# Evaluating ethnic variations in the risk of infections in people with prediabetes and type 2 diabetes: a matched cohort study

Iain M Carey, PhD1 \*, Julia A Critchley DPhil1, Umar A R Chaudhry MBBS1, Stephen DeWilde MD1, Elizabeth S Limb MSc1, Derek G Cook PhD1, Peter H Whincup PhD1, Tess Harris MD1

1 – Population Health Research Institute, St George’s, University of London, London SW17 0RE, United Kingdom

\* Corresponding Author = Iain M Carey, e-mail = i.carey@sgul.ac.uk, Phone = +44 (0)20 8725 5426

## Twitter summary

People with type 2 diabetes or prediabetes are at increased risk of infections, and this risk is similarly elevated across each major ethnic group in England.

## Short running title

Risk of infections in diabetes by ethnicity

## Abstract

### Objective

People living with type 2 diabetes (T2DM) are at higher infection risk, but it is unknown how this risk varies by ethnicity, or whether the risk is similarly observed in people with non-diabetic hyperglycaemia (“prediabetes”).

### Research Design and Methods

We included 527,151 patients in England with T2DM and 273,216 with prediabetes, aged 18-90 and alive on 1st January 2015 on the Clinical Practice Research Datalink. Each were matched to 2 patients without diabetes or prediabetes on age, sex and ethnic group. Infections during 2015-9 were collated from primary care and linked hospitalisation records. Infection incidence rate ratios (IRR) for prediabetes or T2DM were estimated, unadjusted and adjusted for confounders.

### Results

People with T2DM had increased risk for infections presenting in primary care (IRR=1.51, 95%CI 1.51-1.52) and hospitalisations (IRR=1.91, 1.90-1.93). This was broadly consistent overall within each ethnic group, though younger White T2DM patients (age <50) experienced a greater relative risk. Adjustment for socio-economic deprivation, smoking and co-morbidity attenuated associations, but IRRs remained similar by ethnicity. For prediabetes, a significant but smaller risk was observed (primary care IRR=1.35 1.34-1.36, hospitalisation IRR=1.33, 1.31-1.35); these were similar within each ethnicity for primary care infections, but less consistent for infection related hospitalisations.

### Conclusions

The elevated infection risk for people with T2DM appears similar for different ethnic groups and is also seen in people with prediabetes. Infections are a substantial cause of ill-health and health service use for people with prediabetes and T2DM. This has public health implications with rising prediabetes and diabetes prevalence.

## Article Highlights

* People with type 2 diabetes are at higher risk of infections but it is unknown if this varies by ethnicity or whether elevated risks are also seen in people with prediabetes.
* We estimated relative risks of infection for type 2 diabetes compared to non-diabetes and found they were similarly elevated in each major ethnic group in England.
* We also observed an increased relative risk of infections among people with prediabetes compared to people without diabetes.
* Infections are a substantial cause of ill-health and health service use for people with prediabetes and type 2 diabetes.

The proportion of adults in England estimated to have type 2 diabetes (T2DM) has been increasing over recent decades(1), due to a combination of improvements in life expectancy, rising obesity levels(2), and declines in case-fatality(3). Additionally, a growing proportion of UK adults are thought to have non-diabetic hyperglycaemia or “prediabetes”(4,5). Due to improved cardiovascular disease (CVD) risk factor management, CVD mortality among people with T2DM has declined substantially in recent decades, resulting in a larger proportional increase in the burden of other conditions among people with diabetes(6-8). One of these is infectious diseases, which are common in people with diabetes(9), resulting in significant health service use, especially in primary care(10), and which have substantial negative impacts on quality of life(11).

We previously showed that 15% of people with T2DM had a serious infection requiring hospitalisation over a 5.5 year follow-up period, a doubling of risk compared to age-sex matched patients without diabetes(12). However, we and other studies, were unable to assess whether this risk of infection varied by ethnic group. This might arise due to differences in age structure, obesity levels, co-morbidities, other risk factors (such as smoking) or potentially socio-economic status(13). Although few studies have estimated risk by ethnicity among people with diabetes, a higher risk has been reported in the US Black population(14). Assessing the risk among non-white ethnic groups is important since their prevalence of T2DM is markedly higher and the onset is at younger ages(15). Furthermore, little is known about infection risk in the large population of people with prediabetes, where the proportion at younger ages may be higher among non-white ethnic groups(16).

This study therefore aims to extend our earlier work using data from the Clinical Practice Research Datalink (CPRD) in two important ways. Firstly, we take advantage of improved ethnicity recording in the data to compare within different ethnic groups how similar the risk of infection is between people with and without T2DM. Secondly, we investigate for the first time whether an association with infection is also found in people with prediabetes when compared to those without diabetes or prediabetes. Finally, for both of these aims, we additionally provide a picture of the attributable risks of infection, in primary care and for hospitalisations, due to prediabetes and T2DM, both among those with these conditions and across the wider adult population. In doing so, we have chosen to evaluate risks in a period ending just prior to the COVID-19 pandemic before severe disruptions to health service use occurred.

## Research Design and Methods

### Data resource

CPRD is a primary care database in the UK jointly sponsored by the Medicines and Healthcare products Regulatory Agency and the National Institute for Health and Care Research(17). It provides a pseudonymised longitudinal medical record for all registered patients (greater than 99% of the UK population are registered with a General Practitioner), with diagnoses and other clinical information recorded using Read codes. The database recently expanded (CPRD Aurum) to include 16 million currently registered patients(17), with over 80% having their ethnicity recorded(18). Over 90% of contributing CPRD practices in England have consented to their data being linked to external sources; researchers have no access to geographical identifiers such as residential postcode(19). These data sources include HES (Hospital Episodes Statistics), which records every NHS hospital admission in England(20), and the Index of Multiple Deprivation (IMD), a composite small-area (approximately 1500 people) measure used in England for allocation of resources(21). Within CPRD, the distribution of IMD is comparable to the national distribution and provides researchers with a good proxy for individual socio-economic deprivation(22).

