The comorbidity burden of type-2 diabetes mellitus: patterns, clusters and predictions from a large English primary care cohort

Magdalena Nowakowska - research assistant 1,2; Salwa S Zghebi – research associate 1,2; Darren M Ashcroft - professor 1,3; Iain Buchan – honorary professor 2,4; Carolyn Chew-Graham - professor 5; Tim Holt – senior clinical research fellow 6; Christian Mallen - iPCHS director and professor 7; Harm Van Marwijk - professor 8; Niels Peek – professor 9; Rafael Perera-Salazar – professor 8; David Reeves – reader 1,2; Martin K Rutter – senior lecturer and honorary consultant diabetologist 11,12; Stephen F Weng – assistant professor 13,14; Nadeem Qureshi – clinical professor 14; Mamas Mamas – professor 15; Evangelos Kontopantelis – professor 1,2

1. NIHR School for Primary Care Research; Centre for Primary Care, Manchester Academic Health Science Centre (MAHSC), University of Manchester; M13 9PL, Manchester; UK.
2. Division of Population Health, Health Services Research & Primary Care; School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), University of Manchester; M13 9PL, Manchester; UK
3. Division of Pharmacy and Optometry, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), University of Manchester, M13 9PL, Manchester, UK
4. Health eResearch Centre, Division of Informatics, Imaging, and Data Science, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre (MAHSC), University of Manchester; M13 9PL, Manchester; UK
5. Research Institute for Primary Care and Health Sciences, Faculty of Medicine and Health Sciences, Keele University, DJW 1.54a, Staffordshire; UK
6. Nuffield Department of Primary Care Health Sciences, University of Oxford, OX2 6GG, Oxford, UK
7. Arthritis Research UK Primary Care Centre, Research Institute for Primary Care and Health Sciences, Keele University, DJW 1.54a, Staffordshire; UK
8. Division of Primary Care and Public Health, Brighton and Sussex Medical School, University of Sussex, BN1 9PH, Brighton, UK
9. Health eResearch Centre, Division of Informatics, Imaging & Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, M13 9PL, Manchester, UK
10. Centre for Biostatistics, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, M13 9PL, Manchester; UK
11. Division of Diabetes, Endocrinology and Gastroenterology, Faculty of Medicine, Biology and Health, University of Manchester, M13 9PL, Manchester; UK
12. Diabetes Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, M13 0JE, Manchester; UK
13. Division of Primary Care, School of Medicine, University of Nottingham, NG7 2RD, Nottingham, UK
14. Primary Care Stratified Medicine (PriSM) group, Division of Primary Care, School of Medicine, University of Nottingham, NG7 2RD Nottingham, UK
15. Keele Cardiovascular Research group, Centre for Prognosis Research, Institute for Primary Care and Health Sciences, Keele University, ST4 7QB, Stoke-on-Trent, UK

Correspondence:
Evangelos Kontopantelis
5th floor Williamson Building
Division of Population Health, Health Services Research & Primary Care
University of Manchester
Manchester
M13 9PL
Email: e.kontopantelis@manchester.ac.uk
Telephone: 0161-2757646

Word count: 3,775
Abstract

Objective: To quantify and explore comorbidity patterns in patients with type-2 diabetes (T2DM) in the UK at the time of T2DM diagnosis and two, five and nine years after. To estimate the prevalence of six chronic conditions in 2026 and identify clusters of similar conditions using hierarchical clustering analysis.

Design: Population based cohort study

Setting: English general practices contributing to the UK General Practice Research Datalink (CPRD), 2007-2017.

Participants: 102,394 people diagnosed with T2DM between April 2007 and March 2017.

Main outcome measures: Using anonymised primary care electronic health records from CPRD, we investigated the crude and age-standardised prevalence and co-prevalence of 18 chronic conditions at and after T2DM diagnosis. We used linear regression to predict the prevalence of selected conditions for 10 years up to 2027 and hierarchical clustering to group similar conditions. All analyses were stratified by gender and socioeconomic deprivation.

Results: At T2DM diagnosis, 75% of people had at least one additional condition and 31% had at least two. Females had more comorbidities than males and people from the most deprived areas had more conditions than people from the least deprived areas. The prevalence of cardiovascular disease has decreased between 2007 and 2017 and is predicted to decrease further over the next decade. Depression prevalence at the time of T2DM diagnosis has risen in all analysed groups and is predicted to continue to increase to affect over 30% of females and 15% of males by 2026. The cluster analysis showed moderate clustering tendencies in the data with the T2DM-concordand conditions grouped together in all analysed groups.

Conclusion: Most people with T2DM have at least one other chronic condition influencing diabetes progression and management. The great variability in comorbidities patterns between patients emphasises the need for a patient-centred healthcare. Our findings indicate that mental health issues are a growing concern for people with T2DM and there is a need to identify and target interventions that target both physical and mental health in this population.
INTRODUCTION

The prevalence of type-2 diabetes mellitus (T2DM) is increasing both in the UK,¹ and internationally.² Diabetes (all types) is estimated to affect 1 in 11 adults aged between 20 and 79 years, or 415 million adults globally.³ In 2016 it was the seventh leading cause of death worldwide with an estimated 1.6 million deaths directly caused by diabetes.⁴ In the UK over 90% of diabetes cases are T2DM.⁵ Most individuals with T2DM have at least one other chronic condition.⁶ The diabetes related healthcare outcomes, treatment options, care needs and associated cost are complicated by the presence of comorbidities – chronic conditions existing in addition to T2DM.

