Impact of COVID-19 pandemic on incidence of long-term conditions in Wales: a population data linkage study using primary and secondary care health records

Qi, Cathy; Osborne, Tim; Bailey, Rowena; Cooper, Alison; Hollinghurst, Joe; Akbari, Ashley; Crowder, Ruth; Peters, Holly; Law, Rebecca-Jane; Lewis, Ruth; Smith, Deb; Walker, Mark; Edwards, Adrian; Lyons, Ronan

British Journal of General Practice

DOI:
10.3399/BJGP.2022.0353

Published: 01/05/2023

Publisher's PDF, also known as Version of record

Accepted Manuscript

British Journal of General Practice

Impact of COVID-19 pandemic on incidence of long-term conditions in Wales: a population data linkage study

Qi, Cathy; Osborne, Tim; Bailey, Rowena; Cooper, Alison; Hollinghurst, Joe; Akbari, Ashley; Crowder, Ruth; Peters, Holly; Law, Rebecca-Jane ; Lewis, Ruth ; Smith , Deb; Edwards, Adrian; Lyons, Ronan

DOI: https://doi.org/10.3399/BJGP.2022.0353

To access the most recent version of this article, please click the DOI URL in the line above.

Received 05 July 2022
Revised 20 October 2022
Accepted 25 November 2022

Impact of COVID-19 pandemic on incidence of long-term conditions in Wales: a population data linkage study

Authors
Cathy Qi MSc, statistician and research officer, ORCiD: 0000-0002-0147-1985[1]
Tim Osborne MSc, data scientist and research officer[1]
Rowena Bailey MSc, statistician and senior research officer[1]
Alison Cooper MBBCh, MRCGP, MSc, PhD, clinical research fellow[2] https://orcid.org/0000-0001-8660-6721
Joe P Hollinghurst PhD, statistician and senior research officer[1]
Ashley Akbari MSc, associate professor of population data science research[1] https://orcid.org/0000-0003-0814-0801
Ruth Crowder M.Sc.(Econ), chief allied health professions advisor[3]
Holly Peters BMBS, BMedSci, MPH, MRCGP, clinical lecturer in assessments[4]
Rebecca-Jane Law PhD, research and evidence lead[5]
Ruth Lewis PhD, senior lecturer[6]
Deb Smith, patient and public involvement[2]
Adrian Edwards, B.Med.Sci., MB BS, FRCGP, PhD, professor of general practice[2]
Ronan A Lyons OBE, FmedSci, FLSW, MAE, FFPH, MD, professor of public health[1]

1. Population Data Science, Swansea University Medical School, Faculty of Medicine, Health & Life Science, Swansea University
2. Wales COVID-19 Evidence Centre, Division of Population Medicine, Cardiff University
3. Directorate of Primary care and Mental Health, Health and Social Services Group, Welsh Government
4. Centre for Medical Education, Cardiff University
5. Technical Advisory Cell, Health and Social Services Group, Welsh Government
6. North Wales Centre for Primary Care Research, PRIME Centre Wales, Bangor University

Corresponding author: cathy.qi@swansea.ac.uk

Abstract

Background

The COVID-19 pandemic has indirectly impacted health service provisions owing to surge and sustained pressures on the system. The effects of these pressures on the management of long-term or chronic conditions are not fully understood.

Aim

To explore the effects of COVID-19 on the recorded incidence of 17 long-term conditions.

Design and setting

An observational retrospective population data linkage study on the population of Wales using primary and secondary care data within the Secure Anonymised Information Linkage (SAIL) Databank.

Methods

We presented monthly rates of new diagnosis between 2000 and 2021 for each long-term condition. Incidence rates post-2020 were compared to expected rates predicted using time series modelling of pre-2020 trends.

Proportion of annual incidence was presented by socio-demographic factors: age, sex, social deprivation, ethnicity, frailty and learning disability.

Results

We included 5,476,012 diagnoses from 2,257,992 individuals. Incidence rates from 2020 to 2021 were lower than mean expected rates across all conditions. The largest relative deficit in incidence was in chronic obstructive pulmonary disease corresponding to 343 (95% CI: 230 to 456) undiagnosed patients per 100,000 population, followed by depression, type 2 diabetes, hypertension, anxiety disorders and asthma. A GP practice of 10,000 patients might have over 400 undiagnosed long-term conditions.

No notable differences between socio-demographic profiles of post- and pre- 2020 incidences were observed.

Conclusion

There is a potential backlog of undiagnosed patients across multiple long-term conditions. Resources are required to tackle anticipated workload as part of COVID-recovery, particularly in primary care.

