OXFORD

Precision Clinical Medicine, 2022, **5**: pbac015

DOI: 10.1093/pcmedi/pbac015
Research Article

Health consequences of early-onset compared with late-onset type 2 diabetes mellitus

Can Hou[1,2], Huazhen Yang[1,2], Yuanyuan Qu[1,2], Wenwen Chen[1,2], Yu Zeng[1,2], Yao Hu[1,2], K.M. Venkat Narayan[3], Huan Song[1,2,4,*] and Dong Li[3,5,*]

[1]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu 610041, China
[2]Med-X Center for Informatics, Sichuan University, Chengdu 610211, China
[3]Emory Global Diabetes Research Center (EGDRC), Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA
[4]Centre of Public Health Sciences, Faculty of Medicine, University of Iceland, Reykjavík 101, Iceland
[5]Division of Hospital Medicine, School of Medicine, Emory University, Atlanta, GA 30322, USA
*Correspondence: Huan Song, songhuan@wchscu.cn; Dong Li, dong.li2@emory.edu

ABSTRACT

Background: Although cumulating evidence has suggested that early-onset type 2 diabetes mellitus (T2DM) conferred on patients a broader tendency for complications beyond vascular ones, a comprehensive analysis of patterns of complications across all relevant systems is currently lacking.

Method: We prospectively studied 1 777 early-onset (age at diagnosis ≤ 45 years) and 35 889 late-onset (>45 years) T2DM patients with matched unexposed individuals from the UK Biobank. Diabetes-specific and -related complications were examined using phenome-wide association analysis, with patterns identified by comorbidity network analysis. We also evaluated the effect of lifestyle modifications and glycemic control on complication development.

Results: The median follow-up times for early-onset and late-onset T2DM patients were 17.83 and 9.39 years, respectively. Compared to late-onset T2DM patients, patients with early-onset T2DM faced a significantly higher relative risk of developing subsequent complications that primarily affected sense organs [hazard ratio (HR) 3.46 vs. 1.72], the endocrine/metabolic system (HR 3.08 vs. 2.01), and the neurological system (HR 2.70 vs. 1.81). Despite large similarities in comorbidity patterns, a more complex and well-connected network was observed for early-onset T2DM. Furthermore, while patients with early-onset T2DM got fewer benefits (12.67% reduction in pooled HR for all studied complications) through fair glycemic control (median $HbA_{1c} \leq 53$ mmol/mol) compared to late-onset T2DM patients (18.01% reduction), they seemed to benefit more from favorable lifestyles, including weight control, healthy diet, and adequate physical activity.

Conclusions: Our analyses reveal that early-onset T2DM is an aggressive disease resulting in more complex complication networks than late-onset T2DM. Aggressive glucose-lowering intervention, complemented by lifestyle modifications, are feasible strategies for controlling early-onset T2DM-related complications.

Keywords: early-onset type 2 diabetes mellitus, phenome-wide association analysis, comorbidity network analysis, late-onset type 2 diabetes mellitus

Introduction

Early-onset type 2 diabetes mellitus (T2DM) is becoming a fundamental public health issue worldwide. The landmark SEARCH Study[1] reported that early-onset T2DM accounted for 13.7% of all people with diabetes in youth and recorded a 4.8% annual surge in incidence.[2] Early-onset T2DM appears to be a more aggressive disease by unclear mechanisms in terms of frequency of complications that are not commensurate merely with the age of onset.[3,4] Previous studies have primarily focused on the microvascular and macrovascular complications of early-onset T2DM.[5,6] However, recent data suggest that the complications are perhaps widely distributed throughout the human system,[7–10] even in the early stages of the disease, but the evidence-based data gap is the Achilles' heel of these studies.

A way to overcome this challenge is to perform phenome-wide association analyses (PheWAS) to identify the characteristics of subsequent complications, the essential prerequisite to developing effective interventions. Taking advantage of the rich information on demographics and lifestyles as well as the complete medical record in the UK Biobank, we investigated the health consequences of early-onset T2DM (age at diagnosis ≤ 45 years) compared to late-onset T2DM (age at diagnosis > 45 years) through systemically computing the phenotypic risk and patterns of morbidities. Furthermore, we explored whether the heightened risk of complications in early-onset T2DM may be modifiable by glycemic control or favorable lifestyles. To our knowledge, this is the first PheWAS study on T2DM, with potentially far-reaching aims that provide not only an overview of potential complications but also has implications for effective intervention strategies.

Methods

Study design and data sources

We analyzed data from the UK Biobank, a community-based prospective study described in detail elsewhere.[11] Briefly, the UK Biobank recruited 502 507 participants aged between

Received: January 28, 2022. **Accepted:** May 30, 2022. **Published:** 6 June 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of the West China School of Medicine & West China Hospital of Sichuan University. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