Due to similar risk factors, people with T2DM have higher risk of cardiovascular complications,⁷ end-stage renal disease,⁸ and hypertension.⁹ However, individuals with T2DM have also been found to have higher risks of depression,¹⁰ thyroid gland diseases,¹¹ and chronic obstructive pulmonary disease (COPD).¹² People with multiple chronic conditions report a number of barriers to self-care such as physical limitations, lack of knowledge, financial constraints, logistic of obtaining care and the need for social and emotional support.¹³ The specific combination of comorbidities in diabetes (type 1 and 2) patients have been found to impact their ability to prioritise and manage the disease.¹⁴ It is suggested that patients with conditions considered unrelated to diabetes may need additional support in making decisions about care priorities and self-management activities.¹⁵ While the presence of diabetes-“concordant” conditions (i.e. sharing the same management goals), tends to be positively associated with quality of care,¹⁶ certain “discordant” comorbidities, like depression and arthritis, impact on treatment options, posing barriers to lifestyle changes and self-care behaviours recommended for diabetes management.¹⁶ ¹⁷

The specific combinations of conditions present dictate the needs of patients, management priorities and the associated demand on healthcare services. A better understanding of the nature, prevalence and patterns of comorbidities in T2DM patients may provide key insights for managing multimorbidity in primary care, and facilitate a more patient-centred approach in risk assessment and more appropriate and tailored therapeutic interventions. In addition, understanding and forecasting the prevalence of specific comorbidities can inform policy-makers in planning and structuring health services to meet the future demands of the population.

In this study, we explored the comorbidities patterns occurring in patients with T2DM over time, as seen in English primary care. We quantified the prevalence of 18, highly prevalent and well recorded physical and mental health conditions and compare the patterns in subgroups of patients stratified by gender, age and socioeconomic deprivation. Focusing on an incidental cohort of patients with T2DM, we explored the patterns in comorbidity occurrence at the time of T2DM diagnosis and after two, five and nine years of follow-up.
METHODS
Data source
The Clinical Practice Research Datalink (CPRD) is a database of anonymised electronic, primary health records. In January 2017, the CPRD held data on nearly 17 million active and historical patients registered with 714 general practices across the UK. It contains detailed information on diagnoses, referrals, tests and therapy records, which are primarily, recorded using Read clinical codes. Additional data is available for a subset of English practices (nearly 75% of English practices; 58% of all UK CPRD practices) which consented to participate in the CPRD linkage scheme and provided patient-level information. To obtain information on social deprivation at the level of the patient’s postcode we used the linked information on the quintiles from the 2015 Index of Multiple Deprivation (IMD) measure, which aggregates data on income, employment, health and disability, education and training, barriers to housing and services, and crime and living environment.

Study Sample
People registered with a general practice in England meeting CPRD data quality standards, and with the first T2DM Read code recorded at any point between 1st April 2007 and 31st March 2017 were included. The inclusion criteria for this study were: patient registered with a CPRD practice for at least 365 days before T2DM diagnosis, aged 35 years and older, and no recorded diagnostic code for type-1 diabetes mellitus (T1DM). In the UK, T2DM has been incentivised since 2004 through a national pay-for-performance scheme, the Quality and Outcomes Framework (QOF), along with another 20 clinical domains approximately, resulting in uniformity in Read code usage and recording. The index date was defined as the date of first recorded code for T2DM and the follow-up was defined as the time between the index date and the earliest of: date of death, transfer out of practice date, last date of data collection from the practice or the end of study period (31st March 2017). The lists of codes used to establish presence of each comorbidity were downloaded from clinicalcodes.org and CPRD@Cambridge websites.

Defining comorbidities
We selected the following 18 conditions: coronary heart disease (CHD), chronic kidney disease (CKD), atrial fibrillation, stroke, hypertension, heart failure, peripheral vascular disease (PVD), rheumatoid arthritis, cancer, osteoporosis, depression, asthma, chronic obstructive pulmonary disease (COPD), dementia, severe mental illness (SMI), epilepsy, hypothyroidism and learning disability. The reporting of these conditions is financially incentivised under the QOF and consequently they are well-recorded in the CPRD. The presence of asthma, epilepsy and depression were determined using Read codes and prescription data, since these can be acute or resolvable. Each condition was considered to be present at the index date if it satisfied the definition criteria at the time of the T2DM diagnosis (available in supplementary material S1, Table S1.1). Each condition was considered to be present during the follow-up period if it satisfied the definition criteria at the index date or at any time during the follow-up.

Statistical analysis
First, we used descriptive statistics to characterise patients in terms of the total number of comorbidities present at the index date (time of the T2DM diagnosis) and after one year, five years and nine years of follow-up. We examined the total number of comorbidities present at and after the index date, stratified by gender and social deprivation quintiles. Age adjusted prevalence was
calculated using the direct age standardisation to the 2013 European Standard Population using 5-year age bands up to 95+ years old. Differences between means of categorical variables were tested using 2-sample t-tests.

We calculated the age-adjusted prevalence of each condition, stratified by gender, for patients from the least and most deprived areas. We also calculated the crude and age adjusted co-prevalence of each pair of comorbidities for the whole sample and stratified by gender, deprivation (focusing on least and most deprived areas) and age (using 35-54; 55-74 and 75+ year old age bands). We present the co-prevalence for the whole sample in the main paper (Figure 2), while the stratified results are available in supplementary material S2.

We longitudinally calculated the prevalence of each comorbidity in the incidental cohort of patients with T2DM, for financial years (April to March) 2007/08 to 2016/17. To forecast the proportion of people diagnosed with T2DM in the next ten years that will also have a particular comorbidity present at the time of diagnosis we used linear regression on log-transformed, age standardised prevalence. For clarity of results, we present the patterns for the six most prevalent conditions as the prevalence of remaining conditions remained relatively low and stable over the study period.