How this fits in

What is already known on this topic

- Studies have reported reduced recording of long-term or chronic condition incidence early in the COVID-19 pandemic
- Evidence for the presence and severity of lags in diagnoses across multiple long-term conditions during the pandemic, and the current status of these lags is limited

What this study adds

- Over 2020 and 2021, recorded incidence across multiple long-term conditions lagged behind projected expectations, representing a substantial backlog of undiagnosed patients, who are unlikely to be receiving systematic monitoring and management.
- Differences in socio-demographic profile of diagnosed patients post-2020 compared to years pre-2020 were not evident, making targeted catch-up initiatives unlikely to be unfeasible

Introduction

The COVID-19 pandemic has had both direct and indirect impacts on the health and care system.[1] Direct effects are those of COVID-19 related illnesses.[2] Indirect effects are highly heterogenous and include delays in cancer services, postponement of elective surgery, and other non-urgent treatments owing to surge pressures on the system. [1] For example, it has been estimated that around 28 million operations were cancelled or postponed globally during the peak 12 weeks of the pandemic's first wave.[3] Non-urgent treatment impacts include harm from cessation or delay of screening services and management of long-term conditions. [1]

A “long-term” or chronic condition is a condition that cannot presently be cured but is controlled by medication and/or other treatment/therapies for example, diabetes and asthma.[4] Long-term conditions are associated with increasing age and deprivation, and the number of people with multiple long-term conditions (multimorbidity) is increasing.[4] Patients with long-term conditions are more intensive users of health and social care services, and before the pandemic accounted for: 50% of general practice (GP) appointments, 64% of outpatient appointments, and 70% of all inpatient bed days.[4]

In primary care, a call and recall system is used to manage long-term conditions, which is offered to patients after a specific diagnosis is made and recorded in condition registries. Primary care activity was substantially reduced in the early months of the pandemic and when activity returned to more usual levels in 2020, acute care displaced much planned care such as long-term condition monitoring and review.[5] It is unknown whether this has resulted in ongoing delays in diagnosis and management for long-term conditions.

Routinely collected data provide opportunity to examine changes in recorded diagnoses. The Secure Anonymised Information Linkage (SAIL) Databank (www.saildatabank.com) contains data from 84% of the GPs and all hospital inpatient and day case activity in Wales.[6 7 8] We sought to examine historic trends in the incidence rates of 17 long-term conditions, and to compare rates in 2020 and 2021 with expected rates over these two years had previous trends continued without interruption. Further, we sought to examine changes in the characteristics of patients with recorded diagnoses to inform resource allocation.

Methods

This was an observational retrospective study reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Data Sources

Anonymised individual-level, population-scale data sources were accessed within the SAIL Databank.[6 7 8 9 10 11 12] Conditions treated in hospital are recorded using International Classification of Diseases version 10 (ICD-10) codes in the Patient Episode Dataset for Wales (PEDW) dataset. Diagnoses from GP records are coded using Read v2 codes in the Welsh Longitudinal General Practice (WLGP) dataset. The Welsh Demographic Service Dataset (WDSD) was used to link birth, death, sex, and lower layer super output area (LSOA) [13] to records extracted from PEDW and WLGP data. Ethnicity categories were identified from 26 linked data sources (Supplementary Table 1).

Study Cohort

We identified residents of Wales diagnosed for the first time with at least one of 17 long-term conditions between January 2000 and December 2021 using ICD-10 or Read v2 codes (Supplementary Tables 2 and 3). The conditions included were anxiety disorders, asthma, atrial fibrillation, coronary heart disease (CHD), chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), dementia, depression, diabetes mellitus, epilepsy, heart failure, hypertension, inflammatory bowel disease (IBD), osteoporosis, peripheral vascular disease (PVD), rheumatoid arthritis, and stroke & transient ischaemic attack (stroke & TIA). These conditions comprise most of the general practice "Quality and Outcomes (QoF) Framework".[14] Further, we identified individuals diagnosed with three diabetes subtypes (type 1, type 2, undetermined) using an algorithm.[15] "Undetermined type diabetes" was assigned when criteria for type 1 or type 2 were not met.

The final study dataset excluded records missing week of birth or sex, or where the diagnosis date was before birth or after death dates.

Variables

Monthly incidence was derived from the number of individuals diagnosed with a long-term condition for the first time, each month.

Age at the earliest found diagnosis date was categorised (<20/ 20-29/ 30-39/ 40-49/ 50- 59/ 60-69/ 70-79/ 80-89/90+years). Sex was male/female. Ethnic groups were analysed using harmonised Office for National Statistics (ONS) categories (White/Black/Asian/Mixed/Other/Unknown). Deprivation was derived from the LSOA code at the time of diagnosis mapped to the 2019 Welsh Index of Multiple Deprivation (WIMD) [16] and categorised in quintiles (1- most deprived to 5- least deprived). Frailty was based on an internationally established cumulative deficit model which utilises an electronic Frailty Index *(*eFI). [17 18 19] eFI scores were used to categorise individuals as: fit, mild, moderate or severely frail using 10-years of previous WLGP data from date of diagnosis. Individuals without sufficient coverage of GP data were assigned to a missing category. Learning disability status (yes/no) was identified for the study cohort using Read v2 codes (Supplementary Table 4). Socio-economic categories with 1 to 4 counts

were rounded to 5 to prevent accidental disclosure and the excess counts deducted from an unknown/missing/adjacent category.