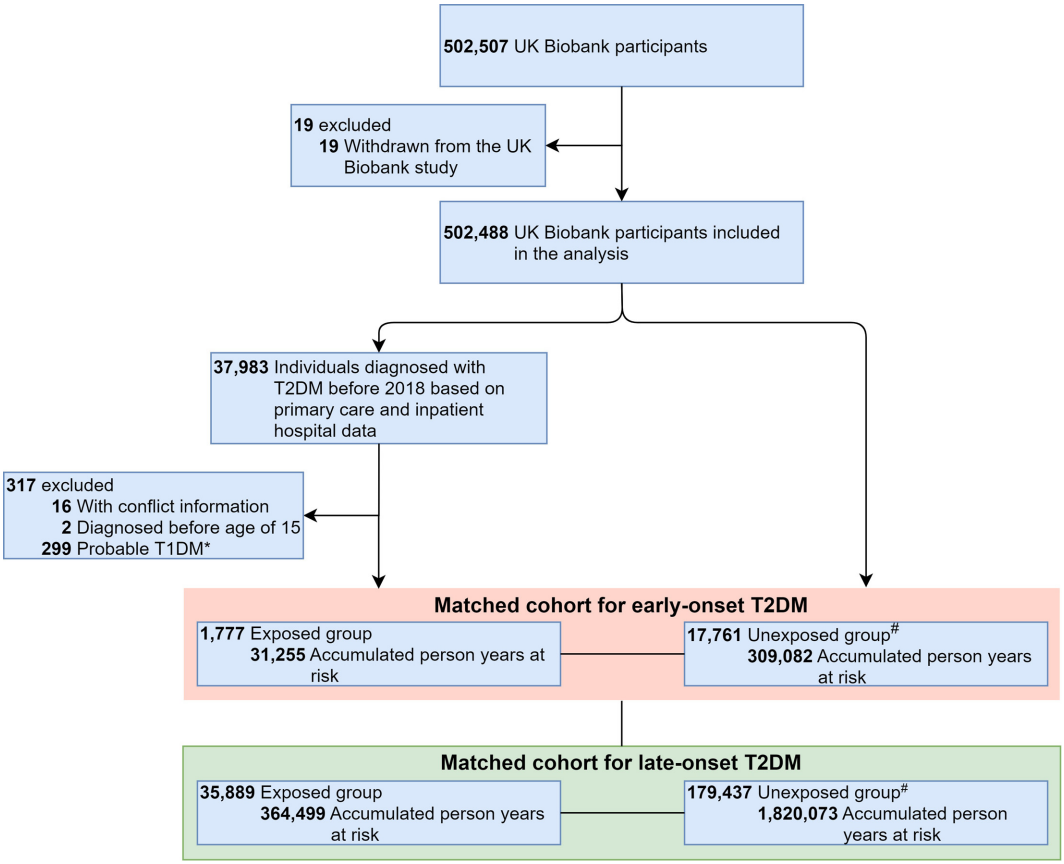


Figure 1. Flowchart for population selection. *The algorithm for identifying probable T1DM patients is shown in Supplementary Fig. 1. #For each T2DM patient, we randomly selected up to 5 (for late-onset T2DM patients) or 10 (for early-onset T2DM patients) individuals without T2DM diagnosis at the diagnosis date of the index patient (i.e. the index date) from the study population, individually matched by date of birth (±1 year), sex, and decile of Townsend deprivation index.

40–69 years between 2006 and 2010, with sociodemographic information, lifestyle factors and medical conditions collected at baseline. The National Health Service (NHS) Digital and the NHS Central Register data were combined to supplement mortality information. Inpatient hospital information from 1997 onwards was obtained from nationwide databases, including the Hospital Episode Statistics, the Patient Episode Database for Wales, and the Scottish Morbidity Record.[11] The primary care records as of 2017 were integrated from multiple general practice (GP) databases, covering 45% of the UK Biobank participants.[12] The UK Biobank has full ethical approval from the NHS National Research Ethics Service (reference number: 16/NW/0274), and the biomedical research ethics committee of West China Hospital approved the current study (reference number: 2019–1171).

Disease categories

Initially, 37 983 individuals were diagnosed with T2DM based on primary care and inpatient records (diagnosis codes are given in Supplementary Table 1, see online supplementary material). We excluded 299 suspected T1DM patients to reduce misdiagnosed or misclassification as T2DM,[13] according to a validated algorithm for T1DM identification in UK Biobank data (Supplementary Fig. 1, see online supplementary material).[14] We further excluded 18 cases with conflicting information or diagnosed before the age of 15, leaving 1 777 early-onset and 35 889 late-onset T2DM patients. We then randomly selected 5 or 10 age, sex, and Townsend deprivation index-matched unexposed individuals for each T2DM patient (Fig. 1). With satisfiable accuracy for T2DM diagnosis in the linked databases (82%–90%),[15–17] the UK Biobank database has been widely used for diabetes-related studies.[18, 19]

Identification of subsequent medical conditions was based on inpatient hospital data, with date of diagnosis recorded as the date of the first identified hospital episode. We restricted the original International Classification of Diseases-10th revision (ICD-10) codes to chapters 1–14 (i.e. excluding the codes related to pregnancy, perinatal conditions, symptoms or signs, or external causes of morbidity and mortality). The ICD-10 codes were mapped to 'phecodes', a coding system considered more relevant to medical conditions mentioned in clinical settings.[20] To sustain sufficient data power, we only used the first level 'phecodes' (3-character 'phecodes'), resulting in 413 specific medical conditions in the analysis (Supplementary Table 2, see online supplementary material). These medical conditions were summarized as diabetes-specific complications (i.e. microvascular complications and macrovascular complications, according to a previous study[21]) and complications related to diabetes, which were further summarized as 13 broad categories based on the mainly affected systems. The diagnoses of medical conditions in inpatient data in the UK have been validated, demonstrating an overall high diagnostic accuracy.[22]

Sociodemographic characteristics and lifestyle factors, including birth, sex, household income, family history, smoking status, drinking status, and fruit/vegetable consumption, were collected at baseline. The anthropometric data used for body mass index (BMI) calculations were measured during the baseline visit.

Each participant was assigned a Townsend deprivation index based on the postal codes provided at recruitment as an index of population-level deprivation.[23] For individuals who completed the physical activity questionnaire, total physical activity was calculated by summing the metabolic equivalent task (MET)-weighted time spent in vigorous, moderate, and walking activity.[24] We further extracted HbA_{1c} (Hemoglobin A_{1c}) test results recorded both at baseline assessment and in the primary care database (Supplementary Table 3, see online supplementary material). We defined 'fair glycemic control' as a median $HbA_{1c} \leq 53$ mmol/mol,[25] calculated based on all available HbA_{1c} tests conducted at least 6 months after the T2DM diagnosis.

Statistical analysis

PheWAS for identifying T2DM-associated complications

Hazard ratio (HR) of diseases per types of complications (i.e. microvascular complications, macrovascular complications, or complications related to diabetes) was first assessed using Cox models, which were stratified by matching identifier (i.e. date of birth, sex, and Townsend deprivation index) and partially or fully adjusted for household income, BMI, smoking, and alcohol-drinking status. PheWAS for specific medical conditions with a prevalence ≥1% were then performed using fully adjusted Cox models. Specifically, we removed individuals with a history of all related medical conditions indicated in the previous study during the analysis of each medical condition of interest (Supplementary Table 2).[20] Medical conditions with P-value < the Bonferroni corrected threshold and HR > 1.0 were considered complications significantly associated with the prior T2DM diagnosis. To quantify the overall magnitude difference between PheWAS results of the matched cohort for early-onset T2DM and those of late-onset T2DM, we calculated the changed percentages of pooled HRs obtained using the inverse-variance weighting method (i.e. $(\frac{HR^{pooled}_{early-onset}}{HR^{pooled}_{typical}} - 1) \times 100\%$).