### Study design and participants

A retrospective matched cohort study design including all patients aged 18 to 90 alive on 1st January 2015 and actively registered for at least one year, from practices where HES linkage was available. A total of 8,722,348 patients from 1,447 practices in England were eligible (Supplemental Fig. 1). Ethical approval for the study was granted by CPRD’s Research Data Governance (protocol number 21\_000592).

We classified patients with prediabetes or T2DM based on information recorded up to 1st January 2015. Diabetes was first identified from Read codes indicating the patient had been previously diagnosed with diabetes (Supplemental Table 1), and then classified into type 1 or 2 using a strategy developed previously (Supplemental Fig. 2)(12). A total of 527,151 T2DM patients were selected (6.0% prevalence). Patients with type 1 (n=33,851) were not included in this analysis, as we did not anticipate identifying differences in infection risk by ethnicity in this group. Prediabetes patients were identified from the remaining population if they fulfilled any of: (i) Read code for “Prediabetes” before 2015; (ii) HbA1c ≥42 mmol/mol (or ≥6%) during 2013-4; (iii) Read code suggesting impaired glucose tolerance during 2013-4. We excluded any prediabetes patients (n=738) if they had received any anti-diabetes medication (except biguanides) before 2015. A total of 273,216 prediabetes patients were selected (3.1% prevalence).

Patients were grouped into 5 broad ethnicity categories (White, South Asian, Black, Mixed/Other and missing) based on recorded Read codes (Supplemental Table 1)(18). In the UK, ethnicity is predominantly self-reported in primary care records. In our data, we were able to classify ethnicity for approximately 90% of patients with prediabetes or T2DM.

For each patient with prediabetes or T2DM we created two distinct sets of patients without prediabetes or diabetes matched on: (i) age, sex, practice and (ii) age, sex, ethnicity. For each of the four sets produced, patients were randomly selected without replacement from the set of all suitable matches. Thus, it was possible for a patient without prediabetes or diabetes to be selected in each of the four matched sets. Over 99% of prediabetes/T2DM patients were matched (Supplemental Fig. 1), and overall, at least one match was found for >98% of patients within each ethnic group. All patients were followed up to the earliest date of: patient death or de-registration, practice leaving CPRD, or 31st December 2019. We also carried out a sensitivity analyses for prediabetes patients who received a diagnosis of diabetes during the study, by censoring their follow-up time on the day of the first diabetes diagnosis.

### Infection outcomes and covariates

We classified and grouped infections broadly along the same lines as our previous study(12). First, we updated an extensive list of Read codes (primary care) and ICD-10 codes (hospital data) for all infection diagnoses (Supplemental Table 1). Secondly, we searched electronically in the data over a 5-year period (2015-2019) for the following: (i) any infection with a prescription in primary care for an antibiotic, antifungal or antiviral within +/- 14 days of the diagnosis; (ii) any new hospital episode where an infection was the primary diagnosis. In the UK, hospital data is organised into finished consultant episodes and assigned a primary diagnosis(20). Subsequent episodes can be assigned to the same admission, with a different primary diagnosis e.g., a hospital acquired infection. For each summary group, only one event was counted within a 90-day period, with multiple codes assumed to be the same event. Additionally, we carried out an analysis with each of the individual infection groups again restricting to one event per group within a 90-day period. We also extracted patient information on smoking history, body mass index (BMI), and co-morbidities as of 1st January 2015. We selected 12 chronic conditions routinely collected as part of the Quality and Outcomes Framework (QOF), a UK wide system for performance management and payment of GPs in primary care (23). These were atrial fibrillation, cancer, chronic obstructive pulmonary disease, coronary heart disease, chronic kidney disease, dementia, epilepsy, heart failure, hypertension, peripheral vascular disease, serious mental Illness and stroke.

### Statistical methods

Conditional Poisson regression compared infection rates during follow-up between patients with prediabetes or T2DM to those without prediabetes/diabetes, with an offset fitted for total days of follow-up time in the study (Stata version 15). These were conditioned on the match-sets, which implicitly controls for the matching factors (age, sex, practice/ethnicity). These were initially fitted without any further adjustment, but we also fitted models that adjusted for socio-economic status (IMD quintile, with quintile 1 representing the most deprived 20% small areas in England), smoking and a count of co-morbidities. To assess the impact of ethnicity as a confounder, we compared results from the ethnicity matched with the non-ethnicity matched analysis. To explore the impact of age and ethnicity as effect modifiers, we fitted stratified models by age group (18-50, 51-70, 71-90) and by ethnicity separately as well as together. For these analyses, we present and compare unadjusted models in the main analysis, as between ethnic group differences in key confounders, such as socio-economic deprivation, will be indirectly controlled as each model only compares within ethnic group. However, we also provide adjusted estimates (by ethnicity) in the supplementary material. Sensitivity analyses explored the impact of additionally adjusting for obesity, and censoring follow-up time for prediabetes patients who received a diagnosis of diabetes during the study period. Finally, attributable risks for infection in people with prediabetes and T2DM, and population attributable risks were estimated for each ethnic group, assuming any observed infection risk is the direct cause of prediabetes or T2DM. These were estimated using models stratified by 10-year age-group (18-29, 30-39 and so on to 80-89) and summed using a weighted average(24). We provide calculations derived from the unadjusted and adjusted models described above.

### Role of the funding source

The study funder had no role in study design, collection, analysis, and interpretation of data; writing of the report or decision to submit the paper for publication. The corresponding author had full access to all the data in the study and final responsibility for the decision to submit for publication.

### Data and resource availability

The data that support the findings of this study are available from Clinical Practice Research Datalink (CPRD) obtained under license from the UK Medicines and Healthcare Products Regulatory Agency (MHRA), but restrictions apply to the availability of these data, which were used under license for the current study and therefore are not publicly available. CPRD data governance and the license to use CPRD data does not allow distribution of patient data directly to other parties. Researchers must apply directly to CPRD for data access (<https://www.cprd.com>). However, code lists generated during the current study are available in the repository <https://10.24376/rd.sgul.21565557>.