Lastly, we selected patients with multimorbidity, defined as two or more condition present at the index date (in addition to T2DM), and we used agglomerative hierarchical clustering to identify groups of similar conditions. Similarity was assessed using the tetrachoric correlation coefficient. Tetrachoric correlation estimates what the correlation for two binary variables would be if they were measured on a continuous scale. We used Ward’s linkage method to group conditions. At each linkage step, Ward’s method finds a pair of clusters that lead to minimum increase in total within-cluster variance after merging. To avoid chaining (low prevalence comorbidities being sequentially linked to exiting clusters), we excluded conditions with prevalence in a given group below 3%. Cluster analysis was stratified by gender, age bands (35 to 54 years, 55 to 74 years and ≥75 years old) and deprivation using the least and most deprived quintiles. We present the results for the whole sample. Stratified results are available in supplementary material S3. To assess the progression in clustering patterns, we performed the cluster analysis for conditions present at the time of T2DM diagnosis and those present at two, five and nine years after. We plotted the results in dendrograms and identified clusters using visual analysis. Dendrograms visually represent the clustering process of the characteristics of interest. The heights at which conditions fuse together correspond to how similar they are. The earlier the branches merge, the more similar the groups of conditions are. We used R version 3.4.2 for the analysis and data preparation.
RESULTS

We identified 102,394 people with incident T2DM during the study period, who met the study inclusion criteria. A flow chart of the process is available in supplementary material S1, Figure S1.1. The median (LQ – 25th centile; UQ – 75th centile) follow-up was 4.9 years (LQ:2.8; UQ: 7.3). Over half of the sample (56.3%) was male with an average (mean ± SD) age at diagnosis of 60.3 (± 12.5) (Table 1). On average, women were diagnosed at an older age (63.7 ± 13.6, p <0.001). People from most deprived areas, were diagnosed with T2DM at a younger age, compared to those from the most affluent areas (59.3 ± 13 vs 63.9 ± 12.8, p<0.001). The age standardised prevalence of one or more comorbid conditions was 33.3% (95% CI: 32.5%; 34.1%) for the least deprived areas and 32.7% (31.7%; 33.3%) in the most deprived areas. For four or more comorbid conditions, the age standardised prevalence was 2.9% (2.7%; 3.1%) in the most affluent areas and 4.4% (4.1%; 4.7%) in the most deprived areas. More females experienced multimorbidity compared to males (1.6 ± 1.4 vs 1.2 ± 1.2, p<0.001). In all subgroups (by sex and deprivation), the proportion of people with zero comorbidities decreased during the follow up period (Figure 1).

Hypertension was the most common condition among all patients, with higher prevalence among females than males (42.8%; 95% CI: 42.3%; 43.3% vs 45.8%; 95% CI: 45%; 46.4%) (Figure 2). In females, the second most prevalent condition was depression, with higher prevalence in females from the most deprived areas (20.2%, 95% CI: 19.3%; 21.1%), compared to those from most affluent areas (15.6%, 95% CI: 14.7%; 16.5%). In males, the second most prevalent condition was CHD with higher prevalence among males from the most deprived areas (13.6%, 95% CI: 12.9%; 14.3%), compared to those from the most affluent areas (10.8%, 95% CI: 10.3%; 11.3%). During follow-up, the prevalence of depression and asthma decreased in all groups whereas the prevalence of all other conditions increased (prevalence rates for SMI, dementia, epilepsy and learning disability was too low to make meaningful comparisons) (Figure 3). Hypertension and CKD had the highest age-standardised co-prevalence rate among all patients, at 12.1% at the time of T2DM diagnosis and 15.4%, 17.8% and 21.5% after two, five and nine years following the T2DM diagnosis (Figure 4).

Our longitudinal analysis showed a steady decrease in the prevalence of hypertension and relatively stable prevalence rates for CHD, CKD, stroke, and atrial fibrillation. The prevalence of depression increased during the study period for all analysed groups. In females the age-standardised prevalence rate of depression increased from 15.9% (95% CI: 14.8%; 17.0%) in 2007 to 21.5% (19.7%; 20.8%) in 2015 and 18.8% (16.8%; 20.8%) in 2016. In males the age-standardised prevalence rate of depression increased from 7.0% (3.4%; 7.6%) in 2007 to 10.4% (9.1%; 11.7%) in 2016. If the current trend continues, depression can affect over a third of females diagnosed with T2DM by 2026 (age-standardised prevalence: 30.7%, 95% CI: 23.9%; 39.4%) and over 15% (95% CI: 13.2%; 18.9%) of males. The prevalence of depression increased from 9.8% (8.5%; 11.1%) in 2007 to 14.9% (11.3%; 16.5%) in 2016 in the most affluent areas and from 13.4% (12.0%; 14.8%) in 2007 to 17.7% (15.3%; 19.6%) in 2015 and to 14.1% (11.5%; 16.7%) in 2016. If current trend continues, depression is predicted to affect 17.9% (11.7%; 27.5%) of people in the most affluent and 21% (15.9%; 29.5%) of people from the most deprived areas by 2026.

The hierarchical cluster analysis showed conditions being grouped into two main clusters; the first composed of atrial fibrillation, heart failure, PVD, CHD, cancer, stroke, hypertension and CKD and the second composed of depression, SMI, COPD, asthma, hypothyroidism, rheumatoid arthritis and osteoporosis. This pattern was similar in all analysed groups with cancer being included in the first
cluster for males, people from most deprived areas, people age 35 to 74 and 75 and over. However, cancer was linked with cluster two in females, people from least deprived areas and people age 55-74. Moderate clustering tendencies have been observed for conditions present at the time of T2DM diagnosis with the agglomerative coefficient around 0.45 with some variations between groups. This decreased slightly when analysing conditions present two, five and nine years after the diagnosis.

