD
Asthma
Cystic Fibrosis
Rheumatoid arthritis
Ankylosing spondylitis
Inflammatory bowel disease
Diabetes Mellitus
Hypertension
Cardiovascular Disease
Chronic Liver Disease
Chronic Kidney Disease
Epilepsy
Severe mental illness
Dementia
Parkinson's disease

Social care variables:

Frailty score, e.g. Electronic Frailty Index (eFI)
Falls
Care Home/Nursing Home Resident
Poor mobility bed ridden or housebound