Outcomes

The primary outcome measure was the monthly incidence rates for each long-term condition. This was derived for the full study period from January 2000 to December 2021. The primary analysis used data from January 2015 to December 2021, the primary outcome was the relative difference between observed and expected incidence rates from 2020 to 2021. The secondary outcome was the annual number and proportion of incident cases by each socio-demographic and clinical subgroup.

Statistical Analysis

Monthly incidence rates were derived from the number of new diagnoses occurring each month x 100,000/population size and presented descriptively for the full study period. Population size was estimated from individuals registered to GPs in Wales on 1st July of each year; a breakdown by age group, sex and social deprivation was presented to check population stability over time. The population size of Wales published by the ONS[20] was extracted to estimate coverage achieved by the GP-registered population size. Three-month rolling averages were derived from the mean rate of the month in question, the previous and the following month.

We fitted a seasonal autoregressive integrated moving average (SARIMA) model on monthly incidence data from January 2015 to December 2019 to predict the expected incidence rate (and 95% CI) for each month in 2020 and 2021. Model selection is described in Supplementary Box 1. The difference between the total observed and predicted (lower and upper 95% CI bound) rates was calculated over the two-year period, and for 2020 and 2021 separately. Percentage differences were (observed - expected)x 100/expected rates.

Counts and percentages of persons by demographic groups were presented for each year from 2000 to 2021, and for 2015-2019 and 2020-2021.

Each of the 17 long-term conditions and three diabetes subgroups was examined and analysed separately. As sensitivity analyses, the primary analysis was repeated on the number of cases, unadjusted for population.

Statistical analyses were performed using R V4.1.2.

Public involvement

A public partner contributed public or patient perspective to stakeholder discussions at each stage of the study, including interpretation of the significance and potential impact of the results.

Results

We identified 5,476,012 diagnoses of long-term conditions between January 2000 and December 2021 belonging to 2,257,992 individuals after minor exclusions (Figure 1). Coverage of the population of Wales using GP data in SAIL (Supplementary Table 5) was high (>80% from 2003, and >85% from 2015). Supplementary Table 6 shows that population demographics in the GP population were generally stable from 2000 to 2021.

A fully interactive dashboard showing incidence counts and rates from 2000 to 2021 for all 17 long-term conditions and diabetes subtypes is available here: https://envhe.shinyapps.io/wales-cec-ltc-incidence/ (source code: https://gitlab.com/envhe/wales-cec-ltc-incidence-shiny-dashboard).

Figure 2 shows monthly incidence rates from 2015 to 2021, and predicted rates from 2020 by condition. There was an abrupt reduction around March to April 2020 across all conditions, followed by a general upward trend in subsequent months. Table 1 shows the difference in the total observed and expected incidence rates over 2020-2021 by condition. Observed incidence was lower than mean expected incidence for all conditions, except type 1 diabetes. Predicted rates are not available for osteoporosis as a SARIMA model was not fitted due to inconsistent trends in 2015-2019 data. Conditions with the largest relative deficit in diagnoses were COPD, depression, type 2 diabetes, hypertension, anxiety disorders, and asthma. Observed rates for COPD were 38.4% (95% CI: 29.5% to 45.4%) lower than expected, corresponding to an undiagnosed population of 343 (95% CI: 230 to 456) per 100,000 individuals. Anxiety disorders had the largest absolute undiagnosed population of 830 (95% CI: 281 to 1379) per 100,000. Compared to 2020, estimated differences for 2021 were similar for COPD and anxiety disorders, and smaller but with larger 95% CIs among most other conditions (Supplementary Table 7). Figure 2 suggests that there may still be an overall lag in diagnoses in 2021 for most conditions. Incidence rates for some conditions were close to pre-pandemic levels by the end of 2021, others (e.g. heart failure and stroke) were approaching predicted rates near the start of 2021 but dropped again towards the end of the year.

The estimated rate of underdiagnosis for diabetes mellitus was 178 (95% CI: 57 to 299) in 2020 and 137 (95% CI: -104 to 378) in 2021, similar to corresponding estimates for type 2 diabetes (168 (95% CI: 72 to 263) in 2020 and 132 (95% CI: -38 to 302) in 2021), while the estimated underdiagnosis for type 1 diabetes was 0 (95% CI: -8 to 7) in 2020 and -3 (-11, 5) in 2021.

Results from analysis of incidence counts unadjusted for population size (Supplementary Tables 8 and 9) were consistent with primary findings. SARIMA model specification and estimated parameters for analysis of incidence rates and counts are shown in Supplementary Tables 10 and 11, respectively.

Supplementary Tables 12 to 31 show annual incidence by socio-demographic factors from 2015 to 2021. The study dashboard (link above) includes data from 2000. There was no notable difference between the distribution of cases among categories in 2020 and/or 2021 compared to preceding years for any of the socio-demographic factors, indicating that though overall rates of diagnosis decreased, influences of socio-demographic characteristics on being diagnosed did not drastically differ pre- and post-2020.

