BMJ Open

Cohort profile: Health in Central Denmark (HICD) cohort - a register-based questionnaire survey on diabetes and related complications in the Central Denmark Region

Lasse Bjerg ,[1,2] Else-Marie Dalsgaard,[1] Kasper Norman,[1] Anders Aasted Isaksen,[3] Annelli Sandbæk[1,3]

To cite: Bjerg L, Dalsgaard E-M, Norman K, *et al*. Cohort profile: Health in Central Denmark (HICD) cohort - a register-based questionnaire survey on diabetes and related complications in the Central Denmark Region. *BMJ Open* 2022;**12**:e060410. doi:10.1136/bmjopen-2021-060410

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-060410).

Received 20 December 2021
Accepted 17 June 2022

[1]Steno Diabetes Center Aarhus, Aarhus University Hospital, Aarhus N, Denmark
[2]Viborg Regional Hospital, Viborg, Midtjylland, Denmark
[3]Department of Public Health, Aarhus University, Aarhus, Midtjylland, Denmark

Correspondence to
Dr Lasse Bjerg; lassehan@rm.dk

ABSTRACT

Purpose The Health in Central Denmark (HICD) cohort is a newly established cohort built on extensive questionnaire data linked with laboratory data and Danish national health and administrative registries. The aim is to establish an extensive resource for (1) gaining knowledge on patient-related topics and experiences that are not measured objectively at clinical health examinations and (2) long-term follow-up studies of inequality in diabetes and diabetes-related complications.

Participants A total of 1.3 million inhabitants reside in the Central Denmark Region. Using register data and a prespecified diabetes classification algorithm, we identified 45 507 persons aged 18–75 years with prevalent diabetes on 31 December 2018 and a group without diabetes of equal size matched by sex, age and municipality. A 90-item questionnaire was distributed to eligible members of this cohort on 18 November 2020 (estimated time required for completion: 15–20 min).

Findings to date We invited 90 854 persons to take part in the survey, of whom 51 854 answered the questionnaire (57.1%). Among these respondents, 2,832 persons had type 1 diabetes (55.9%), 21,140 persons had type 2 diabetes (53.2%), while 27,892 persons were part of the matched group without diabetes (60.4%). In addition to questionnaire data, the cohort is linked to nationwide registries that provide extensive data on hospital diagnoses and procedures, medication use and socioeconomic status decades before enrolment while laboratory registries has provided repeated measures of biochemical markers, for example, lipids, albuminuria and glycated haemoglobin up to 10 years before enrolment.

Future plans The HICD will serve as an extensive resource for studies on patient-related information and inequality in type 1 diabetes and type 2 diabetes. Follow-up is planned to continue for at least 10 years and detailed follow-up questionnaires, including new topics, are planned to be distributed during this period, while registry data are planned to be updated every second year.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ The Health in Central Denmark (HICD) cohort is an ongoing diabetes research resource with information on more than 50 000 persons with and without diabetes.
⇒ The HICD cohort is built on questionnaire data combined with routinely collected nationwide health and administrative registry data and laboratory data.
⇒ Sequential updates of registry data and laboratory data will be performed for at least 10 years and detailed follow-up questionnaires are planned during this period.
⇒ The HICD cohort represents a heterogenic population with a high degree of variability, for example, in terms of diabetes duration and diabetes treatment.
⇒ The HICD cohort allows researchers to explore long-term associations that focus on patient-related health information, development of classical and non-classical diabetes-related complications, and inequality in diabetes.

INTRODUCTION

Diabetes is affecting a large proportion of the global population and the prevalence is expected to rise in years to come.[1] Both type 1 diabetes and type 2 diabetes are associated with a wide range of complications, including cardiovascular disease, diabetic kidney disease, retinopathy and neuropathy,[2] and both are associated with increased mortality compared with the general population.[3] Furthermore, diabetes is associated with increased risk of mental health problems and lower quality of life.[4] Physicians tend to focus on the classical complications (eg, cardiovascular disease, diabetic kidney disease, retinopathy and neuropathy) that can be measured objectively, but from the patient's perspective, the subjective daily life challenges may be more important. Undoubtedly, there has been a massive development in diabetes care within the last decades reflected by the decreasing incidence of classical complications and mortality.[3 5] The decreasing mortality and complication incidence in both