Comorbidity network analyses

The comorbidity network was constructed to investigate the diversity of disease clusters associated with a prior diagnosis of early-onset and late-onset T2DM.[26] All possible pairs were compiled using identified complications in the PheWAS step. We estimated relative risk (RR) and ϕ-correlation as the measures of strength of comorbidity associations for pairs with a prevalence ≥1.0%. The comorbidity networks were then constructed based on the pairs of complications with RR ≥ 2.0 and ϕ-correlation ≥ 0.1. Next, to clarify the comorbidity patterns, the networks were partitioned into groups with high intrinsic connectivity (i.e. modules), using the community detection algorithm Louvain.[27]

Sub-analyses and sensitivity analyses

We performed sub-analyses to evaluate the prioritized blood glucose control strategies on subsequent complications for early-onset and late-onset T2DM patients. Likewise, we explored the potential modifying effect of lifestyle factors by performing sub-analyses for T2DM patients with adequate or inadequate fruit/vegetable consumption,[28] for those with BMI ≥ 29.9 or < 29.9, and for those with total physical activity ≤ 798 or > 798 MET.min/week. Furthermore, we also repeated the sub-analyses for different lifestyle factors among T2DM patients with poor glycemic control (n = 1 089 and 8 901 for early-onset and late-onset T2DM, respectively).

In sensitivity analyses, to test the robustness of the main results with the definition of T2DM, we repeated the PheWAS after removing T2DM patients with any documented diagnosis of other types of diabetes (codes listed in Supplementary Table 3), and those with only one documented diagnosis of T2DM. Additionally, with a concern that the observed difference in the RR of subsequent complications after early-onset and late-onset T2DM was mainly attributed to longer disease duration (or more extended surveillance period) for early-onset than late-onset T2DM patients, we further re-assessed these estimates by restricting to the first 10 years of follow-up. Finally, to relieve the concern that the observed results could be largely influenced by systematic difference of comorbidity status between early-onset and late-onset T2DM patients, we also re-assessed these estimates after additionally excluding T2DM patients with a history of hypertension, disorders of lipoid metabolism, overweight, obesity and other hyperalimentation, chronic kidney disease, or cardiovascular disease at the index date, along with their matched controls.

All the statistical analyses were conducted using SciPy (version 1.4.1), Statsmodels (version 0.11.1), and Lifelines (version 0.25.2) in Python 3.8.

Results

The matched cohort for early-onset T2DM included 1 777 exposed patients and 17 761 matched unexposed individuals. In the matched cohort for late-onset T2DM, the numbers of exposed patients and matched unexposed individuals were 35 889 and 179 437, respectively (Fig. 1 and Table 1). In addition to an equal sex distribution (60.32% *vs.* 59.54%), better socioeconomic conditions (−1.26 *vs.* 0.95 for median Townsend deprivation index), and more smokers (55.12% *vs.* 42.26%) and drinkers (91.05% *vs.* 83.57%) were observed in typical, compared with early-onset T2DM patients. A higher percentage of obesity (61.34% *vs.* 56.86% for BMI ≥ 29.9), low physical activity (29.77% *vs.* 26.60%), and poor glycemic control (61.28% *vs.* 24.80%) were found in early-onset T2DM patients.

Results of PheWAS

The median follow-up times for early-onset and late-onset T2DM patients were 17.83 and 9.39 years, respectively. Except for diabetes-specific complications, early-onset T2DM was also significantly associated with increased risk of subsequent diabetes-related complications in multiple systems of the body (Fig. 2 and Supplementary Table 4, see online supplementary material). Compared with typical T2DM, the top three involved high-risks systems in order were sense organs (HR: 2.86 *vs.* 1.59), endocrine/metabolic (HR: 3.09 *vs.* 2.01), and respiratory system (HR: 1.73 *vs.* 1.49), respectively. Among 413 medical conditions widely distributed in all systems of the body, a total of 181 and 142 passed the 1% prevalence threshold. PheWAS indicated up to 101 diabetes-specific or -related complications of various systems among early-onset T2DM patients and 129 among late-onset T2DM patients (Supplementary Table 5, see online supplementary material), although most of the complications were significantly associated with both types of T2DM (Fig. 3). The overall difference in association magnitude (i.e. between matched cohorts for early-onset and late-onset T2DM) was 44.36%. Notably, we observed substantial differences for complications specific to diabetes (overall magnitude difference: 131.80% for microvascular and 125.42% for macrovascular complications; see HRs for specific complications in Supplementary Table 5).

Table 1. Basic characteristics of the individuals.The values are reported as median (lower quantile–upper quantile) for continues variables or number (%) for categorical variables.

Characteristics	Matched cohort for early-onset T2DM		Matched cohort for late-onset T2DM	
	Diabetes patients (n = 1 777)	Matched individuals[b] (n = 17 761)	Diabetes patients (n = 35 889)	Matched individuals[b] (n = 179 437)
Age at index date, years	42.76 (40.51–44.41)	42.82 (40.50–44.49)	62.02 (56.12–67.52)	62.01 (56.08–67.50)
Townsend deprivation index[a]	0.95 (−1.96–3.62)	0.80 (−2.06–3.25)	−1.26 (−3.17–2.01)	−1.26 (−3.18–1.86)
Follow-up time, years	17.83 (13.89–22.06)	17.74 (13.55–22.29)	9.39 (5.90–13.76)	9.27 (5.92–13.66)
Sex				
Female	719 (40.46%)	7 190 (40.48%)	14 241 (39.68%)	71 203 (39.68%)
Male	1 058 (59.54%)	10 571 (59.52%)	21 648 (60.32%)	108 234 (60.32%)
Body mass index (kg/m^2)				
<24.1	119 (6.70%)	4 460 (25.11%)	2 105 (5.87%)	38 954 (21.71%)
24.1–29.9	516 (29.04%)	8 597 (48.40%)	12 913 (35.98%)	94 368 (52.59%)
≥29.9	1 090 (61.34%)	4 539 (25.56%)	20 408 (56.86%)	44 927 (25.04%)
Unknown	52 (2.93%)	165 (0.93%)	463 (1.29%)	1188 (0.66%)
Smoking status				
Ever	751 (42.26%)	7 976 (44.91%)	19 781 (55.12%)	89 724 (50.00%)
Never	998 (56.16%)	9 672 (54.46%)	15 714 (43.79%)	88 492 (49.32%)
Unknown	28 (1.58%)	113 (0.64%)	394 (1.10%)	1 221 (0.68%)
Drinking status				
Ever	1 485 (83.57%)	16 688 (93.96%)	32 677 (91.05%)	171 408 (95.53%)
Never	272 (15.31%)	968 (5.45%)	2 973 (8.28%)	7 354 (4.10%)
Unknown	20 (1.13%)	105 (0.59%)	239 (0.67%)	675 (0.38%)
Total physical activity[c]				
Low	529 (29.77%)	3 562 (20.06%)	9 547 (26.60%)	34 692 (19.33%)
Normal or high	820 (46.15%)	10 873 (61.22%)	17 604 (49.05%)	108 265 (60.34%)
Unknown	428 (24.09%)	3 326 (18.73%)	8 738 (24.35%)	36 480 (20.33%)
Fruit/vegetable consumption[d]				
Inadequate	1 206 (67.87%)	13 030 (73.36%)	24 071 (67.07%)	121 620 (67.78%)
Adequate	546 (30.73%)	4 619 (26.01%)	11 574 (32.25%)	57 132 (31.84%)
Unknown	25 (1.41%)	112 (0.63%)	244 (0.68%)	685 (0.38%)
Household income (£)				
<18 000	539 (30.33%)	3 255 (18.33%)	11 199 (31.20%)	43 494 (24.24%)
18 000–52 000	666 (37.48%)	7 644 (43.04%)	13 945 (38.86%)	77 070 (42.95%)
>52 000	234 (13.17%)	4 612 (25.97%)	3 779 (10.53%)	29 796 (16.61%)
Unknown	338 (19.02%)	2 250 (12.67%)	6 966 (19.41%)	29 077 (16.20%)
Glycemic control[e]				
Poor	1 089 (61.28%)	NA[f]	8 901 (24.80%)	NA
Good	545 (30.67%)		11 622 (32.38%)	
Unknown	143 (8.05%)		15 366 (42.82%)	