Type 1 diabetes was the only condition with an estimated mean net gain in incidence of 8.6% (95% CI: -22.8% to 83.3%). Given that type 1 diabetes is diagnosed in younger patients (around 75% under 50 years old), we investigated whether diagnosis trends differed between younger (<50) and older (>50) populations (Supplementary Figure 1). Most conditions were rare in under 50s (monthly rate <10 per 100,000), but among the remaining conditions, trends within age groups were similar to aggregate trends, including for depression, anxiety and asthma.

As further post-hoc exploration, Supplementary Figures 2 and 3 show that incidence trends by sex and social deprivation groups were also similar.

Discussion

Summary

From 2020 to 2021, there were deficits in recorded incidences across multiple long-term conditions, likely an indirect effect of the COVID-19 pandemic. Increasing demand and workforce vacancies could have affected availability of appointments and postponed diagnostic tests. A typical general practice of 10,000 patients might have over 400 undiagnosed long-term conditions (some potentially occurring in the same individuals). Observed incidence for some conditions (e.g. heart failure and stroke) increased and declined again during 2021, this could reflect changes in healthcare pressures between the alpha wave (September 2020 to March 2021) and the delta wave (June 2021 to December 2021) in Wales. Other conditions were approaching pre-pandemic levels towards the end of 2021 (e.g. asthma), which could reflect condition-specific 'catch-up' activity but an excess would be needed to reach net expected numbers.

Strengths and limitations

Our work included multiple conditions, mostly selected from the QoF framework, previously used to monitor and reward performance in primary care, thus electronic coding quality is generally good though this can vary between individual clinicians and practices. Overall data coverage was close to the full population of Wales.

The assumption that trends in 2015-2019 would persist if COVID-19 had not occurred could not be tested. Possible interactions between COVID-19 and prognosis were not accounted, for example, excess mortality could partially explain the persistent reduction in incidence and could have led to an overestimation of expected rates. However, given that underdiagnosis is evident in a wide range of conditions and in those aged <50, non-presentation and recording may be the biggest issue.

Comparison with existing literature

Observational studies conducted in Spain have reported reduced incidence of multiple chronic diseases in 2020, [21] and substantial reductions in clinical indicators for control and treatment of chronic disease in March and April 2020. [22] A UK based study using primary care data reported reduced incidences of depression (47.1%) and anxiety (40.8%) in Wales, Scotland and Northern Ireland, especially among working age adults registered at practices in more deprived areas.[23] Our work included longer-term data showing there is likely still a lag for most conditions as services have resumed pre-pandemic activity. Further, the pandemic has exacerbated an already high prevalence of undiagnosed COPD.[24 25] UK pandemic guidance to postpone tests which may increase the respiratory transmission of viral infections including spirometry, likely contributed.[26] This might also explain the difference in lag towards the end of 2021 between asthma and COPD, since spirometry is needed to diagnose COPD while a diagnosis of asthma is

based more on the clinical history. Reductions in hospital admissions for infectious exacerbation of COPD following the national lockdown in Wales[27] could also in part explain the reduction in incidence rates.

The absence of deficits in recorded incidence for type 1 diabetes is likely condition-specific rather than owing to a younger patient population since type 1 diabetes inevitably presents soon after symptom onset, and there were no indications that overall trends were confounded by age. Other studies have reported increased incidence in 2020-2021, mostly in younger patients (<18 years) [28 29 30 31] and increased risk following COVID-19 infections [28 29] though it is unclear if the association is causative.

Implications for research and/or practice

Rectifying this backlog of case identification and consequent management deficits is likely to require specific strategic and operational planning at the level of primary care organisations. Targeted catch-up initiatives are unlikely to be feasible due to the lack of socio-demographic characterisation of the missing diagnoses. Consideration for specific resource allocation to enable healthcare staff time to be committed to searching records, testing and screening risk groups (e.g. across cardio-vascular conditions) is needed. Governments and policymakers may need to identify such specific funding to tackle this workload as part of COVID-recovery, alongside other higher profile patient needs such as cancer care and elective surgery. General or condition-specific patient advocacy organisations and charitable foundations may have a role in 'championing' for patients with potentially relevant symptoms to present to primary care (as advocated also for example with potential cancer symptoms[32)], or to seek attendance and 'health checks' among infrequent attenders.

Further research

Further research is undergoing to identify what deficits in condition management, health outcomes and health service impacts have occurred.