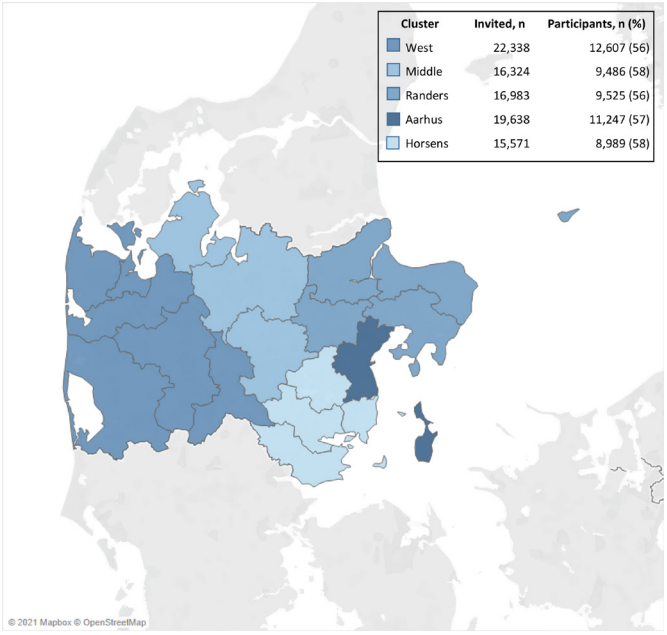

Figure 1 The Health in Central Denmark cohort was invited from the entire Central Denmark Region. This region is divided in five clusters. The box in the top corner provides the number of invitees and the number of participants from each cluster.

type 1 diabetes and type 2 diabetes is mainly driven by systematic intervention based on knowledge from solid randomised trials.[6–8] However, inequality in diabetes management exists and may have negative consequences.[9 10] Persons with low socioeconomic position have higher risk of developing type 2 diabetes and they have higher risk of developing complications. This is most prominent in type 2 diabetes but also in type 1 diabetes there is socioeconomic inequality in mortality, morbidity and diabetes management.[9] As an example, those with low socioeconomic status are less likely to be monitored, more likely to receive poor diabetes management, and have higher mortality than those with high socioeconomic status.[9 11] Therefore, there is a strong need for large-scale cost-effective surveys, that shed light both on the individual patient's perspective, while also focusing on inequality in diabetes management and its long-term consequences.

Health in Central Denmark (HICD) is a cohort comprising survey data collected through electronic questionnaires combined with national register and laboratory data. The Danish healthcare system is publicly funded and all Danish residents have free access to most healthcare services.[12] Danish national registries hold information on routinely collected administrative data and contacts with social services. Also, healthcare services are extensively documented at an individual level in national healthcare registries. Since 1968, all Danish residents are assigned a unique civil personal registration number (CPR number)[12 13] at birth or on immigration. Due to the use of the CPR number as personal identifier in the data sources, all available data can be combined on an individual level. The self-reported information by electronic questionnaires combined with nationwide register data will provide an opportunity to explore key factors in the development of diabetes and diabetes-related complications. The HICD cohort holds register data collected decades before inclusion and will be followed for a minimum of 10 years for a variety of clinical and sociodemographic outcomes in the existing registries. Follow-up questionnaires are planned in the coming years, allowing knowledge to be gained on both short-term and long-term consequences of diabetes. Due to the nature of the data, this can be conducted both as retrospective and prospective studies.

This paper provides a description of the HICD cohort as an extensive diabetes research resource with information on more than 50 000 persons with and without diabetes. This groundwork allows national and international researchers to explore long-term associations that are not traditionally within the scope of clinical health examinations and diabetes research, such as those involving patient-related health information, development of classical and non-classical diabetes-related complications and inequality in diabetes.

Cohort description

The HICD cohort is based on data collected in the Central Denmark Region, which is one of five administrative regions in the Danish tax-funded healthcare system (figure 1). The region has 1.3 million inhabitants (22% of the entire Danish population) and a mix of rural and urban municipalities. The cohort consists of adult persons between 18–75 years with register-classified prevalent diabetes on 31 December 2018 and a matched group without diabetes. We distributed the electronic questionnaire to the eligible members of the cohort on 18 November 2020 and the survey data collection ended on 7 February 2021.