[a]Townsend deprivation index was assigned to each individual based on their postcode location; a greater index score implies a greater degree of deprivation.
[b]Up to 5 (for early-onset T2DM) or 10 (for late-onset T2DM) individuals free of T2DM diagnosis at the index date were randomly selected and individually matched to each patient by date of birth (within 1 year), sex, and decile of Townsend deprivation index.
[c]Total physical activity amount was calculated by summing the MET weighted time spent in vigorous, moderate, and walking activity, while low physical activity was defined as total physical activity amount ≤798 MET.min/week.
[d]Inadequate fruit/vegetable consumption was defined as eating <5 portions of fruit and vegetables per day.
[e]HbA_{1c} levels were retrieved from UK Biobank baseline assessment and primary care data (using codes listed in Supplementary Table 3), and only HbA_{1c} measurements conducted at least 6 months after T2DM diagnosis were used to calculate each T2DM patient's median HbA_{1c} level; T2DM patients were further divided into good glycemic control (≤53 mmol/mol) group and poor glycemic control (>53 mmol/mol) group according to their median HbA_{1c} levels.
[f]NA: Not available

Comorbidity network diversity

Among 9 870 possible comorbidity pairs, 1 025 and 479 were involved in the comorbidity network for early-onset and late-onset T2DM, respectively, after filtering based on prevalence, RR, and ϕ-correlation (Fig. 4). With 28 additional nodes and 605 additional links exclusive to early-onset T2DM (Supplementary Fig. 2, see online supplementary material), the comorbidity network for early-onset T2DM was more complex and tightly connected than that for late-onset T2DM. The modularity analysis indicated that the two networks shared similar comorbidity patterns, consisting of modules predominated by macrovascular complications (Fig. 4, dark yellow nodes), musculoskeletal (pink nodes), and digestive (brown nodes). However, early-onset T2DM was associated with increased interlinks and complications in each module and the presence of converged modules [modules predominated by genitourinary (light brown nodes), and microvascular complications (light pink nodes)].

Sub-analyses and sensitivity analyses

Analysis by glycemic control condition showed significant reductions in risk of subsequent complications that were associated with fair glycemic control in both matched cohorts. As such the reduction was more pronounced for late-onset (pooled HRs 1.98–1.62, 18.01% reduction) than early-onset (pooled HRs 2.67–2.33, 12.67% reduction) T2DM, and a bigger magnitude difference was seen in the fair glycemic control group than in the poor glycemic control group (43.47% vs. 34.71%, Supplementary Fig. 3, see online supplementary material). Furthermore, the sub-analyses