Condition	2020 and 2021			
	Observed	Predicted (95% CI)	Change (95% CI)	% Change (95% CI)
COPD	549	892 (779, 1005)	-343 (-456, -230)	-38.4 (-45.4, -29.5)
Depression	1800	2512 (2194, 2830)	-712 (-1031, -394)	-28.3 (-36.4, -17.9)
Type 2 diabetes	837	1136 (871, 1401)	-300 (-565, -34)	-26.4 (-40.3, -3.9)
Hypertension	1663	2231 (1979, 2483)	-568 (-820, -316)	-25.5 (-33, -16)
Anxiety disorders	2503	3333 (2784, 3882)	-830 (-1379, -281)	-24.9 (-35.5, -10.1)
Asthma	756	1006 (898, 1114)	-250 (-358, -142)	-24.9 (-32.2, -15.9)
Diabetes mellitus	999	1314 (952, 1676)	-315 (-677, 47)	-24 (-40.4, 4.9)
Rheumatoid arthritis	148	192 (142, 243)	-45 (-95, 6)	-23.1 (-39, 4)
PVD	341	430 (375, 485)	-90 (-145, -35)	-20.8 (-29.8, -9.2)
Inflammatory bowel disease	147	183 (152, 214)	-36 (-67, -5)	-19.8 (-31.4, -3.4)
Undetermined type diabetes	123	147 (116, 178)	-24 (-55, 7)	-16.3 (-31, 6.1)
CHD	671	774 (680, 869)	-103 (-198, -9)	-13.3 (-22.8, -1.3)
Heart failure	756	871 (753, 990)	-116 (-234, 3)	-13.3 (-23.6, 0.4)
CKD	1462	1678 (1496, 1861)	-217 (-399, -34)	-12.9 (-21.5, -2.3)
Epilepsy	159	182 (143, 220)	-23 (-61, 16)	-12.4 (-27.9, 11.4)
Atrial fibrillation	1158	1304 (1145, 1463)	-146 (-305, 13)	-11.2 (-20.8, 1.1)
Stroke & TIA	592	647 (554, 740)	-55 (-148, 38)	-8.5 (-20, 6.9)
Dementia	1050	1135 (991, 1279)	-85 (-229, 59)	-7.5 (-17.9, 6)
Type 1 diabetes	41	38 (22, 53)	3 (-12, 19)	8.6 (-22.8, 83.3)

Table 1. Total observed and predicted incidence rate per 100,000 population in 2020 and 2021. Conditions are ordered from largest to smallest relative (%) change between observed and predicted rates. COPD: chronic obstructive pulmonary disease. PVD: peripheral vascular disease. CHD: coronary heart disease. CKD: chronic kidney disease. Stroke & TIA: stroke & transient ischaemic attack.

Funding

This work was funded by the Wales COVID-19 Evidence Centre, funded by Health and Care Research Wales. This work was supported by the Con-COV team funded by the Medical Research Council (grant number: MR/V028367/1). This work was supported by Health Data Research UK, which receives its funding from HDR UK Ltd (HDR-9006) funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation (BHF) and the Wellcome Trust. This work was supported by the ADR Wales programme of work. The ADR Wales programme of work is aligned to the priority themes as identified in the Welsh Government's national strategy: Prosperity for All. ADR Wales brings together data science experts at Swansea University Medical School, staff from the Wales Institute of Social and Economic Research, Data and Methods (WISERD) at Cardiff University and specialist teams within the Welsh Government to develop new evidence which supports Prosperity for All by using the SAIL Databank at Swansea University, to link and analyse anonymised data. ADR Wales is part of the Economic and Social Research Council (part of UK Research and Innovation) funded ADR UK (grant ES/S007393/1).

Ethical approval

All research conducted has been completed under the permission and approval of the SAIL independent Information Governance Review Panel (IGRP) project number 0911.

Competing interests

All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; AE declare role as the Director of Wales Covid-19 Evidence Centre as part of university employment, receiving no further payments; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. We would like to acknowledge all the data providers who make anonymised data available for research. We would also like to acknowledge Mark Walker (senior medical officer at the Directorate of Primary care and Mental Health, Health and Social Services Group, Welsh Government) for his contribution to this study.

References

1. Technical Advisory Group, Welsh Government. Five Harms Arising from COVID-19: Considerations of Potential Baseline Measures. Published 2021 July 9. Available from: https://gov.wales/sites/default/files/publications/2021-07/technical-advisory-group-5-harms-arising-from-covid-19_0.pdf

2. Sanchez-Ramirez DC, Normand K, Zhaoyun Y, Torres-Castro R. Long-Term Impact of COVID-19: A Systematic Review of the Literature and Meta-Analysis. *Biomedicines*. 2021;9(8):900. Published 2021 Jul 27. doi:10.3390/biomedicines9080900

3. Carr A, Smith JA, Camaradou J, Prieto-Alhambra D. Growing backlog of planned surgery due to covid-19. *BMJ.* 2021;372:n339. Published 2021 Feb 9. doi:10.1136/bmj.n339

4. Department of Health, NHS. Long Term Conditions Compendium of Information *Third Edition*. Published 2012 May 30. Available from: https://www.gov.uk/government/publications/long-term-conditions-compendium-of-information-third-edition

5. Watt T, Firth Z, Fisher R, Thorlby R, Kelly E. Use of primary care during the COVID-19 pandemic; Patient-level data analysis of the impact of COVID-19 on primary care activity in England. The Health Foundation. Published 2020 September 17. Available from: https://www.health.org.uk/news-and-comment/charts-and-infographics