Population

We identified all inhabitants in Central Denmark Region between the ages 18 and 75 years on 31 December 2018 in the Danish Civil Register (n=942 572).[13] From this population, all persons with diabetes were identified using national Danish registers. Persons with diabetes were identified based on hospital diagnoses (International Classification of Diseases (ICD)-10 codes DE10-DE14) from the Danish National Patient Register,[14 15] redeemed glucose-lowering medications were retrieved from the Danish National Prescription Register[16] (ATC-code (Anatomic Therapeutic Chemical code) A10, except Saxenda), diabetes-specific podiatrist services from the Danish National Health Service Register[17] and laboratory results for glycated haemoglobin (HbA1c) (≥48 mmol/mol) from the Clinical Laboratory Information System in Central Denmark Region, LABKA II[18] (NPU27300 and NPU03835) (table 1).

In total, 47 507 persons with diabetes were identified in Central Denmark Region (figure 2). From the remaining population in Central Denmark Region (n=895 065), a group without diabetes matching the population with diabetes on gender, age and municipality was drawn by

Table 1 Register data available in Health in Central Denmark cohort

Register data	Main variables	Data available*
Danish Civil Register[13]	Unique 10-digit personal identification.	1977–2018
Danish National Patient Register[14]	ICD-10-codes and procedures. Dates of admission, discharge and outpatient visit.	1977–2018
Danish National Health Service Registry[17]	Code for provider and type of consultation. Service provided. Date of visit.	1990–2019
Danish National Prescription Register Statistic[16]	Anatomic Therapeutic Chemical Classification system. Number, daily doses, date of redemption.	1995–2019
Clinical Laboratory Information System (LABKA), The Central Denmark Region[18]	Nomenclature, Properties, and Units (NPU) coding system for types of test. Test results and units. Date and time of sample and test result.	2011–2018
Register of Laboratory Results for Research[27]	NPU coding system for types of test. Test results and units. Date and time of sample and test result.	2008–2019
Cause of Death Register[26]	Date and cause of death	1970–2019
Statistics Denmark[28]	Sex. Age. Data of birth. Household. Family. Civil status. Residence. Country of origin. Citizenship. Occupation (6-digit). Socioeconomic classification (3-digit). Family equivalent income. Disposal income. Highest completed education (4-digit).	2000–2019

*The Health in Central Denmark is planned to update linked registry data every second year during follow-up.
ICD-10, International Classification of Diseases, 10th edition.

use of a simple random sample, that is, the matched group without diabetes is not matched directly with an individual with diabetes (figure 2).

All persons who migrated out of the region, were registered with an address protection, or died between 31 December 2018 (the date the register data were extracted) and 18 November 2020 (the date the survey was distributed) were excluded from the study. In total, the final study population consisted of 90 854 persons; 44 659 persons with diabetes and 46 195 without diabetes (figure 2). The cohort is registered in the Central Denmark Region internal register of research projects.

Patient and public involvement

The study design and questionnaire development was conceptualised by the authors of this manuscript and discussed with healthcare professionals at Steno Diabetes Center Aarhus. Patients were not involved in the study design. The results of the study will be presented at national and international conferences, research papers and relevant dissemination channels including local and social media, and via the website: www.hicd.rm.dk

Invitation

Using a prespecified algorithm, the definition of type 1 diabetes was based on type-specific diabetes diagnoses among primary diagnoses from medical specialty departments and purchases of glucose-lowering medication. In detail, for a person with diabetes to be classified as type 1 diabetes, they had to either (1) have purchased insulins and never any other types of glucose-lowering medication, and have a least one type 1-specific diagnosis, or (2) have a majority of type 1-specific diagnoses from endocrinology departments (or from medical departments, if no records of contacts to endocrinology departments), and a purchase of insulin within 180 days after onset of diabetes, and a cumulative insulin-dose more than twice the combined dose of other types of glucose-lowering medications, as measured by defined daily doses. In both cases, they had to have purchased insulin within the last year prior to inclusion. Those who were not classified as having type 1 diabetes were classified as type 2 diabetes.