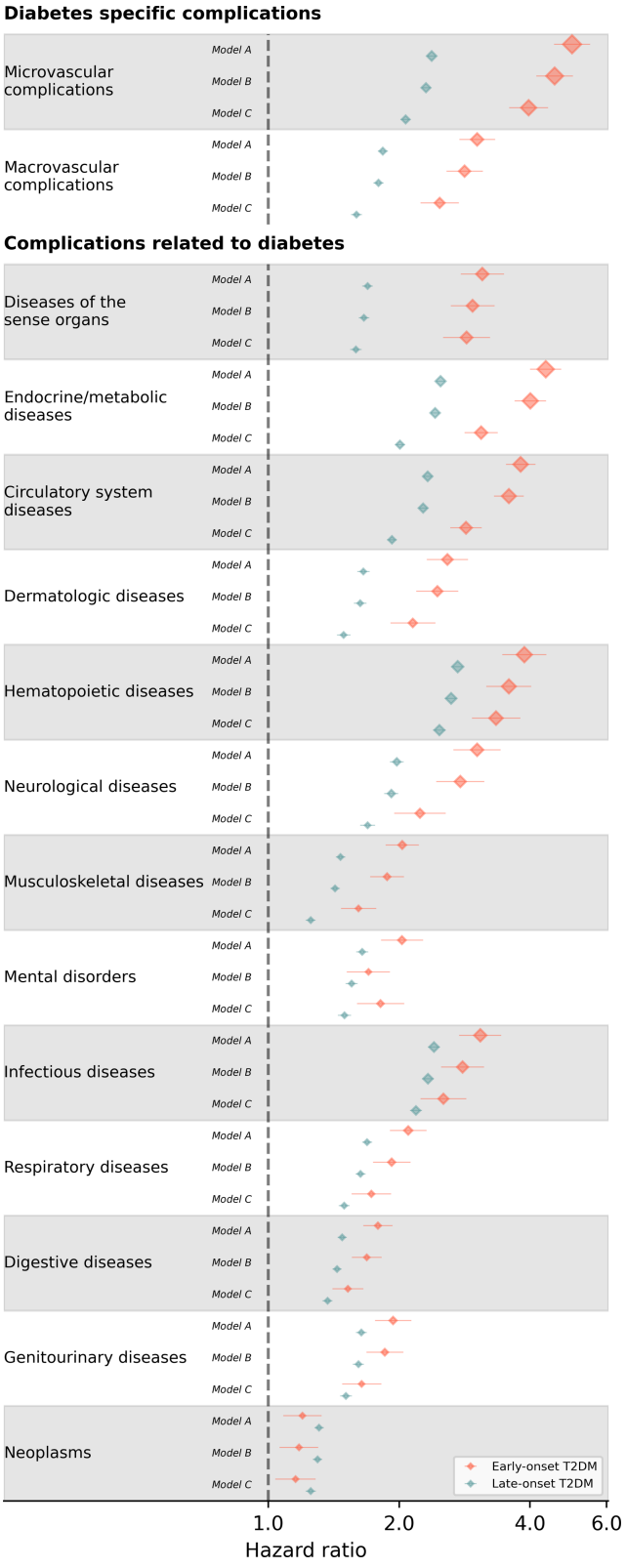


Figure 2. Risks of complications after T2DM diagnosis according to category. Estimates of HRs with confidence intervals were calculated from different conditional Cox models. Cox models were stratified by matching identifier (i.e. date of birth, sex, and Townsend deprivation index) (model A) and further adjusted for household income (model B) or household income, BMI, smoking and drinking status (model C). Detailed results are shown in Supplementary Table 4.

revealed significant effects of modifiable lifestyle factors (i.e. weight control, healthy diet, and physical activity; Supplementary Fig. 4, see online supplementary material). In contrast to the glycemic control condition, with generally greater HR reductions observed for early-onset than late-onset T2DM, the magnitude differences were consistently more minor in healthier lifestyle groups. The beneficial effects of these modifiable lifestyle factors were consolidated in the sub-analysis of T2DM patients with poor glycemic control only, with consistently higher HR reductions observed for early-onset compared with late-onset T2DM (Supplementary Fig. 5, see online supplementary material).

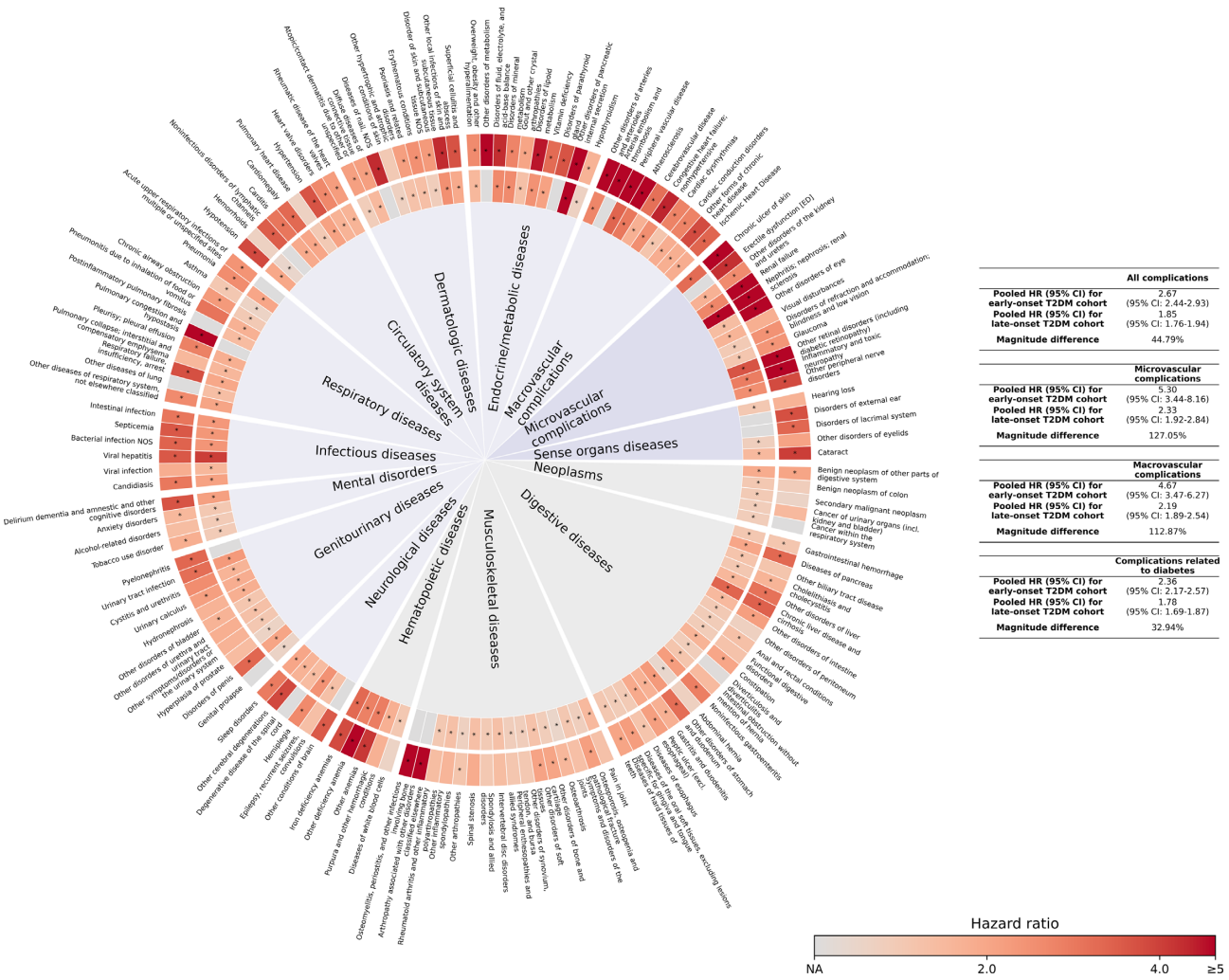
In the sensitivity analyses, either applying a stricter criterion for ascertaining T2DM diagnosis, restricting the maximum follow-up time to 10 years, or additionally excluding T2DM patients with comorbidity changed the results of PheWAS substantially (Supplementary Fig. 6, see online supplementary material). However, the magnitude difference was slightly reduced in the first sensitivity analysis (from 44.36% to 35.02%), while being slightly increased in the last one (from 44.36% to 56.43%). Due to a significant reduction in sample size in the sub-analyses and sensitivity analyses, *P*-values of many complications failed to reach the Bonferroni corrected threshold.