6. Lyons RA, Jones KH, John G, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak*. 2009;9:3. Published 2009 Jan 16. doi:10.1186/1472-6947-9-3

7. Ford DV, Jones KH, Verplancke JP, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Serv Res*. 2009;9:157. Published 2009 Sep 4. doi:10.1186/1472-6963-9-157

8. Jones KH, Ford DV, Thompson S, Lyons RA. A Profile of the SAIL Databank on the UK Secure Research Platform. *Int J Popul Data Sci*. 2019;4(2):1134. Published 2019 Nov 20. doi:10.23889/ijpds.v4i2.1134

9. Jones KH, Ford DV, Jones C, et al. A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: a privacy-protecting remote access system for health-related research and evaluation. *J Biomed Inform*. 2014;50(100):196-204. doi:10.1016/j.jbi.2014.01.003

10. UK Secure eResearch Platform -SAIL databank. [cited 2019 Aug 13]. Available from: https://saildatabank.com/wp-content/uploads/UKSeRP_Brochure_v1.5.pdf

11. Rodgers SE, Demmler JC, Dsilva R, Lyons RA. Protecting health data privacy while using residence-based environment and demographic data. *Health Place*. 2012;18(2):209-217. doi:10.1016/j.healthplace.2011.09.006

12. Rodgers SE, Lyons RA, Dsilva R, et al. Residential Anonymous Linking Fields (RALFs): a novel information infrastructure to study the interaction between the environment and individuals' health. *J Public Health (Oxf)*. 2009;31(4):582-588. doi:10.1093/pubmed/fdp041

13. Office for National Statistics licensed under the Open Government Licence v.3.0 (2018)

14. Welsh Government. Quality outcome framework (QOF), General Medical Services Contract. Available from: https://statswales.gov.wales/Catalogue/Health-and-Social-Care/NHS-Primary-and-Community-Activity/GMS-Contract

15. Rafferty J, Stephens JW, Atkinson MD, et al. A retrospective epidemiological study of type 1 diabetes mellitus in Wales, UK between 2008 and 2018. Int J Popul Data Sci. 2021;6(1):1387. Published 2021 Apr 15. doi:10.23889/ijpds.v6i1.1387

16. ONS. Welsh Index of Multiple Deprivation (full Index update with ranks): 2019. Release 2019 November 27. Updated 2020 October 1. Available from: https://gov.wales/welsh-index-multiple-deprivation-full-index-update-ranks-2019

17. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. *The Scientific World Journal.* 2001;1:323–336.

18. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age and Ageing.* 2016;45(3):353–360.

19. Hollinghurst J, Fry R, Akbari A, et al. External validation of the electronic Frailty Index using the population of Wales within the Secure Anonymised Information Linkage Databank. *Age and ageing.* 2019;48(6):922–926.

20. ONS. Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland. Release 2021 June 25. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

21. Sisó-Almirall A, Kostov B, Sánchez E, Benavent-Àreu J, González de Paz L. Impact of the COVID-19 Pandemic on Primary Health Care Disease Incidence Rates: 2017 to 2020. *Ann Fam Med*. 2022;20(1):63-68. doi:10.1370/afm.2731

22. Coma E, Mora N, Méndez L, et al. Primary care in the time of COVID-19: monitoring the effect of the pandemic and the lockdown measures on 34 quality of care indicators calculated for 288 primary care practices covering about 6 million people in Catalonia. *BMC Fam Pract*. 2020;21:208

23. Carr MJ, Steeg S, Webb RT, et al. Effects of the COVID-19 pandemic on primary care-recorded mental illness and self-harm episodes in the UK: a population-based cohort study. *Lancet Public Health. 2021;6*(2), e124-e135.

24. Almagro P, Soriano JB. Underdiagnosis in COPD: a battle worth fighting. *Lancet Respir Med.* 2017;5(5):367-368. doi:10.1016/S2213-2600(17)30133-9

25. Bastin AJ, Starling L, Ahmed R, et al. High prevalence of undiagnosed and severe chronic obstructive pulmonary disease at first hospital admission with acute exacerbation. *Chron Respir Dis*. 2010;7(2):91-97. doi:10.1177/1479972310364587

26. PCRS. Diagnostic work up of the patient presenting with respiratory symptoms during the COVID-19 pandemic. 2020 June 01. Available from: https://www.pcrs-uk.org/sites/pcrs-uk.org/files/Diagnosis%20of%20asthma%20and%20COPD%20during%20Covid.pdf

27. Alsallakh MA, Sivakumaran S, Kennedy S, et al. Impact of COVID-19 lockdown on the incidence and mortality of acute exacerbations of chronic obstructive pulmonary disease: national interrupted time series analyses for Scotland and Wales. *BMC Med*. 2021;19(1):124. Published 2021 May 17. doi:10.1186/s12916-021-02000-w

28. McKeigue PM, McGurnaghan S, Blackbourn L, et al. Relation of Incident Type 1 Diabetes to Recent COVID-19 Infection: Cohort Study Using e-Health Record Linkage in Scotland [published online ahead of print, 2022 Jul 26]. *Diabetes Care*. 2022;dc220385. doi:10.2337/dc22-0385