In cooperation with Statistics Denmark, an invitation, including a link to a web-based self-administrated questionnaire, was sent out to all persons in the study

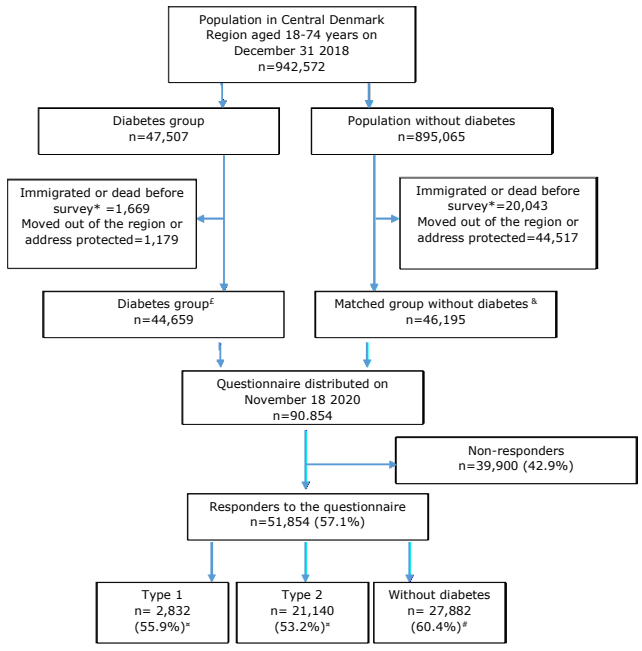

Figure 2 Flow of study population, Health in Central Denmark. £, Diabetes group—identified through algorithm. &, Matched group without diabetes—matched with diabetes group on sex, age and municipality. *, register-based population data was derived on 31 December 2018 and survey was distributed on 18 November 2020. ¤, percentage of individuals with type 1 diabetes or type 2 diabetes. #, percentage of individuals without diabetes.

population on 18 November 2020. In total, 90% of the study population received the invitation via digital mail while those exempted from digital mail received the invitation by postal delivery. An electronic reminder was sent after 2 weeks and a postal reminder was sent to all persons who had not yet responded to the questionnaire after 4 weeks. Finally, 6 weeks after the first invitation, persons who had only answered part of the web-based questionnaire received a request to complete the questionnaire. During data collection, a hotline was established at Steno Diabetes Center Aarhus to answer any questions from invitees, and all invitees had access to a webpage with detailed information on the survey. Lastly, Statistics Denmark could be contacted directly during the data collection period. By answering the questionnaire, respondents gave their informed consent. Among persons who completed the entire questionnaire, 20 cash prizes were drawn. Data collection ran until 7 February 2021. By this time, the HICD cohort consisted of 51 854 persons (57.1% of the study population), who had responded to the invitation and filled in questionnaire data.

Questionnaire description

We collected self-reported data on anthropometric measures, lifestyle habits, well-being, social life, symptoms and diseases, sleeping problems, dyspnoea, dental status, sexual dysfunction, fertility problems and family history of diabetes. Among persons with self-reporting diabetes, questions about diabetes-specific problems and self-management were asked. In the following section a short description of validated questionnaires is provided. The distributed questionnaire included 90 items, and the estimated time for completion of the questionnaire was 15–20 min for each respondent.

Well-Being Index 5-items[19] is a short, self-administered questionnaire of well-being over the last 2 weeks. It consists of five positively worded items that are rated on a 6-point Likert scale, ranging from 0 (at no time) to 5 (all of the time). The raw scores are transformed to a score from 0 to 100, with lower scores indicating worse well-being. A score of ≤50 indicates poor well-being and suggests further investigation into possible symptoms of depression. A score of 28 or below is indicative of depression. The Short Form Health Survey[20] is a validated health-related quality-of-life questionnaire consisting of 12 questions assessing the impact of health on everyday life. It produces separate scores of overall mental and physical well-being and is often used as a quality-of-life measure. The Insomnia Severity Index[21 22] is a validated brief instrument containing seven questions designed to assess the severity of both night-time and day-time components of insomnia, with higher sum score indicating more severe insomnia problems. The Brief Sexual Symptoms Checklist (BSSC)[23] is a screening tool to explore sexual dysfunction. The screening tool consists of separate checklists for men (BSSC-M) and women (BSSC-W). The checklists are not sum scores but should be handled as single items. Problem Areas in Diabetes, 5-items scale[24] is a validated self-reported instrument that assesses diabetes-related emotional distress and covers a range of negative emotional problems for persons with diabetes. The five items are rated from 0 (not a problem) to 4 (serious problem). The sum score ranges from 0 to 20 with higher score indicating higher diabetes-related emotional distress. The Patient Activation Measure (PAM-13)[25] assesses the individual's self-reported knowledge, skills and confidence regarding diabetes and measures patient activation in self–care, for example, internal capability to make health-promoting actions. PAM-13 has been found to be reliable and valid.[25] Items scores range from 0 (not applicable), 1 (strongly disagree), 2 (disagree), 3 (agree) to 4 (strongly agree), and the sum score is transferred to a score from 0 to 100 with a higher score indicating higher patient activation.