**DISCUSSION**

We showed important changes in the comorbidity patterns in a large real-world cohort of people living with T2DM, using data from the UK primary care. Our findings are relevant to patients, clinicians and policy-makers, and can inform on the healthcare needs and how best to prioritise and deliver care for people with T2DM. We identified alarming levels and trends of depression prevalence in these patients, which we estimated will continue to grow over the next decade. This could have major consequences for how to offer these patients integrated care. Health systems will have to respond to a growing need for diagnosis and management of mental health problems among people with T2DM, underpinned with established links between depression and poor glycaemic control, treatment adherence, diabetes complications, and mortality. The differences in comorbidity patterns observed in groups stratified by gender and social deprivation highlight the need to address the present and increasing health inequalities, particularly with higher multimorbidity prevalence in patients from more deprived areas.

**Strengths and limitations of the study**

To the best of our knowledge, this is the largest study of multimorbidity in patients with T2DM in England. The quality of the data is very high for our study period, primarily due to data recording in line with the QOF and the financial incentives offered to UK primary care for the management of chronic and other conditions such as T2DM. However, the study has limitations. Firstly, due to low prevalence of some conditions in general (such as learning disabiity) and in specific groups (like osteoporosis in males), some comorbidities were excluded from the cluster analysis for all or some strata. However, all conditions were included in the frequency analysis which provides a starting point for the analysis of grouping patterns of specific conditions. Secondly, we selected only 18 conditions for which recording quality was high, but patients may have additional comorbidities impacting on their disease management and quality of life. To identify patients with depression we used algorithm analysing prescriptions as well as diagnostic codes. We were unable to discriminate uses of antidepressants for other conditions such as obsessive-compulsive or bipolar disorders, therefore patients with other mental health conditions might have been incorporated wrongly into the depression group. The predictions of future prevalence rates were obtained from linear regression models, which are dependent on certain assumptions such as the linearity of the trend. Lastly, some of the conditions we modelled may be present but undiagnosed in our cohort. Some diagnostic criteria were also changed during the study period, for example the diagnostic criteria for hypertension. Therefore, the average number of comorbidities calculated in our sample is likely to be underestimated both due to the finite set of conditions we used and to non-diagnosis in practice.
Prior studies

We found that almost 75% of patients had at least one additional comorbidity at the time of T2DM diagnosis, and 44% had at least two comorbidities. Prevalence of multimorbidity was lower than reported in some clinical trials (90%)\textsuperscript{32} or studies using administrative data (91.4\%\textsuperscript{(26)}, (84.6\%)\textsuperscript{(27)}, but higher than in others (44%).\textsuperscript{35} However, our population was younger than in some studies and we analysed a large but not exhaustive list of conditions. As expected, the burden of comorbidity increased with age, however, contrary to previous research,\textsuperscript{(4,21)} which found higher age-standardised prevalence of multimorbidity in males or no gender difference, we found that the burden was higher in females. This reflects the pattern in the general population which shows that females tend to have more comorbid conditions than males.\textsuperscript{36} This difference may relate to the surveillance bias with females being more likely to visit a general practitioner and therefore have a recorded diagnosis of comorbidity. In addition, previous studies tend to focus on conditions regarded as diabetes-concordant such as cardiovascular diseases and CKD.\textsuperscript{(4)} Females with T2DM were found to have lower probability of these having conditions and higher prevalence of depression, which we included in our study.\textsuperscript{33} The presence of mental health problems may have a significant impact on the ability of the patient to manage their condition, progression of T2DM.\textsuperscript{10,28,30} Our findings of high and increasing prevalence of depression in patients with T2DM, implies that the inclusion of mental health conditions is essential in multimorbidity studies for this population.

Prevalence of comorbidities

We observed a higher burden of comorbidity among people from the most deprived areas compared with the most affluent areas. Differences were also observed in the prevalence of specific conditions, notably higher prevalence of depression, CHD, asthma and COPD among people from the most deprived areas. This is consistent with other studies and may be explained by the higher prevalence of risk factors such as smoking, obesity and alcohol consumption.\textsuperscript{37–40}

We found a significant increase in the prevalence of T2DM-comorbid depression, which is expected to rise even more over the next 10 years. The rising prevalence of depression and the large gender gap has also been observed for the general population.\textsuperscript{41} However, people with both T2DM and depression may require tailored approaches of treatment for both conditions.

Our hierarchical clustering analysis showed that conditions regarded as diabetes-concordant (stroke, atrial fibrillation, CKD, CHD, hypertension, PVD and heart failure) tend to group together in all analysed groups. Cancer has been linked with different condition groups, depending on the analysed stratum. This may be due to the fact that we grouped all types of cancer into one condition. However, specific types of cancer may be more prevalent in different groups and be linked with the conditions sharing common risk factors. We also found that conditions present after two, five and nine years from T2DM diagnosis had lower clustering tendencies than conditions present at the time of the diagnosis. This reflects the increasing health complexity associated with aging and disease progression.

Implications

Most people with T2DM have at least one other condition that can influence the self-management of diabetes and its progression. We found high prevalence of T2DM-concordant conditions such as hypertension, coronary heart disease and chronic kidney disease as well as T2DM-discordant conditions such as COPD and depression. The complexity of needs, specific to the patients’
multimorbidity patterns as well as socio-economic situation, has to be considered when developing and providing comprehensive and precise care for people with T2DM. With growing prevalence of T2DM, these complexities have to be taken into account when planning future care services, particularly given the higher cost of treating people with multimorbidity and the lead-times for developing appropriately skilled multi-disciplinary care teams. Further research is needed to identify the best course of action for treating people with multimorbidity, as recent research shows that existing multimorbidity interventions are not particularly effective for improving quality of life.