## Results

### Study population

Table 1 summarises the baseline characteristics of patients with prediabetes and T2DM by ethnicity. Approximately 70% of patients were classified as white ethnicity, followed by about 10% as South Asian ethnicity. In the wider population, crude prevalence was highest for South Asian ethnicity (5.5% prediabetes, 11.1% T2DM), with this difference more striking at younger ages (Supplemental Fig. 3 and Supplemental Table 2). Patients with prediabetes were more likely to be female for all ethnicities, whereas people with T2DM were more likely to be male (except for Black ethnicity). The White ethnicity group had a mean age of 67 years for both prediabetes and T2DM. Non-white ethnicities were on average 10-13 years younger for prediabetes, and 5-7 years younger for T2DM, and consequently had less co-morbidities. For both prediabetes and T2DM, non-White ethnicities were more likely to live in deprived areas, with 4-in-10 in the Black ethnicity group residing in the most deprived quintile. Among those with prediabetes, Black patients had the highest average recorded BMI, with 48.3% being greater than 30kg/m2, whilst for those with T2DM, White people had the highest BMI levels (54.3% greater than 30kg/m2).

We also compared baseline characteristic differences between patients with prediabetes or T2DM and age-sex-ethnicity matched patients without prediabetes/diabetes (Supplemental Table 3). Both patients with prediabetes and T2DM were more likely to live in deprived areas, be more obese, have a history of smoking and have co-morbidities than patients without prediabetes or diabetes. The relationship with socio-economic deprivation was further explored within ethnic group (Supplemental Fig. 4) and showed that the association of greater deprivation with prediabetes or T2DM, compared to patients without prediabetes/diabetes, is maintained despite differences in overall deprivation between ethnic groups.

### Overall findings for infections

There were significant increases in infection risk for both patients with T2DM and prediabetes compared to patients without prediabetes/diabetes (Table 2). Comparisons using the age-sex-ethnicity matched versus the age-sex-practice matched comparison group yielded similar results, and all analyses from this point use the age-sex-ethnicity matched group. For T2DM, the (unadjusted) relative risks of infection were higher for hospitalisations (IRR=1.91, 95%CI 1.90-1.93) than primary care infections (IRR=1.51, 95%CI 1.51-1.52). The relative association with primary care infections was similar by sex, but slightly higher in women for hospitalisations (IRR=2.02 95%CI 1.99-2.05 vs IRR=1.83 95%CI 1.80-1.85). Larger relative associations were found at younger ages for both outcomes (e.g., IRR=2.96 95%CI 2.85-3.08 for hospitalisations and T2DM for ages 18-50) where the infection rates among non-diabetes patients were comparatively lower. Even after accounting for the greater number of co-morbidities in people with T2DM (as well as differences in smoking and socio-economic deprivation), associations with both infection outcomes were still observed among people with T2DM. Adjusting for obesity only explained a small proportion of the observed association and was less influential than co-morbidity in the models (Supplemental Table 4).

For people with prediabetes, (unadjusted) infection risks were smaller overall and similar for primary care (IRR=1.35 95%CI 1.34-1.36) and hospitalisations (IRR=1.33, 95%CI 1.31-1.35) and showed similar gradation with age. After adjustment for confounders, the association with hospitalisation was reduced further, especially among older ages. Censoring prediabetes patients diagnosed with diabetes during follow-up had minimal impact on these associations (Supplemental Table 4).

### Infection findings by ethnic group

Figure 1 plots crude infection rates (primary care, hospitalisations) in prediabetes (orange bars) and T2DM (blue bars) patients by ethnicity. These are compared to matched patients without diabetes/prediabetes (white bars). Among patients with T2DM, infections in primary care were highest for South Asian people (235.1 per 1,000 per year), while hospitalisation infections were highest for the White group (68.0 per 1,000 per year). However, compared to patients without prediabetes/diabetes in the same ethnic group, the relative increase in risk for both infection outcomes were broadly similar across ethnic groups (e.g., IRRs for hospitalisation were: South Asian=1.98 95%CI 1.91-2.06, Black=1.87 95%CI 1.76-2.00, Mixed/Other=1.98 95%CI 1.88-2.08, White=1.88 95%CI 1.86-1.98). For prediabetes, the pattern of a similar increase in infection risk by ethnic group was observed for infections in primary care but was less consistent for hospitalisations where Black people with prediabetes had no statistically significant increase in risk compared to Black people without prediabetes/diabetes (IRR=1.07, 95%CI 0.96-1.18).

As people of non-white ethnicities with prediabetes or T2DM were on average younger (and healthier) in our data than white people with the same conditions, we stratified the (unadjusted) ethnic specific IRRs by age (Figure 2), and also adjusted the IRRs for potential confounders (Supplemental Tables 5 and 6). These reveal that for T2DM, there tends to be a higher relative risk among the youngest ages (18-50) in the White ethnic group (IRR=2.12 95%CI 2.07-2.16 for primary care infections, IRR=3.23 95%CI 3.08-3.40 for hospitalisations) compared to the risks found in non-White ethnicities. Above age 50 the IRRs are generally similar between ethnicities for both prediabetes and T2DM, though the Black ethnicity prediabetes group showed no increased risk with hospitalisations. Adjusting for confounders within each ethnic group attenuated all associations, but not did alter findings made on comparisons made between ethnic group.

Finally, we investigated associations with specific infection types by ethnicity in prediabetes and T2DM (Supplemental Tables 7-8). Within each ethnic group, associations were consistently observed for every infection group apart from extremely rare ones where power was low. Upper respiratory tract infections were almost twice as common in South Asian T2DM patients compared to other ethnic groups with T2DM. However, when compared to South Asian patients without diabetes/prediabetes the relative increase in risk was more similar (IRR=1.43, 95% 1.40-1.46), albeit still higher than the corresponding relative risk within the White ethnic group (IRR=1.21, 95% CI 1.20-1.22).