Discussion

The soaring prevalence of early-onset T2DM among adolescents and young adults has become a global public health threat.[29, 30] Fully understanding this trend, and systematic identification of the characteristics of subsequent complications, is crucial for developing effective interventions. In this large, matched cohort study based on the UK Biobank, in contrast to the traditional view that early-onset T2DM mainly involves macrovascular and microvascular complications, our study initially found that the complications of early-onset T2DM extensively involved a wide range of body systems even at the very early stage of the disease, although there were shared comorbidity patterns with late-onset T2DM. It suggested that a more aggressive intervention strategy for individuals at the early stage of T2DM needs to be urgently considered to reduce complications.

Our phenome-wide based analyses added systematic and comprehensive phenotypic data on all diabetes-related complications. We revealed the potential impact of a younger age of onset on subsequent complications with the absence of other traditional risk factors. Besides replicating previous findings that compared with late-onset T2DM patients, early-onset T2DM patients were found to be at higher risk for subsequent microvascular and macrovascular complications, including cardiovascular diseases, nephropathy, and retinopathy.[31–34] Our study further showed that, the RRs for other diabetes-related complications, such as delirium, dementia, other cerebral degenerations, anemias, osteoarthrosis, and noninfectious gastroenteritis, were also higher among early-onset T2DM patients, although the excess risks were smaller compared with those of microvascular and macrovascular complications.

Tight blood glucose control is the principle recommendation for late-onset T2DM to prevent and manage diabetes-related complications. Unexpectedly, our data showed that fair glycemic control might not achieve sufficient effectiveness in reducing the risk of subsequent complications in the early-onset T2DM population compared with the late-onset T2DM population, due to an apparently lower impact on diabetes-related complications. On the contrary, average weight or effective weight control (BMI controlled within 29.9), sufficient physical activity, and



	All complications
Pooled HR (95% CI) for early-onset T2DM cohort	2.67 (95% CI: 2.44-2.93)
Pooled HR (95% CI) for late-onset T2DM cohort	1.85 (95% CI: 1.76-1.94)
Magnitude difference	44.79%

	Microvascular complications
Pooled HR (95% CI) for early-onset T2DM cohort	5.30 (95% CI: 3.44-8.16)
Pooled HR (95% CI) for late-onset T2DM cohort	2.33 (95% CI: 1.92-2.84)
Magnitude difference	127.05%

	Macrovascular complications
Pooled HR (95% CI) for early-onset T2DM cohort	4.67 (95% CI: 3.47-6.27)
Pooled HR (95% CI) for late-onset T2DM cohort	2.19 (95% CI: 1.89-2.54)
Magnitude difference	112.87%

	Complications related to diabetes
Pooled HR (95% CI) for early-onset T2DM cohort	2.36 (95% CI: 2.17-2.57)
Pooled HR (95% CI) for late-onset T2DM cohort	1.78 (95% CI: 1.69-1.87)
Magnitude difference	32.94%

Figure 3. Risks of complications after T2DM diagnosis. The outer and inner rings show the point estimates of HRs of the corresponding complications for early-onset and late-onset T2DM respectively. Cox models were stratified by matching identifier (i.e. date of birth, sex, and Townsend deprivation index) and adjusted for household income, BMI, smoking and drinking status. Complications with P-value < Bonferroni corrected threshold are marked with an asterisk. Detailed results are shown in Supplementary Table 5.

adequate fruit/vegetable intake were significantly associated with a decreased risk of subsequent complications among early-onset T2DM patients, and were at comparable levels to those for late-onset T2DM. Therefore, we initially advocated a more aggressive glycemic control strategy for early-onset T2DM. Our data also validated that remission of early-onset T2DM is possible through weight management. Notably, due to the high prevalence of obesity in adolescents and young adults, the incidence of early-onset T2DM is expected to increase dramatically in the coming decades. Therefore, clinical guidelines focusing specifically on the treatment of early-onset T2DM and prevention of its complications are urgently needed.

The strengths of the current study include the use of a large community-based cohort to construct matched cohorts for early-onset and late-onset T2DM separately and long-term follow-up for all the subsequent complications through complete linkage to inpatient data for the UK Biobank participants since 1997. In addition, the availability of enriched data on sociodemographic information, lifestyle factors, and GP visits enabled consideration of many vital confounders in the model and further sub-analyses by modifiable lifestyles. Nevertheless, caution is needed in interpreting our results. The primary concern of the present study was that the observed difference in the RR of subsequent complications after early-onset and late-onset T2DM could be mainly attributed to poor glycemic control status or longer follow-up time in early-onset T2DM patients. To relieve such a concern, we conducted the sub-analysis among T2DM patients with the same glycemic control status and sensitivity analysis, restricting the follow-up period to the first 10 years, and the results were generally comparable with the main results. Another concern was that the individuals included in the exposed groups could be patients with other types of DM that were misclassified as T2DM, as the ascertainment of T2DM was based solely on register data. To address this concern, we used a validated algorithm to exclude probable T1DM patients. In the sensitivity analysis, we further excluded T2DM patients with any documented diagnosis of other DM types or with only one documented diagnosis of T2DM, and such alteration did not change the main results substantially. Although many important confounders, such as sociodemographic factors, lifestyle factors, and comorbidity have been considered in the analyses, residual confounding due to undiagnosed comorbidities or other unmeasured confounders could have biased the study's results to some extent. Meanwhile, as the primary care data only covered half of the UK Biobank participants, and inpatient data were only