Our analysis shows that cardiovascular conditions may become less prevalent among people with T2DM; however, clinicians will have to identify and manage the rising burden of comorbid mental health problems. Currently, services targeting people with T2DM are geared towards cardiovascular conditions. The growing burden of mental health conditions will require restructuring of the services and workforce planning.
Contributorship statement
EK, SZ and MN originally designed the study. MN performed the statistical analyses. MN wrote the manuscript and all co-authors critically edited the manuscript.
EK is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Competing interests
DMA has received grant funding from Abbvie and the Leo Foundation. MKR has received educational grant support from MSD and Novo Nordisk; has modest stock ownership in GSK; and has consulted for Roche. Remaining authors have no competing interests to declare.

Ethical approval
Not applicable

Patient involvement
Not applicable

Funding
This study was funded by the National Institute for Health Research (NIHR) School for Primary Care Research (SPCR), under the title “Quantifying severity of chronic conditions in English Primary Care using the Clinical Practice Research Datalink” (project no. 331 and 353). This paper presents independent research funded by the National Institute for Health Research (NIHR). The study funders had no role in the study design, data collection, analysis or interpretation, in the writing of the paper or in the decision to submit the paper for publication. CDM is funded by a National Institute for Health Research (NIHR) Research Professorship (NIHR-RP- 2014-04- 026), the NIHR Collaborations for Leadership in Applied Health Research and Care West Midlands and the NIHR School for Primary Care Research.

Data sharing agreement
This study is based on data from the Clinical Practice Research Datalink (CPRD) obtained under licence from the UK Medicines and Healthcare products Regulatory Agency (MHRA). The data is provided by patients and collected by the NHS as part of their care and support. The study was approved by the independent scientific advisory committee (ISAC) for CPRD research (reference number: 18_097). The interpretation and conclusions contained in this study are those of the authors alone and not necessarily those of the MHRA, NHS, the National Institute for Health Research or the Department of Health. Due to licence restrictions we cannot share the data.

Copyright
The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.
Bibliography

1. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. Diabetes, Obes Metab. 2017 Nov 1;19(11):1537–45.

2. World Health Organization. Global Report on Diabetes. Vol. 978, ISBN. 2016.

3. International Diabetes Federation. IDF Diabetes Atlas Eighth Edition 2017. International Diabetes Federation. 2017.

4. Global Health Estimates. Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2016. WHO. Geneva: World Health Organization; 2016.

5. Holman N, Young B, Gadsby R. Current prevalence of Type 1 and Type 2 diabetes in adults and children in the UK. Diabet Med. 2015;32(9):1119–20.

6. Iglay K, Hannachi H, Joseph Howie P, Xu J, Li X, Engel SS, et al. Prevalence and co-prevalence of comorbidities among patients with type 2 diabetes mellitus. Curr Med Res Opin. 2016 Jul 2;32(7):1243–52.

7. The Emerging Risk Factors Collaboration, Sarwar N, Gao P, Seshasai SRK, Gobin R, Kaptoge S, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet. 2010 Jun 26;375(9733):2215–22.

8. United States Renal Data System. International Comparisons. In United States Renal Data System. 2014 USRDS annual data report: Epidemiology of kidney disease in the United States. Bethesda (MD): National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2014.

9. Waerber B, Feihl F, Rulope L. Diabetes and hypertension. Blood Press. 2001;10(5–6):311–21.

10. de Groot M, Anderson R, Freedland KE, Clouse RE, Lustman PJ. Association of depression and diabetes complications: a meta-analysis. Psychosom Med. 2001;63(4):619–30.

11. Vondra K, Vrbikova J, Dvorakova K. Thyroid gland diseases in adult patients with diabetes mellitus. Minerva Endocrinol. 2005 Dec;30(4):217–36.

12. Feary JR, Rodrigues LC, Smith CJ, Hubbard RB, Gibson JE. Prevalence of major comorbidities in subjects with COPD and incidence of myocardial infarction and stroke: a comprehensive analysis using data from primary care. Thorax. 2010 Nov 1;65(11):956–62.

13. Bayliss EA, Steiner JF, Fernald DH, Crane LA, Main DS. Descriptions of barriers to self-care by persons with comorbid chronic diseases. Ann Fam Med. 2003;1(1):15–21.

14. Kerr EA, Heisler M, Krein SL, Kabeto M, Langa KM, Weir D, et al. Beyond comorbidity counts: how do comorbidity type and severity influence diabetes patients’ treatment priorities and self-management? J Gen Intern Med. 2007 Dec;22(12):1635–40.

15. Ricci-Cabello I, Stevens S, Kontopantelis E, Dalton ARH, Griffiths RI, Campbell JL, et al. Impact of the Prevalence of Concordant and Discordant Conditions on the Quality of Diabetes Care in Family Practices in England. Ann Fam Med. 2015 Nov;13(6):514–22.

16. Ciechanowski PS, Katon WJ, Russo JE. Depression and diabetes: impact of depressive symptoms on adherence, function, and costs. Arch Intern Med. 2000 Nov 27;160(21):3278–
17. Krein SL, Heisler M, Piette JD, Makki F, Kerr EA. The effect of chronic pain on diabetes patients' self-management. Diabetes Care. 2005 Jan;28(1):65–70.

18. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). Int J Epidemiol. 2015 Jun;44(3):827–36.