### Attributable Risk Estimates

Attributable risk fractions by ethnic group, due to prediabetes and T2DM were estimated for primary care and hospital infections (Supplemental Table 9), and then weighted accordingly to create population wide estimates. The attributable fractions by ethnic group among patients with prediabetes or T2DM derived from the IRRs in Figure 1 were 32-36% for primary care infections and 49-56% for hospitalisation infections for T2DM, while for prediabetes they were similar for primary care infections (24-32%) but lower for hospitalisation infections, especially among black people with prediabetes (7%). The percentage of infections amongst all adults in the population attributable to prediabetes or T2DM was 5.8% for primary care and 10.0% for hospitalisations. When estimated by ethnic group, these were highest among South Asian people (11.6% and 17.4% respectively).

## Conclusions

Our study has two key findings. Firstly, we have shown that the relative risks for infection associated with T2DM appear to be broadly similar within each ethnic group, this remains true at different ages, and after adjustment for potential confounders. As in our previous work(12), compared to people without prediabetes/diabetes of the same age, sex and now additionally ethnicity, the risk of hospitalisation for infection was roughly doubled among people with T2DM, and about 50% higher for infections requiring primary care contact and an associated prescription. Secondly, we have shown that increased risks for infection are also present in people with prediabetes, albeit lower – about a 30% increase for both infection outcomes when compared to people without diabetes/prediabetes.

### Strengths and limitations

Overall, a major strength of our study is the extremely large sample size (8 million total adults, 750,000 with prediabetes and T2DM). The large number of patients with prediabetes is a likely result of the increased emphasis on both vascular health checks and specifically diabetes prevention and screening in primary care(25), though it still may be underestimating the true scale of prediabetes in the general population, as this continues to rise globally(26). However, our sample is still likely highly representative of people living with prediabetes or T2DM in England prior to the COVID-19 pandemic(17). Previous CPRD analyses(12,27,28), have been based on an earlier dataset (GOLD) which underrepresented major urban areas with ethnic diversity. By utilising the newer Aurum database, with higher overall recording of patient ethnicity, we were able to successfully match patients by broad ethnic group unlike previous studies(27); a stronger design for assessing the pattern of between ethnic differences in risk. In our study, matching was highly successful overall and in each ethnic group; over 98% of people living with prediabetes or T2DM had at least one patient without diabetes matched on age and sex. Additionally, we selected patients without diabetes or prediabetes matched on practice rather than ethnicity to establish the impact of ethnicity as a confounder in our earlier findings. Although about 10% of patients with T2DM or prediabetes could not be assigned an ethnicity, we retained these patients in the analysis to assess any potential bias. These patients tended to be older and more affluent, largely resembling the characteristics of the White ethnicity group, suggesting that we captured a high proportion of non-White ethnicities with T2DM or prediabetes.

Importantly, our analysis was able to consider the impact of age as an effect modifier and show that relative risks were higher in the youngest age group, where the baseline risks of infection in the population without prediabetes or diabetes tends to be much lower. For example, people under 50 with T2DM were three times more likely to be hospitalised with an infection than people without prediabetes or diabetes. However, the increase in infection risk with age in the population reference group is such that the absolute risk differences are still greater at older ages despite the lower relative risk. Since non-white ethnic groups in the UK are younger on average, we were able to investigate how age was modifying the relationships with the overall ethnic-specific associations. This revealed that younger people with T2DM of White ethnicity were at greater risk (both relative and absolute differences), while people with prediabetes from Black and Afro-Caribbean ethnic minorities had little or no excess risk of hospitalisation for infection. We further considered the impact of factors more common in people with prediabetes or T2DM (socio-economic deprivation, smoking, co-morbidity), even though some of their co-morbidity such as heart disease may be a result of their diabetes. Although this may represent an overadjustment, our conclusions regarding ethnicity and infection risk were not altered.

Clearly, primary care recording of infection outcomes is pragmatic and imperfect, based on clinical diagnoses. However, the magnitude of infection risk was stronger for infections resulting in a hospitalisation, where infection recording is more often supported by laboratory findings. Identification of both diabetes and prediabetes is dependent on general practice consultations; prediabetes in particular may be significantly under-diagnosed in primary care(4), resulting in misclassification of exposure, and potentially underestimating the association with infection risk. However, the pattern of infection risk identified for this patient group appear similar, albeit lower, than that for people with T2DM. Our study is based on retrospective cohort data from a period ending just prior to the COVID-19 pandemic. Incidence and recording of infections, especially in primary care, has likely been profoundly affected by the disruption to normal health service delivery during the pandemic so we believe it is more informative to assess non-COVID infection risk prior to this date. Future studies could focus on how this has been affected since 2020.

### Comparisons with other studies

Ethnic differences in infection risk have been a topic of considerable recent interest since persistent mortality differences by ethnicity were noted from the early stages of the COVID-19 pandemic(29). Subsequently, a higher risk of testing positive for SARS-CoV-2 among non-white ethnic groups was also confirmed(30). To the best of our knowledge, few studies have considered whether infections in general might differ by ethnicity in people with diabetes. The higher prevalence of tuberculosis infection reported among Hispanic and Asian people in the US has been hypothesised to explain the association between diabetes and tuberculosis(31). While this might reflect poor health care access, socio-economic deprivation or potentially publication bias (since most studies only report on individual infections), a recent US cohort study reported higher risks of infection-related hospitalisation associated with diabetes in younger people and in people of Black ethnicity(14). However, the study size (about 1,500 people with diabetes) meant it could not stratify by ethnicity or consider whether differences in age among the Black population might explain ethnic variation in infection risk.

In our study, we were able to report on differences in infection rates by ethnic group at different ages, which suggested that non-White ethnicities were less likely to be hospitalised for an infection. In the UK, overall hospitalisation rates have been shown to vary by ethnicity across different conditions(32) likely due to pervasive inequalities and differences in health seeking behaviours(33). For example, people from Black African and Caribbean ethnicities have higher reported hospitalisations for diabetes and endocrine disorders relative to the white population, but their overall hospitalisation rates are significantly lower, possibly reflecting barriers to access within the healthcare system(32). By contrast, some specific infection groups we reported on were more common in non-White ethnicities, such as upper respiratory infections in South Asian people. These may reflect generally higher household size found for Bangladeshi and Pakistani people in the UK(34), representing a greater risk of transmission among household members. However, the higher or lower infection rates by ethnicity were true irrespective of diabetes, so estimated relative risks of infection for prediabetes or T2DM versus people without diabetes tended to be similar within each ethnic group (with the possible exception of a greater infection risk among younger White people with T2DM). We did not assess whether diabetes clinical management or severity differed by ethnicity, though a previous study found improved risk factor recording and faster time to starting anti-diabetes medications among ethnic minority groups(35).