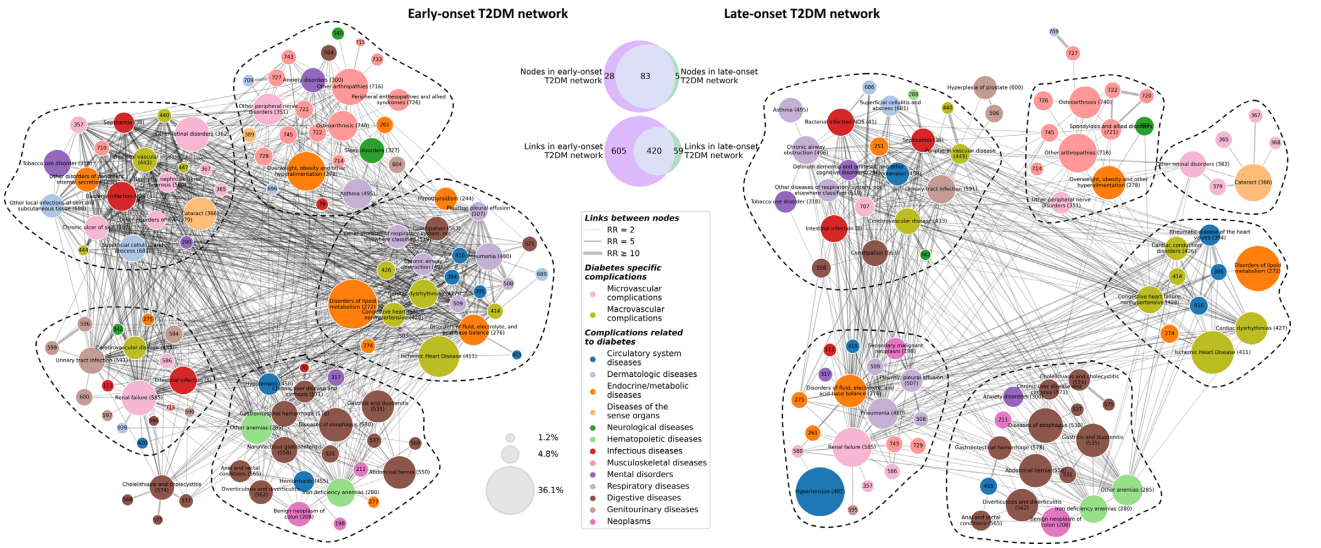


Figure 4. Comorbidity networks for early-onset and late-onset T2DM. The number of nodes and links in each network and their logical relation are shown in the Venn diagrams in the upper center of the figure. Each node represents a diabetes-associated complication, while the size and color of the node indicate the prevalence and the category of the corresponding complication (see legend). Reference nodes with percentages representing the prevalence scales are shown in the lower center of the figure. The width of the link represents the strength of comorbidity association, measured by RR. Reference links with RR are shown in the legend. Complications with prevalence ≥5% are labeled with its phenotype name along with the 'phecode', while others are labeled only with the 'phecode' (see Supplementary Table 2 for a description of each 'phecode'). The two networks were partitioned into several modules using the Louvain algorithm, and nodes belonging to the same module are grouped together and separated from other nodes using dashed lines.

available from 1997 onwards, the identified time of first T2DM diagnosis could be later than the actual time of onset, resulting in some patients with early-onset T2DM being misclassified as late-onset T2DM patients. Other limitations of the current study include that some important variables used for matching or model adjustment (e.g. Townsend deprivation index, smoking, and drinking status) were measured at the UK Biobank study baseline and may not reflect the individuals' status at the same time as the time of T2DM diagnosis. Additionally, the small number of early-onset T2DM patients leads to a limited statistical power to detect a significantly elevated HR for some of the studied complications in the early-onset cohort, especially in the further sub-analyses. Finally, the generalizability of the study results to a broader population was limited by the inherent flaws of the UK Biobank study, including low response rate and oversampling of the White population.[35]

Taken as a whole, we systemically reviewed the health consequences of early-onset T2DM and compared the results with those of late-onset T2DM, based on matched cohorts in the UK Biobank. Our data reveal that early-onset T2DM is more challenging in terms of glycemic control, with a higher risk of severe complications. Therefore, we advocate that guidelines for the intensive management of complications and glycemic control criteria are urgently needed.

Supplementary data

Supplementary materials are available at *PCMEDI Journal* online.

Acknowledgements

This work is supported by 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (Grant No. ZYYC21005 to HS), and by the National Natural Science Foundation of China (Grant No. 81971262 to H.S.). We express our sincere thanks to all the sponsors and team members involved in West China Biomedical Big Data Center and Med-X Center for Informatics, Sichuan University. Part of this research was done using the UK Biobank Resource under Application 54803.

Author contributions

H.S. and D.L. are responsible for the study's concept and design. W.C., Y.Z., and Y.H. carried out data and project management. C.H., H.Y., and Y.Q. performed data cleaning and analysis. C.H., H.Y., Y.Q., H.S., and D.L. interpreted the data. C.H., K.M.V.N., H.S., and D.L. drafted the manuscript. All the authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflict of interest

None declared.

References

1. Hamman RF, Bell RA, Dabelea D, *et al.* The search for diabetes in youth study: rationale, findings, and future directions. *Diabetes Care* 2014;**37**:3336–44. doi: 10.2337/dc14-0574.
2. Divers J, Mayer-Davis EJ, Lawrence JM, *et al.* Trends in incidence of type 1 and type 2 diabetes among youths - selected counties and Indian reservations, United States, 2002-2015. *MMWR Morbidity and Mortality Weekly Report* 2020;**69**:161–5. doi: 10.15585/mmwr.mm6906a3.
3. Twig G, Zucker I, Afek A, *et al.* Adolescent obesity and early-onset type 2 diabetes. *Diabetes Care* 2020;**43**:1487–95. doi: 10.2337/dc19-1988.
4. Huang JX, Liao YF, Li YM. Clinical features and microvascular complications risk factors of early-onset type 2 diabetes