19. Noble M, Wright G, Smith G, Dibben C. Measuring multiple deprivation at the small-area level. Environ Plan A. 2006;38:169–85.

20. Doran T, Kontopantelis E, Valderas JM, Campbell S, Roland M, Salisbury C, et al. Effect of financial incentives on incentivised and non-incentivised clinical activities: longitudinal analysis of data from the UK Quality and Outcomes Framework. BMJ. 2011 Jun 28;342:d3590.

21. Springate DA, Kontopantelis E, Ashcroft DM, Olier I, Parisi R, Chamapiwa E, et al. ClinicalCodes: An Online Clinical Codes Repository to Improve the Validity and Reproducibility of Research Using Electronic Medical Records. Petersen I, editor. PLoS One. 2014 Jun 18;9(6):e99825.

22. CPRD @ Cambridge. CPRD @ Cambridge - Code List [Internet]. 2017 [cited 2018 Jul 9]. Available from: http://www.phpc.cam.ac.uk/pcu/cprd_cam/codelists/

23. Eurostat. Revision of the European Standard Population Report of Eurostat’s task force. 2013 edition. Luxemburg; 2013.

24. Cornell JE, Pugh JA, Williams, Jr JW, Kazis L, Lee AFS, Parchman ML, et al. Multimorbidity Clusters: Clustering Binary Data From Multimorbidity Clusters: Clustering Binary Data From a Large Administrative Medical Database. Appl Multivar Res. 2009 Jan 13;12(3):163.

25. Uebersax JS. Introduction to the tetrachoric and polychoric correlation coefficients. [Internet]. Web. 2015 [cited 2018 Sep 27]. Available from: http://www.john-uebersax.com/stat/tetra.htm

26. Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc. 1963;58(301):236–44.

27. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [Internet]. 2017. Available from: https://www.r-project.org/

28. Lustman PJ, Anderson RJ, Freedland KE, de Groot M, Carney RM, Clouse RE. Depression and poor glycemic control: a meta-analytic review of the literature. Diabetes Care. 2000 Jul;23(7):934–42.

29. Ciechanowski PS, Katon WJ, Russo JE, Hirsch IB. The relationship of depressive symptoms to symptom reporting, self-care and glucose control in diabetes. Gen Hosp Psychiatry. 2005 Jul;25(4):246–52.

30. Katon WJ, Rutter C, Simon G, Lin EHB, Ludman E, Ciechanowski P, et al. The association of comorbid depression with mortality in patients with type 2 diabetes. Diabetes Care. 2005 Nov;28(11):2668–72.

31. Olier I, Springate DA, Ashcroft DM, Doran T, Reeves D, Planner C, et al. Modelling Conditions and Health Care Processes in Electronic Health Records: An Application to Severe Mental
Illness with the Clinical Practice Research Datalink. Harezlak J, editor. PLoS One. 2016 Feb 26;11(2):e0146715.

32. Teljeur C, Smith SM, Paul G, Kelly A, O’Dowd T. Multimorbidity in a cohort of patients with type 2 diabetes. Eur J Gen Pract. 2013 Mar 22;19(1):17–22.

33. Alonso-Morán E, Orueta JF, Fraile Esteban JJ, Artegaitia Axpe JM, Marqués González ML, Toro Polanco N, et al. The prevalence of diabetes-related complications and multimorbidity in the population with type 2 diabetes mellitus in the Basque Country. BMC Public Health. 2014 Oct 10;14:1059.

34. Luijks H, Schermer T, Bor H, van Weel C, Lagro-Janssen T, Biermans M, et al. Prevalence and incidence density rates of chronic comorbidity in type 2 diabetes patients: an exploratory cohort study. BMC Med. 2012 Dec 29;10(1):128.

35. Struijs JN, Baan CA, Schellevis FG, Westert GP, van den Bos GAM. Comorbidity in patients with diabetes mellitus: impact on medical health care utilization. BMC Health Serv Res. 2006 Jul 4;6:84.

36. Violan C, Foguet-Boreu Q, Flores-Mateo G, Salisbury C, Blom J, Freitag M, et al. Prevalence, determinants and patterns of multimorbidity in primary care: A systematic review of observational studies. PLoS One. 2014;9(7).

37. Lorant V, Deliège D, Eaton W, Robert A, Philippot P, Ansseau M. Socioeconomic Inequalities in Depression: A Meta-Analysis. Am J Epidemiol. 2003 Jan 15;157(2):98–112.

38. Floud S, Balkwill A, Moser K, Reeves GK, Green J, Beral V, et al. The role of health-related behavioural factors in accounting for inequalities in coronary heart disease risk by education and area deprivation: prospective study of 1.2 million UK women. BMC Med. 2016 Dec 13;14(1):145.

39. Duran-Tauleria E, Rona RJ, Weiss K. Geographical and socioeconomic variation in the prevalence of asthma symptoms in English and Scottish children. Thorax. 1999 Jun 1;54(6):476–81.

40. Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of chronic obstructive pulmonary disease in England: a national study of 51,804 patients. Br J Gen Pract. 2010 Jul;60(576):277–84.

41. Randall C, Corp A, Self A. Measuring National Well-Being: Life in the UK. Off Natl Stat. 2014;(February):1–52.

42. Brilleman SL, Purdy S, Salisbury C, Windmeijer F, Gravelle H, Hollinghurst S. Implications of comorbidity for primary care costs in the UK: a retrospective observational study. Br J Gen Pract. 2013 Apr;63(609):e274-82.