### Implications of study findings

Management of infections in people living with diabetes is clearly a common problem for patients and doctors but had not always received much attention before the pandemic. Whilst NICE guidelines for people with T2DM provide little mention of infections(36), we have shown that attributable risks in those living with T2DM are substantial (34% for primary care, 50% for hospitalisations) and consistent by ethnic group. While some of this attributable risk may be explained by their poorer overall health, it does not change the fact that these patients experience a large burden of infections by having T2DM. Our analyses also demonstrated that higher risks were also present, though somewhat reduced, in people with prediabetes in both healthcare settings. This suggests that any infection risk does not begin with a diagnosis of diabetes but is instead on a continuum and already present in this large group of patients. Given the high and rising prevalence of prediabetes and T2DM(3,4,26) this is of considerable importance and results in a substantial population burden affecting people with diabetes and prediabetes and health services. We estimate that about 6% of all adult infections treated in primary care and 10% of hospitalisations due to infection are statistically attributable to prediabetes or T2DM, and these will continue to increase with rising prevalence(2). The South Asian population experiences the greatest estimated burden (12% and 17% respectively) as a result of the higher prevalence of prediabetes and T2DM at all ages.

There has been little previous development of interventions to reduce infection risks in people with diabetes; infection outcomes have not been included in most major diabetes management trials(11). Therefore, robust observational data may instead provide an initial path to developing interventions, potentially based on increased self-management of risk factors and awareness by promptly identifying infections and knowing when to seek professional help, which could lead to reduced hospital admissions for some infections. Enhanced glucose monitoring, offering better diabetes control among the highest risk individuals, might improve infection outcomes, but this has not been clearly demonstrated and thus remains a gap in our knowledge. As we have previously shown that poor glycaemic control is associated with the risk of serious infections(37), future work could establish whether any association with HbA1c level is also present in people with prediabetes.

In conclusion, we have used large routine health databases in England to reveal that the increased infection risk associated with T2DM is generally consistent across different ethnic groups and also observed in people with prediabetes. The burden of infections attributable to T2DM and prediabetes is significant and will continue to have public health implications with their rising prevalence. Efforts to reduce infection risk in people with diabetes remains an important challenge in all ethnic groups.

## Acknowledgments

We thank our helpful Patient Study Advisory Group for their input, particularly to discussion and interpretation of the results.

### Funding and assistance

This study is funded by the National Institute for Health and Care Research (NIHR) - Research for Patient Benefit Programme (NIHR202213) and supported by the NIHR Applied Research Collaboration South London (NIHR ARC South London) at King’s College Hospital NHS Foundation Trust. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

### Author contributions

IC, JC, TH, SDW, UC and DC were involved in the conception and design of the study. JC led on funding acquisition. IC led the data curation and statistical analysis. All authors were involved in the interpretation of data for the work. IC and JC wrote the first draft of the manuscript with input from all co-authors. All authors approved the final version for publication. IC 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.

### Duality of interest disclosure

We declare no competing interests.