29. Barrett CE, Koyama AK, Alvarez P, et al. Risk for Newly Diagnosed Diabetes >30 Days After SARS-CoV-2 Infection Among Persons Aged <18 Years - United States, March 1, 2020-June 28, 2021. *MMWR Morb Mortal Wkly Rep*. 2022;71(2):59-65. Published 2022 Jan 14. doi:10.15585/mmwr.mm7102e2

30. Unsworth R, Wallace S, Oliver NS, et al. New-Onset Type 1 Diabetes in Children During COVID-19: Multicenter Regional Findings in the U.K. *Diabetes Care*. 2020;43(11):e170-e171. doi:10.2337/dc20-1551

31. Passanisi S, Salzano G, Aloe M, et al. Increasing trend of type 1 diabetes incidence in the pediatric population of the Calabria region in 2019-2021. *Ital J Pediatr*. 2022;48(1):66. Published 2022 May 4. doi:10.1186/s13052-022-01264-z

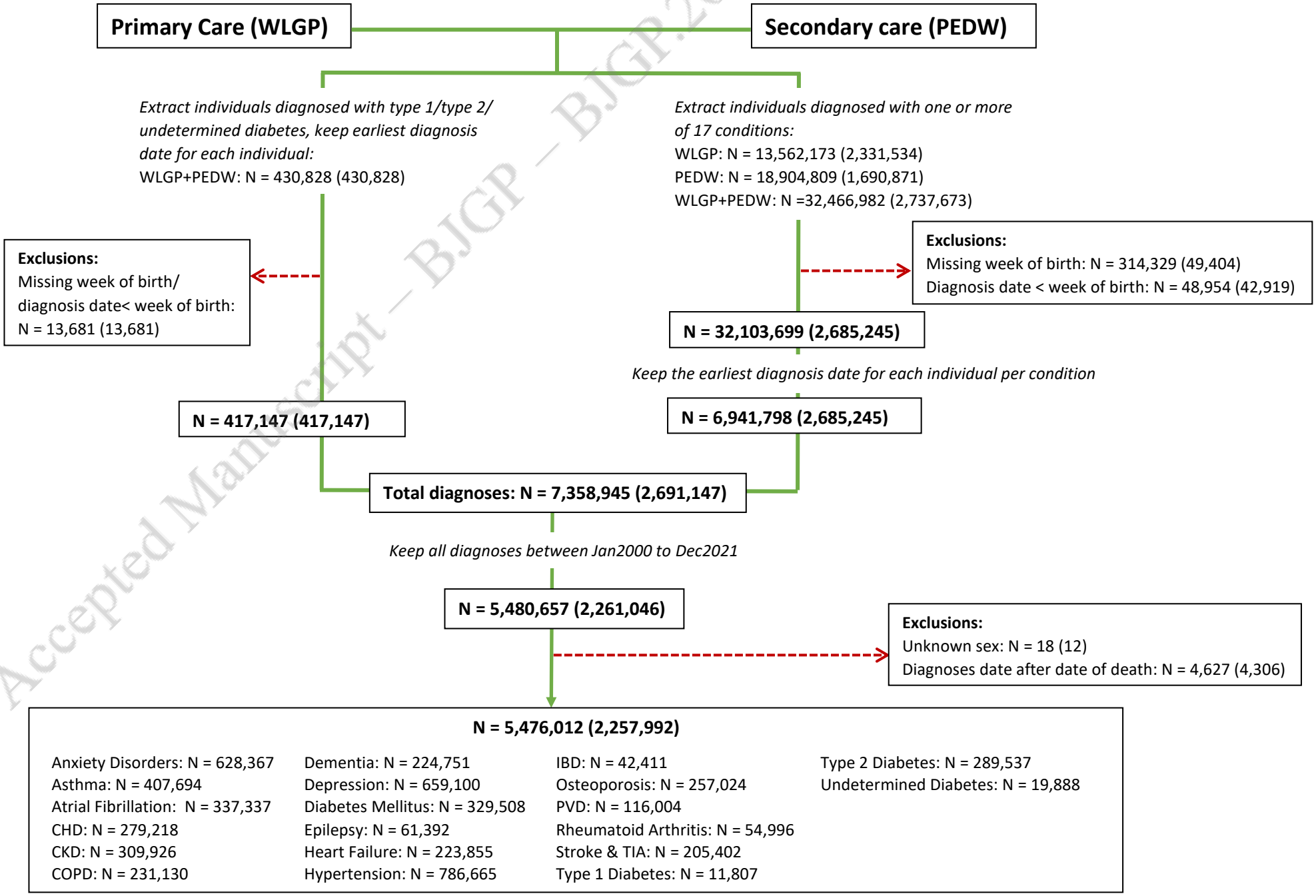
32. Smith P, Moody G, Clarke E, et al. Protocol for a feasibility study of a cancer symptom awareness campaign to support the rapid diagnostic centre referral pathway in a socioeconomically deprived area: Targeted Intensive Community-based campaign To Optimise Cancer awareness (TIC-TOC). *BMJ Open*. 2022;12(10):e063280. Published 2022 Oct 12. doi:10.1136/bmjopen-2022-063280