43. Salisbury C, Man M-S, Bower P, Guthrie B, Chaplin K, Gaunt DM, et al. Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach. Lancet. 2018;

44. Boult C, Leff B, Boyd CM, Wolff JL, Marsteller JA, Frick KD, et al. A Matched-Pair Cluster-Randomized Trial of Guided Care for High-Risk Older Patients. J Gen Intern Med. 2013 May 10;28(5):612–21.
Table 1: Descriptive statistics on patients with Type-2 Diabetes Mellitus (T2DM) and additional comorbidity

|                          | N (%) | Age (mean ± SD) | Follow-up period (median [LQ;UQ]) | Number of comorbidities at T2DM diagnosis (mean ± SD) | Number of comorbidities 2 years after T2DM diagnosis (mean ± SD) (sample surviving 2 years) | Number of comorbidities 5 years after T2DM diagnosis (mean ± SD) (sample surviving 5 years) | Number of comorbidities 9 years after T2DM diagnosis (mean ± SD) (sample surviving 9 years) |
|--------------------------|-------|-----------------|-----------------------------------|--------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| **Total cohort**         | 102394 (100) | 62.1 ±13.1      | 4.9 (2.8 ; 7.3)                   | 1.4 ±1.3                                               | 1.5 ± 1.4 (84,350)                                                                               | 1.6 ± 1.4 (50,475)                                                                               | 1.7 ± 1.4 (8,977)                                                                               |
| **Gender**              |       |                 |                                   |                                                        |                                                                                                |                                                                                                |                                                                                                |
| Females                 | 44764 (43.7)  | 63.7 ±13.6      | 4.9 (2.7 ; 7.3)                   | 1.6 ±1.4                                               | 1.7 ± 1.4 (36,669)                                                                               | 1.8 ± 1.4 (21,830)                                                                               | 1.9 ± 1.5 (3,942)                                                                               |
| Males                   | 57630 (56.3)  | 60.7 ±12.5      | 5 (2.8 ; 7.3)                     | 1.2 ±1.2                                               | 1.4 ± 1.3 (47,681)                                                                               | 1.5 ± 1.3 (28,645)                                                                               | 1.6 ± 1.4 (5,035)                                                                               |
| **Age bands**           |       |                 |                                   |                                                        |                                                                                                |                                                                                                |                                                                                                |
| 35-54 years             | 31545 (30.8)  | 46.8 ±5.2       | 5.1 (2.9 ; 7.4)                   | 0.8 ±0.9                                               | 0.9 ± 1 (26,368)                                                                                | 1 ± 1 (16,106)                                                                                  | 1.1 ± 1 (2,893)                                                                                |
| 55-74 years             | 51288 (50.1)  | 64.2 ±5.6       | 5.2 (3 ; 7.5)                     | 1.4 ±1.2                                               | 1.6 ± 1.3 (42,950)                                                                               | 1.7 ± 1.3 (26,618)                                                                               | 1.9 ± 1.4 (4,871)                                                                               |
| 75+ years               | 19561 (19.1)  | 81 ±4.9         | 4.1 (2.1 ; 6.5)                   | 2.3 ±1.6                                               | 2.5 ± 1.6 (15,032)                                                                               | 2.6 ± 1.6 (7,751)                                                                               | 2.8 ± 1.5 (1,213)                                                                               |
| **IMD quintiles**       |       |                 |                                   |                                                        |                                                                                                |                                                                                                |                                                                                                |
| Quintile 1 – Least deprived | 19110 (18.7) | 63.9 ±12.8      | 5 (2.8 ; 7.3)                     | 1.3 ±1.3                                               | 1.5 ± 1.3 (15,756)                                                                               | 1.6 ± 1.4 (9,574)                                                                               | 1.7 ± 1.4 (1,682)                                                                               |
| Quintile 2              | 20722 (20.2)  | 63.4 ±13        | 5.1 (2.8 ; 7.4)                   | 1.4 ±1.3                                               | 1.5 ± 1.3 (17,223)                                                                               | 1.6 ± 1.4 (10,500)                                                                               | 1.7 ± 1.4 (1,878)                                                                               |
| Quintile 3              | 21572 (21.1)  | 62.7 ±13        | 4.9 (2.8 ; 7.3)                   | 1.4 ±1.3                                               | 1.5 ± 1.3 (17,811)                                                                               | 1.6 ± 1.4 (10,605)                                                                               | 1.8 ± 1.4 (1,884)                                                                               |
| Quintile 4              | 21393 (20.9)  | 61 ±13.2        | 4.9 (2.7 ; 7.2)                   | 1.4 ±1.3                                               | 1.5 ± 1.4 (17,489)                                                                               | 1.7 ± 1.4 (10,334)                                                                               | 1.7 ± 1.4 (1,839)                                                                               |
| Quintile 5 – Most deprived | 19597 (19.1) | 59.3 ±13        | 4.8 (2.7 ; 7.3)                   | 1.4 ±1.4                                               | 1.6 ± 1.4 (16,071)                                                                               | 1.7 ± 1.4 (9,462)                                                                               | 1.7 ± 1.4 (1,694)                                                                               |

LQ;UQ: lower quintile: Upper quintile; SD: standard deviation; T2DM – type-2 diabetes mellitus, Dx – diagnosis, IMD – Index of Multiple Deprivation
**Figure 1:** Age-standardised (top) and crude (bottom) prevalence of zero, one, two, three and four or more comorbidities present in patients with T2DM at the time of T2DM diagnosis and after two, five and nine years of follow up. Stratified by gender and deprivation.