## Tables

### Table 1 – Baseline characteristics of all patients with prediabetes and type 2 diabetes by ethnicity

|  |  |  |  |
| --- | --- | --- | --- |
|  | Prediabetes (n=273,216) |  | Type 2 Diabetes (527,151) |
|  | South Asian | Black | Mixed/ Other | White | Ethnicity unknown |  | South Asian | Black | Mixed/ Other | White | Ethnicity unknown |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Total,** n(% overall) | 27,368 (10.0) | 13,366 (4.9) | 15,331 (5.6) | 189,195 (69.3) | 27,956 (10.2) |  | 54,913 (10.4) | 22,533 (4.3) | 30,289 (5.8) | 368,253 (69.9) | 51,163 (9.7) |
| Estimated prevalence in the population (%)† | 5.5 | 4.9 | 3.3 | 3.2 | 1.8 |  | 11.1 | 8.3 | 6.6 | 6.2 | 3.2 |
|   |  |  |  |  |  |  |  |  |  |  |  |
| **Matching\***, n(%) |  |  |  |  |  |  |  |  |  |  |  |
|  Age-sex-practice | 27,246 (99.6) | 13,338 (99.8) | 15,300 (99.8) | 188,920 (99.9) | 27,922 (99.9) |  | 53,776 (97.9) | 22,248 (98.7) | 30,076 (99.3) | 367,264 (99.7) | 50,935 (99.6) |
|  Age-sex-ethnicity | 27,366 (100.0) | 13,364 (100.0) | 15,329 (100.0) | 189,170 (100.0) | 27,949 (100.0) |  | 53,763 (97.9) | 22,465 (99.7) | 30,237 (99.8) | 368,206 (100.0) | 51,141 (100.0) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Sex**, n(%) |  |  |  |  |  |  |  |  |  |  |  |
|  Female | 14,267 (52.1) | 7,351 (55.0) | 8,298 (54.1) | 97,068 (51.3) | 13,900(49.7) |  | 25,214 (45.9) | 11,389 (50.5) | 14,298 (47.2) | 159,901 (43.4) | 22,319 (43.6) |
|  Male | 13,101 (47.9) | 6,015 (45.0) | 7,033 (45.9) | 92,127 (48.7) | 14,056(50.3) |  | 29,699 (54.1) | 11,144 (49.5) | 15,991 (52.8) | 208,352 (56.6) | 28,844 (56.4) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Age** |  |  |  |  |  |  |  |  |  |  |  |
|  Mean (sd) | 54.6 (14.0) | 56.3 (13.6) | 57.1 (13.3) | 67.3 (12.6) | 68.4 (14.3) |  | 60.0 (13.2) | 62.1 (13.8) | 61.5 (13.3) | 67.6 (12.4) | 68.5 (13.4) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Quintile of socio-economic deprivation**, n(%) |  |  |  |  |  |  |  |  |  |  |  |
|  1 – Least Deprived | 2,354 (8.6) | 291 (2.2) | 1,311 (8.6) | 38,950 (20.6) | 6,573 (23.5) |  | 4,902 (8.9) | 591 (2.6) | 2,909 (9.6) | 68,139 (18.5) | 10,927 (21.4) |
|  5 – Most Deprived | 8,478 (31.0) | 5,457 (40.8) | 4,330 (28.2) | 33,350 (17.6) | 3,997 (14.3) |  | 16,744 (30.5) | 9,519 (42.2) | 8,976 (29.6) | 78,777 (21.4) | 8,814 (17.2) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **BMI (kg/m2)** |  |  |  |  |  |  |  |  |  |  |  |
|  Mean (sd) | 28.2 (5.3) | 30.7 (6.2) | 29.5 (6.1) | 29.8 (6.3) | 29.2 (6.3) |  | 28.5 (5.3) | 30.6 (6.1) | 29.8 (6.1) | 31.5 (6.6) | 30.7 (6.6) |
|  BMI >30, n(%) | 8,256 (30.2) | 6,460 (48.3) | 6,046 (39.4) | 79,821 (42.2) | 10,046 (35.9) |  | 17,721 (32.2) | 10,793 (47.9) | 12,633 (41.7) | 200,101 (54.3) | 24,436 (47.8) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Smoking,** n(%) |  |  |  |  |  |  |  |  |  |  |  |
|  Never | 19,899 (72.7) | 8,547 (64.0) | 8,572 (55.9) | 66,756 (35.3) | 11,013 (39.4) |  | 36,620 (66.7) | 13,511 (60.0) | 15,555 (51.4) | 114,601 (31.1) | 17,395 (34.0) |
|  Ex | 4,647 (17.0) | 3,201 (24.0) | 4,391 (28.6) | 91,151 (48.2) | 12,481 (44.7) |  | 12,999 (23.7) | 6,764 (30.0) | 10,584 (34.9) | 197,724 (53.7) | 25,371 (49.6) |
|  Current | 2,795 (10.2) | 1,604 (12.0) | 2,358 (15.4) | 31,173 (16.5) | 4,323 (15.5) |  | 5,278 (9.6) | 2,246 (10.0) | 4,138 (13.7) | 55,823 (15.2) | 7,982 (15.6) |
|  |  |  |  |  |  |  |  |  |  |  |  |
| **Comorbidities**, n(%)**‡** |  |  |  |  |  |  |  |  |  |  |  |
|  0  | 16,183 (59.1) | 6,128 (45.9) | 7,722 (50.4) | 56,268 (29.7) | 8,292 (29.7) |  | 19,756 (36.0) | 6,367 (28.3) | 9,602 (31.7) | 78,098 (21.2) | 12,395 (24.2) |
|  1 to 2 | 10,163 (37.1) | 6,650 (49.8) | 6,857 (44.7) | 105,479 (55.8) | 15,169 (54.3) |  | 29,970 (54.6) | 13,772 (61.1) | 17,443 (57.6) | 219,638 (59.6) | 28,075 (54.9) |
|  >2 | 1,022 (3.7) | 588 (4.4) | 752 (4.9) | 27,448 (14.5) | 4,495 (16.1) |  | 5,187 (9.5) | 2,394 (10.6) | 3,244 (10.7) | 70,517 (19.2) | 10,693 (20.9) |

Footnote: Index of Multiple Deprivation was not available for 174 (0.1%) people with prediabetes and 377 (0.1%) with type 2 diabetes. BMI was not available for 6,221 (2.3%) of patients with prediabetes and 3,034 (0.6%) with type 2 diabetes. Smoking was not available for 305 (0.1%) people with prediabetes and 560 (0.1%) of those with type 2 diabetes.

† - Crude prevalence among all CPRD patients aged 18-90 actively registered on 1/1/2015 for at least one year. \* - Percentage of patients with prediabetes or type 2 diabetes who have been matched to at least patient without diabetes, ‡ - Count of the following: Atrial fibrillation, cancer, chronic obstructive pulmonary disease, coronary heart disease, chronic kidney disease, dementia, epilepsy, heart failure, hypertension, peripheral vascular disease, serious mental Illness (e.g. psychosis, schizophrenia & bipolar affective disorder), stroke/TIA.

### Table 2 – Infection rates and incidence rate ratios in patients with prediabetes, type 2 diabetes, and matched patients without prediabetes or diabetes, overall and stratified by age and sex