Mellitus. *Current Medical Science* 2019;**39**:754–8. doi: 10.1007/s11596-019-2102-7.
5. Song SH. Early-onset type 2 diabetes: high lifetime risk for cardiovascular disease. *The Lancet Diabetes & Endocrinology* 2016;**4**:87–8. doi: 10.1016/S2213-8587(15)00390-3.
6. Huo X, Gao L, Guo L, *et al.* Risk of non-fatal cardiovascular diseases in early-onset versus late-onset type 2 diabetes in China: a cross-sectional study. *The Lancet Diabetes & Endocrinology* 2016;**4**:115–24. doi: 10.1016/S2213-8587(15)00508-2.
7. Rempel JD, Packiasamy J, Dean HJ, *et al.* Preliminary analysis of immune activation in early onset type 2 diabetes. *Int J Circumpolar Health* 2013;**72**:21190. doi: 10.3402/ijch.v72i0.21190.
8. Dibato JE, Montvida O, Zaccardi F, *et al.* Association of cardiometabolic multimorbidity and depression with cardiovascular events in early-onset adult type 2 diabetes: a multiethnic study in the U.S. *Diabetes Care* 2021;**44**:231–9. doi: 10.2337/dc20-2045.
9. Song SH. Complication characteristics between young-onset type 2 versus type 1 diabetes in a UK population. *BMJ Open Diabetes Research & Care*. 2015;**3**:e000044. doi: 10.1136/bmjdrc-2014-000044.
10. Barbiellini Amidei C, Fayosse A, Dumurgier J, *et al.* Association between age at diabetes onset and subsequent risk of dementia. *JAMA* 2021;**325**:1640–9. doi: 10.1001/jama.2021.4001.
11. Sudlow C, Gallacher J, Allen N, *et al.* UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015;**12**:e1001779. doi: 10.1371/journal.pmed.1001779.
12. UK Biobank. Primary care linked data 2020 [Available from: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary_care_data.pdf.
13. de Lusignan S, Khunti K, Belsey J, *et al.* A method of identifying and correcting miscoding, misclassification and misdiagnosis in diabetes: a pilot and validation study of routinely collected data. *Diabet Med* 2010;**27**:203–9. doi: 10.1111/j.1464-5491.2009.02917.x.
14. Eastwood SV, Mathur R, Atkinson M, *et al.* Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS One* 2016;**11**:e0162388. doi: 10.1371/journal.pone.0162388.
15. Van Staa TP, Abenhaim LJP. The quality of information recorded on a UK database of primary care records: a study of hospitalizations due to hypoglycemia and other conditions. *Pharmacoepidemiology & Drug Safety* 1994;**3**:15–21. doi: 10.1002/pds.2630030106.
16. Williams DR, Fuller JH, Stevens LK. Validity of routinely collected hospital admissions data on diabetes. *Diabet Med* 1989;**6**:320–4. doi: 10.1111/j.1464-5491.1989.tb01172.x.
17. Anwar H, Fischbacher CM, Leese GP, *et al.* Assessment of the under-reporting of diabetes in hospital admission data: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabet Med* 2011;**28**:1514–9. doi: 10.1111/j.1464-5491.2011.03432.x.
18. Thomas NJ, Jones SE, Weedon MN, *et al.* Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *The Lancet Diabetes & Endocrinology* 2018;**6**:122–9. doi: 10.1016/S2213-8587(17)30362-5.
19. Said MA, Verweij N, van der Harst P. Associations of combined genetic and lifestyle risks with incident cardiovascular disease and diabetes in the UK Biobank study. *JAMA cardiology* 2018;**3**:693–702. doi: 10.1001/jamacardio.2018.1717.
20. Wu P, Gifford A, Meng X, Li X, *et al.* Mapping ICD-10 and ICD-10-CM Codes to phecodes: workflow development and initial evaluation. *JMIR medical informatics* 2019;**7**:e14325. doi: 10.2196/14325.
21. Chen HY, Kuo S, Su PF, Wu JS, *et al.* Health care costs associated with macrovascular, microvascular, and metabolic complications of type 2 diabetes across time: estimates from a population-based cohort of more than 0.8 million individuals with up to 15 years of follow-up. *Diabetes Care* 2020;**43**:1732–40. doi: 10.2337/dc20-0072.
22. Burns EM, Rigby E, Mamidanna R, Bottle A, *et al.* Systematic review of discharge coding accuracy. *J Public Health* 2012;**34**:138–48. doi: 10.1093/pubmed/fdr054.
23. Townsend P, Phillimore P, Beattie A. Health and deprivation: inequality and the North: Routledge; 1988.
24. Cassidy S, Chau JY, Catt M, *et al.* Cross-sectional study of diet, physical activity, television viewing and sleep duration in 233,110 adults from the UK Biobank; the behavioural phenotype of cardiovascular disease and type 2 diabetes. *BMJ Open* 2016;**6**:e010038. doi: 10.1136/bmjopen-2015-010038.
25. Arslanian S, Bacha F, Grey M, *et al.* Evaluation and management of youth-onset type 2 diabetes: a position statement by the American Diabetes Association. *Diabetes Care* 2018;**41**:2648–68. doi: 10.2337/dci18-0052.
26. Hidalgo CA, Blumm N, Barabasi AL. Christakis NA. A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 2009;**5**:e1000353. doi: 10.1371/journal.pcbi.1000353.
27. De Meo P, Ferrara E, Fiumara G, *et al.* Generalized louvain method for community detection in large networks. *2011 11th international conference on intelligent systems design and applications*; 2011: IEEE. Córdoba, Spain.
28. Nelson KM, Reiber G, Boyko EJ, *et al.* Diet and exercise among adults with type 2 diabetes: findings from the third national health and nutrition examination survey (NHANES III). *Diabetes Care* 2002;**25**:1722–8. doi: 10.2337/diacare.25.10.1722.
29. Zhang Y, Luk AOY, Chow E, *et al.* High risk of conversion to diabetes in first-degree relatives of individuals with young-onset type 2 diabetes: a 12-year follow-up analysis. *Diabet Med* 2017;**34**:1701–9. doi: 10.1111/dme.13516.
30. Xie X, Bai G, Liu H, *et al.* Early predictors in the onset of type 2 diabetes at different fasting blood glucose levels. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 2021;**14**:1485–92. doi: 10.2147/DMSO.S301352.
31. Al-Saeed AH, Constantino MI, Molyneaux L, *et al.* An inverse relationship between age of type 2 diabetes onset and complication risk and mortality: the impact of youth-onset type 2 diabetes. *Diabetes Care* 2016;**39**:823–9. doi: 10.2337/dc15-0991.
32. Echouffo-Tcheugui JB, Niiranen TJ, McCabe EL, *et al.* An early-onset subgroup of type 2 diabetes: a multigenerational, prospective analysis in the framingham heart study. *Diabetes Care* 2020;**43**:3086–93. doi: 10.2337/dc19-1758.
33. Middleton TL, Constantino MI, Molyneaux L, *et al.* Young-onset type 2 diabetes and younger current age: increased susceptibility to retinopathy in contrast to other complications. *Diabet Med* 2020;**37**:991–9. doi: 10.1111/dme.14238.
34. Hillier TA, Pedula KL. Complications in young adults with early-onset type 2 diabetes: losing the relative protection of youth. *Diabetes Care* 2003;**26**:2999–3005. doi: 10.2337/diacare.26.11.2999.
35. Fry A, Littlejohns TJ, Sudlow C, *et al.* Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol* 2017;**186**:1026–34. doi: 10.1093/aje/kwx246.