T2DM – type-2 diabetes mellitus, Dx – diagnosis, IMD – Index of Multiple Deprivation
**Figure 2:** Crude prevalence of chronic conditions among females and males with T2DM from the least and most deprived areas at the time of T2DM diagnosis

IMD – Index of Multiple Deprivation; CHD – coronary heart disease; CKD - chronic kidney disease; COPD - chronic obstructive pulmonary disease; PVD – peripheral vascular disease; SMI – severe mental illness
Figure 3: Age-standardised prevalence of chronic conditions among females and males with T2DM from the least and most deprived areas at the time of T2DM diagnosis

IMD – Index of Multiple Deprivation; CHD – coronary heart disease; CKD - chronic kidney disease; COPD - chronic obstructive pulmonary disease; PVD – peripheral vascular disease; SMI – severe mental illness
Figure 4: Age-standardised prevalence of chronic conditions among females and males with T2DM from the least and most deprived areas two, five and nine years after T2DM diagnosis.

IMD – Index of Multiple Deprivation; CHD – coronary heart disease; CKD - chronic kidney disease; COPD - chronic obstructive pulmonary disease; PVD – peripheral vascular disease; SMI – severe mental illness.
### Figure 5: Crude and age-standardised co-prevalence of chronic conditions among people with T2DM diagnosis and two, five and nine years after.

#### All patients at T2DM Dx

| Condition | Age standardised prevalence | Least prevalent | Most prevalent |
|-----------|----------------------------|----------------|---------------|
| Atrial Fibrillation | 0.7 | 0.1 | 4.8 |
| Asthma | 0.3 | 0.1 | 1.0 |
| CKD | 0.3 | 0.1 | 1.0 |
| COPD | 0.3 | 0.1 | 1.0 |
| Depression | 0.3 | 0.1 | 1.0 |
| Diabetes | 0.3 | 0.1 | 1.0 |
| Hypertension | 0.3 | 0.1 | 1.0 |
| Hypothyroidism | 0.3 | 0.1 | 1.0 |
| Osteoporosis | 0.3 | 0.1 | 1.0 |
| Peripheral vascular disease | 0.3 | 0.1 | 1.0 |
| Rheumatoid arthritis | 0.3 | 0.1 | 1.0 |
| Stroke | 0.3 | 0.1 | 1.0 |

#### All patients 2 years after T2DM Dx

| Condition | Age standardised prevalence | Least prevalent | Most prevalent |
|-----------|----------------------------|----------------|---------------|
| Atrial Fibrillation | 0.6 | 0.1 | 5.6 |
| Asthma | 0.3 | 0.1 | 1.0 |
| CKD | 0.3 | 0.1 | 1.0 |
| COPD | 0.3 | 0.1 | 1.0 |
| Depression | 0.3 | 0.1 | 1.0 |
| Diabetes | 0.3 | 0.1 | 1.0 |
| Hypertension | 0.3 | 0.1 | 1.0 |
| Hypothyroidism | 0.3 | 0.1 | 1.0 |
| Osteoporosis | 0.3 | 0.1 | 1.0 |
| Peripheral vascular disease | 0.3 | 0.1 | 1.0 |
| Rheumatoid arthritis | 0.3 | 0.1 | 1.0 |
| Stroke | 0.3 | 0.1 | 1.0 |

#### All patients 5 years after T2DM Dx

| Condition | Age standardised prevalence | Least prevalent | Most prevalent |
|-----------|----------------------------|----------------|---------------|
| Atrial Fibrillation | 0.4 | 0.1 | 5.8 |
| Asthma | 0.3 | 0.1 | 1.0 |
| CKD | 0.3 | 0.1 | 1.0 |
| COPD | 0.3 | 0.1 | 1.0 |
| Depression | 0.3 | 0.1 | 1.0 |
| Diabetes | 0.3 | 0.1 | 1.0 |
| Hypertension | 0.3 | 0.1 | 1.0 |
| Hypothyroidism | 0.3 | 0.1 | 1.0 |
| Osteoporosis | 0.3 | 0.1 | 1.0 |
| Peripheral vascular disease | 0.3 | 0.1 | 1.0 |
| Rheumatoid arthritis | 0.3 | 0.1 | 1.0 |
| Stroke | 0.3 | 0.1 | 1.0 |

#### All patients 9 years after T2DM Dx

| Condition | Age standardised prevalence | Least prevalent | Most prevalent |
|-----------|----------------------------|----------------|---------------|
| Atrial Fibrillation | 0.4 | 0.1 | 5.8 |
| Asthma | 0.3 | 0.1 | 1.0 |
| CKD | 0.3 | 0.1 | 1.0 |
| COPD | 0.3 | 0.1 | 1.0 |
| Depression | 0.3 | 0.1 | 1.0 |
| Diabetes | 0.3 | 0.1 | 1.0 |
| Hypertension | 0.3 | 0.1 | 1.0 |
| Hypothyroidism | 0.3 | 0.1 | 1.0 |
| Osteoporosis | 0.3 | 0.1 | 1.0 |
| Peripheral vascular disease | 0.3 | 0.1 | 1.0 |
| Rheumatoid arthritis | 0.3 | 0.1 | 1.0 |
| Stroke | 0.3 | 0.1 | 1.0 |

### Notes
- T2DM – type-2 diabetes mellitus; Dx - diagnosis; CHD – coronary heart disease; CKD - chronic kidney disease; COPD - chronic obstructive pulmonary disease; PVD – peripheral vascular disease; SMI – severe mental illness
Figure 6: Observed and predicted prevalence of selected conditions present at the time of type-2 diabetes mellitus (T2DM) diagnosis stratified by gender and deprivation.
CHD – Coronary Heart disease; CKD – Chronic Kidney Disease
Figure 7: Cluster analysis of comorbidities in people with T2DM

CHD – coronary heart disease; CKD - chronic kidney disease; COPD - chronic obstructive pulmonary disease; PVD – peripheral vascular disease; SMI – severe mental illness