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Infection Outcome | Non-DM matched comparison\* | PreDM | Non-DM\* | Prediabetes (PreDM) vs. non-diabetes\* |  | T2DM | Non-DM\* | Type 2 Diabetes (T2DM) vs. non-diabetes\*  |
|  |  | Rate† | Rate† | IRR1‡ | 95% CI | IRR2‡ | 95% CI |  | Rate† | Rate† | IRR1‡ | 95% CI | IRR2‡ | 95% CI |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **Primary care** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - All | Age-sex-practice  | 194.6 | 144.7 | 1.34 | 1.33-1.35 | 1.25 | 1.24-1.26 |  | 215.3 | 145.0 | 1.50 | 1.49-1.51 | 1.33 | 1.32-1.34 |
| - All | Age-sex-ethnicity | 194.7 | 144.9 | 1.35 | 1.34-1.36 | 1.25 | 1.24-1.26 |  | 215.1 | 143.3 | 1.51 | 1.51-1.52 | 1.32 | 1.32-1.33 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - Females | Age-sex-ethnicity | 225.9 | 169.2 | 1.34 | 1.33-1.35 | 1.25 | 1.24-1.26 |  | 257.0 | 170.7 | 1.52 | 1.51-1.53 | 1.33 | 1.32-1.34 |
| - Males | Age-sex-ethnicity | 161.2 | 118.9 | 1.36 | 1.35-1.38 | 1.25 | 1.23-1.26 |  | 181.6 | 121.4 | 1.50 | 1.49-1.52 | 1.31 | 1.30-1.32 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - 18 to 50 years | Age-sex-ethnicity | 182.8 | 107.1 | 1.71 | 1.68-1.74 | 1.60 | 1.57-1.64 |  | 211.1 | 104.4 | 2.01 | 1.98-2.05 | 1.80 | 1.77-1.83 |
| - 51 to 70 years | Age-sex-ethnicity | 178.9 | 127.1 | 1.41 | 1.39-1.43 | 1.29 | 1.27.1.30 |  | 200.6 | 123.7 | 1.62 | 1.61-1.63 | 1.39 | 1.38-1.40 |
| - 71 to 90 years | Age-sex-ethnicity | 220.8 | 184.7 | 1.20 | 1.19-1.21 | 1.12 | 1.11-1.13 |  | 235.0 | 181.7 | 1.31 | 1.30-1.32 | 1.17 | 1.16-1.18 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Hospitalisations  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - All | Age-sex-practice  | 44.9 | 34.5 | 1.31 | 1.29-1.33 | 1.11 | 1.09-1.13 |  | 64.4 | 37.5 | 1.81 | 1.79-1.82 | 1.41 | 1.40-1.43 |
| - All | Age-sex-ethnicity | 44.9 | 34.5 | 1.33 | 1.31-1.35 | 1.11 | 1.09-1.13 |  | 64.4 | 35.8 | 1.91 | 1.90-1.93 | 1.45 | 1.43-1.46 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - Females | Age-sex-ethnicity | 45.8 | 34.4 | 1.35 | 1.32-1.38 | 1.14 | 1.11-1.16 |  | 67.6 | 35.8 | 2.02 | 1.99-2.05 | 1.50 | 1.48-1.52 |
| - Males | Age-sex-ethnicity | 44.1 | 34.6 | 1.30 | 1.27-1.33 | 1.08 | 1.06-1.11 |  | 61.9 | 35.7 | 1.83 | 1.80-1.85 | 1.40 | 1.38-1.42 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| - 18 to 50 years | Age-sex-ethnicity | 21.9 | 10.9 | 2.02 | 1.91-2.14 | 1.72 | 1.62-1.83 |  | 32.1 | 11.0 | 2.96 | 2.85-3.08 | 2.25 | 2.15-2.35 |
| - 51 to 70 years | Age-sex-ethnicity | 29.5 | 20.4 | 1.47 | 1.43-1.50 | 1.16 | 1.13-1.19 |  | 44.9 | 20.5 | 2.29 | 2.25-2.33 | 1.58 | 1.55-1.61 |
| - 71 to 90 years | Age-sex-ethnicity | 75.4 | 63.3 | 1.21 | 1.19-1.23 | 1.04 | 1.02-1.06 |  | 100.8 | 63.8 | 1.68 | 1.66-1.70 | 1.33 | 1.31-1.35 |

\* - Non-DM are patients without diabetes or prediabetes matched on age-sex-practice or age-sex-ethnicity. † - Crude infection rate per 1,000 per year. ‡ - Incidence rate ratio compared to non-diabetes. IRR1 is not adjusted (besides matching factors). IRR2 additionally adjusts for index of multiple deprivation, smoking and number of co-morbidities. Note: Prediabetes or type 2 diabetes patients are only included in analysis if they have a matched non-diabetes patient: prediabetes with age-sex-practice match N=274,917, Prediabetes with age-sex-ethnicity match N=275,408, type 2 diabetes with age-sex-practice match N=529,678, type 2 diabetes with age-sex-ethnicity match N=531,596.

## Figure Legends

### Figure 1 - Infection rates and incidence rate ratios in prediabetes, type 2 diabetes & non-diabetes patients, by ethnicity

IRR = Incidence Rate Ratios (with 95% confidence intervals), unadjusted.

Note: Non-diabetes are patients without diabetes or prediabetes matched on age, sex and ethnicity.

### Figure 2 – Incidence rate ratios for infections in prediabetes and type 2 diabetes patients vs non-diabetes patients, stratified by ethnicity and age

Orange symbols = Prediabetes, Blue symbols = Type 2 diabetes. Circles = Primary Care, Triangles = Hospitalisations.

IRR = Incidence Rate Ratios (with 95% confidence intervals), unadjusted.

Note: Non-diabetes are patients without diabetes or prediabetes matched on age, sex and ethnicity.

## References

1. Whicher CA, O’Neill S, Holt RIG. Diabetes in the UK: 2019. *Diabetic Medicine* 2020;37(2):242-47.

2. Holman N, Forouhi NG, Goyder E, Wild SH. The Association of Public Health Observatories (APHO) Diabetes Prevalence Model: estimates of total diabetes prevalence for England, 2010-2030. *Diabetic Medicine* 2011;28(5):575-82.

3. Holden SE, Jenkins-Jones S, Morgan CL, et al. Prevalence, glucose control and relative survival of people with Type 2 diabetes in the UK from 1991 to 2013. *Diabetic Medicine* 2017;34(6):770-80.

4. Mainous AG, Tanner RJ, Baker R, Zayas CE, Harle CA. Prevalence of prediabetes in England from 2003 to 2011: population-based, cross-sectional study. *Bmj Open* 2014;4(6).

5. Ravindrarajah R, Reeves D, Howarth E, et al. Epidemiology and determinants of non-diabetic hyperglycaemia and its conversion to type 2 diabetes mellitus, 2000-2015: cohort population study using UK electronic health records. *Bmj Open* 2020;10(9):10.

6. Pearson-Stuttard J, Bennett J, Cheng YJ, et al. Trends in predominant causes of death in individuals with and without diabetes in England from 2001 to 2018: an epidemiological analysis of linked primary care records. *The Lancet Diabetes & Endocrinology* 2021;9(3):165-73.

7. Tomic D, Shaw JE, Magliano DJ. The burden and risks of emerging complications of diabetes mellitus. *Nature Reviews Endocrinology* 2022;18(9):525-39.

8. Gregg EW, Li YF, Wang J, et al. Changes in Diabetes-Related Complications in the United States, 1990-2010. *New England Journal of Medicine* 2014;370(16):1514-23.