33. Schaffer AL, Dobbins TA, Pearson SA. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. *BMC Med Res Methodol*. 2021;21(1):58. Published 2021 Mar 22. doi:10.1186/s12874-021-01235-8

Figure titles

Figure 1. Study flowchart: numbers presented are number of diagnoses (number of individuals). Data were extracted in two ways: (1) via using a 'diabetes algorithm' to identify individuals diagnosed with type 1, type 2 or undetermined type diabetes, (2) via using ICD-10 and Read codes to identify individuals diagnosed with one or more of 17 conditions (including diabetes mellitus). For (1), the identification algorithm selected the earliest diagnosis date per individual. For (2), the number of diagnoses refers to the number of unique diagnosis dates available, where a diagnosis date is defined as having one or more diagnosis codes recorded on that day. The final dataset included the earliest recorded diagnosis date for each individual per condition.
WLGP: Welsh Longitudinal General Practice. PEDW: Patient Episode Database for Wales. CHD: coronary heart disease. CKD: chronic kidney disease. COPD: chronic obstructive pulmonary disease. IBD: inflammatory bowel disease. PVD: peripheral vascular disease. Stroke & TIA: stroke & transient ischaemic attack.

Figure 2. Monthly observed number of diagnoses per 100,000 population from 2015 to 2021 for 17 long-term conditions and three diabetes subtypes (type 1/type 2/undetermined). For 2020 and 2021, monthly predicted number of diagnoses per 100,000 are also shown with 95% CIs indicated by the shaded region. Monthly observed data is overlaid with three-month rolling averages (solid line) .



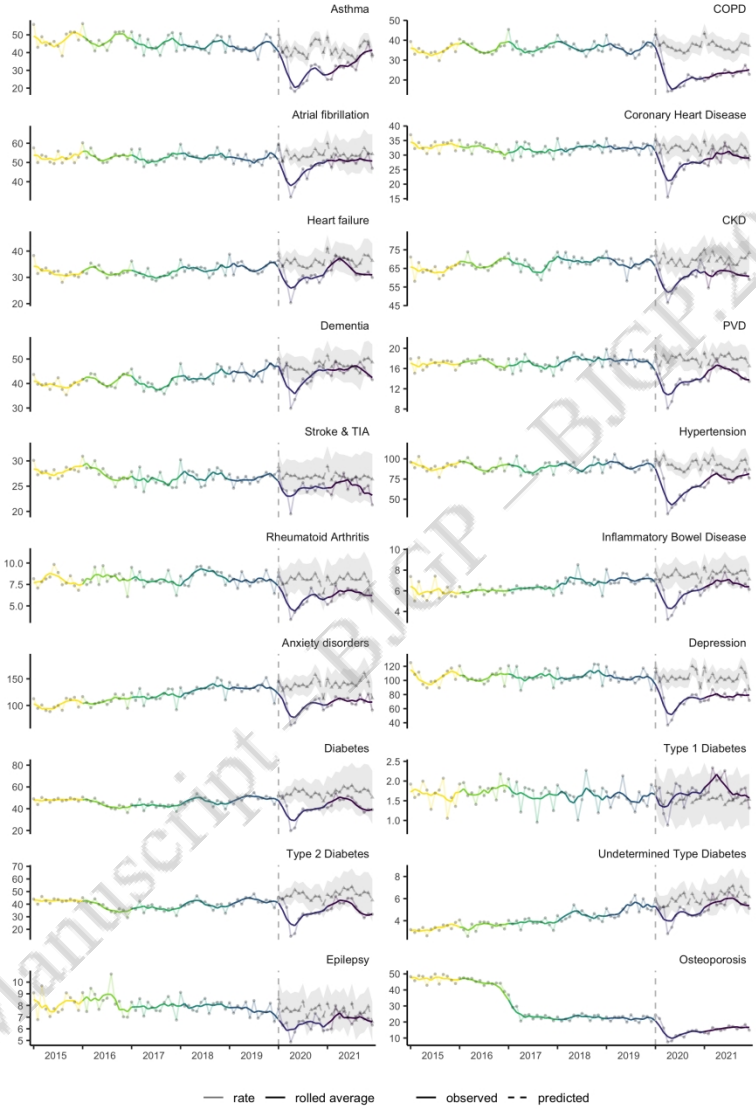


Figure 2. Monthly observed number of diagnoses per 100,000 population from 2015 to 2021 for 17 long-term conditions and three diabetes subtypes (type 1/type 2/undetermined). For 2020 and 2021, monthly predicted number of diagnoses per 100,000 are also shown with 95% CIs indicated by the shaded region. Monthly observed data is overlaid with three-month rolling averages (solid line)

874x1237mm (72 x 72 DPI)