9. Shah BR, Hux JE. Quantifying the risk of infectious diseases for people with diabetes. *Diabetes Care* 2003;26(2):510-13.

10. McDonald HI, Thomas SL, Millett ERC, Nitsch D. CKD and the Risk of Acute, Community-Acquired Infections Among Older People With Diabetes Mellitus: A Retrospective Cohort Study Using Electronic Health Records. *Am J Kidney Dis* 2015;66(1):60-68.

11. Pearson-Stuttard J, Blundell S, Harris T, Cook DG, Critchley J. Diabetes and infection: assessing the association with glycaemic control in population-based studies. *Lancet Diabetes Endocrinol* 2016;4(2):148-58.

12. Carey IM, Critchley JA, DeWilde S, et al. Risk of Infection in Type 1 and Type 2 Diabetes Compared With the General Population: A Matched Cohort Study. *Diabetes Care* 2018;41(3):513-21.

13. Niedzwiedz CL, O’Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Medicine* 2020;18(1):160.

14. Fang M, Ishigami J, Echouffo-Tcheugui JB, et al. Diabetes and the risk of hospitalisation for infection: the Atherosclerosis Risk in Communities (ARIC) study. *Diabetologia* 2021;64(11):2458-65.

15. Goff LM. Ethnicity and Type 2 diabetes in the UK. *Diabetic Medicine* 2019;36(8):927-38.

16. Zhu Y, Sidell MA, Arterburn D, et al. Racial/Ethnic Disparities in the Prevalence of Diabetes and Prediabetes by BMI: Patient Outcomes Research To Advance Learning (PORTAL) Multisite Cohort of Adults in the U.S. *Diabetes Care* 2019;42(12):2211-19.

17. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol* 2019;48(6):1740-40g.

18. Carey IM, Cook DG, Harris T, et al. Risk factors for excess all-cause mortality during the first wave of the COVID-19 pandemic in England: A retrospective cohort study of primary care data. *PLoS One* 2021;16(12).

19. Padmanabhan S, Carty L, Cameron E, et al. Approach to record linkage of primary care data from Clinical Practice Research Datalink to other health-related patient data: overview and implications. *Eur J Epidemiol* 2019;34(1):91-99.

20. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). *Int J Epidemiol* 2017;46(4):1093-+.

21. GOV.UK. English indices of deprivation. 2020: <https://www.gov.uk/government/collections/english-indices-of-deprivation>. Accessed 06-01-2023.

22. Mahadevan P, Harley M, Fordyce S, et al. Completeness and representativeness of small area socioeconomic data linked with the UK Clinical Practice Research Datalink (CPRD). *J Epidemiol Community Health* 2022;76(10):880-86.

23. NHS Digital. Quality and Outcomes Framework. NHS Digital; 2022: <https://digital.nhs.uk/data-and-information/data-collections-and-data-sets/data-collections/quality-and-outcomes-framework-qof>. Accessed 13-04-2022.

24. Hanley JA. A heuristic approach to the formulas for population attributable fraction. *J Epidemiol Community Health* 2001;55(7):508-14.

25. Palladino R, Tabak AG, Khunti K, et al. Association between pre-diabetes and microvascular and macrovascular disease in newly diagnosed type 2 diabetes. *BMJ Open Diab Res Care* 2020;8(1):e001061.

26. International Diabetes Federation. IDF Diabetes Atlas 10th Edition 2021: <https://diabetesatlas.org/atlas/tenth-edition/>. Accessed 20-01-2023.

27. Wright AK, Welsh P, Gill JMR, et al. Age-, sex- and ethnicity-related differences in body weight, blood pressure, HbA(1c) and lipid levels at the diagnosis of type 2 diabetes relative to people without diabetes. *Diabetologia* 2020;63(8):1542-53.

28. McDonald HI, Nitsch D, Millett ERC, Sinclair A, Thomas SL. New estimates of the burden of acute community-acquired infections among older people with diabetes mellitus: a retrospective cohort study using linked electronic health records. *Diabetic Medicine* 2014;31(5):606-14.

29. Aldridge RW, Lewer D, Katikireddi SV, et al. Black, Asian and Minority Ethnic groups in England are at increased risk of death from COVID-19: indirect standardisation of NHS mortality data. *Wellcome Open Res* 2020;5(88):88.

30. Mathur R, Rentsch CT, Morton CE, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *The Lancet* 2021;397(10286):1711-24.

31. Haddad MB, Lash TL, Castro KG, et al. Tuberculosis Infection Among People With Diabetes: United States Population Differences by Race/Ethnicity. *Am J Prev Med* 2020;58(6):858-63.

32. Petersen J, Kandt J, Longley PA. Ethnic inequalities in hospital admissions in England: an observational study. *BMC Public Health* 2021;21(1):862.

33. Marmot M, Allen J, Boyce T, Goldblatt P, Morrison J. Health equity in England: the Marmot review 10 years on.: The Health Foundation; 2020: <https://www.health.org.uk/publications/reports/the-marmot-review-10-years-on>. Accessed 20-01-2023.

34. GOV.UK. Overcrowded households. 2020: <https://www.ethnicity-facts-figures.service.gov.uk/housing/housing-conditions/overcrowded-households/latest>. Accessed 28-11-2022.

35. Mathur R, Palla L, Farmer RE, Chaturvedi N, Smeeth L. Ethnic differences in the severity and clinical management of type 2 diabetes at time of diagnosis: A cohort study in the UK Clinical Practice Research Datalink. *Diabetes Res Clin Pract* 2020;160.

36. Type 2 diabetes in adults: management. NICE guideline [NG28]. National Institute for Health and Care Excellence; 2017: <https://www.nice.org.uk/guidance/NG28>. Accessed 30-06-2017.

37. Critchley JA, Carey IM, Harris T, et al. Glycemic Control and Risk of Infections Among People With Type 1 or Type 2 Diabetes in a Large Primary Care Cohort Study. *Diabetes Care* 2018;41(10):2127-35.

## Figure 1



## Figure 2