Register data

Register data are available at least 10 years before inclusion and for most register more than 20 years. Register data in HICD will be updated frequently during follow-up until the end of 2030. Table 1 presents the registers in more detail. In brief, data were linked using the unique personal number in the Danish Civil Register.[12 13] The Danish National Patient Register[14 15] holds information on all hospital contacts, including diagnoses according to the ICD-10 coding system, admission dates, discharge dates as well as outpatient visits. The Danish National

Health Service Register[17] holds information on all visits to health professionals in the primary care sector. The Danish National Prescription Register[16] holds information on medication purchases. Cause of death will be obtained from the Cause of Death Register.[26]

The Danish Register for Laboratory Results for Research and The regional Clinical Laboratory Information System[18 27] holds laboratory test results from clinical laboratories, including blood samples and urine samples from 2011 onwards. It holds information on both test result, unit, and date of the sample.

The demographic information provided by Statistics Denmark is updated once per year. The HICD is enriched and will be updated with individual-level data on demography, migration, household, family, income, education and labour-marked affiliation (table 1).[28]

Follow-up

Follow-up is planned to continue for at least 10 years. Detailed follow-up questionnaires, including both repeated questionnaire items and new topics, are planned to be distributed during this period, while registry data are planned to be updated every second year. Follow-up questionnaires will be sent out to the initial invited population even if the first questionnaire was not completed. We expect the response rate to decline during the study period. In order to mitigate this, we will collaborate with Statistics Denmark, who possess expertise in both collecting questionnaire data and in maximising participation rate. In future analyses of follow-up data, we will handle missing data according to the research questions, and several methods are available. Among methods that could be relevant to apply are imputation, weighting and analysis using mixed-effect methods. To validate our cohort against the general population, non-responder analyses will be performed using register data.

Characteristics of study population

In total, 51 854 responded to the questionnaire (57.1% of the study population). Uptake in the HICD cohort was similar across geographical clusters in the Central Denmark Region, that is, between 56% and 58% of the identified study population responded to the questionnaire from each of five areas within the region (figure 1). The response percentage to the questionnaire survey was 55.9% in persons with type 1 diabetes, 53.2% in persons with type 2 diabetes and 60.4% of those in the group without diabetes, respectively (figure 2).

A detailed description of all responders is presented in table 2 while diabetes-specific characteristics are presented in table 3. Using the aforementioned prespecified algorithm, 2832 (5.6%) were identified with type 1 diabetes and 21 140 (40.7%) were identified with type 2 diabetes, while 27 882 (53.7%) persons were in the matched group without diabetes. Among the 2832 persons with type 1 diabetes, 54.6% were men, the median age at inclusion was 53 years, 48.0% never smoked, 15.4% had a sedentary activity level and 16.3% had a body mass index (BMI) above 30 kg/m^2. Out of 21 140 persons with type 2 diabetes, 59.1% were men, the median age was 66 years at inclusion, 37.4% never smoked, 24.7% had a sedentary activity level, while 47.3 had a BMI above 30 kg/m^2. In the group of persons without diabetes, 15 714 (56.4%) were men and the median age at inclusion was 65 years. In the latter group, 45.0% never smoked, 11.1% had a sedentary activity level and 18.6% had a BMI above 30 kg/m^2. In total, 52.0% reported hypertension, with the highest percentage among persons with type 2 diabetes. Also, the self-reported percentages of acute myocardial infarction, stroke and cancer were highest among persons with type 2 diabetes.

Non-responder characteristics

In total, 42.9% of the invited study population did not respond to the survey. The non-respond patterns were similar across the three groups (type 1 diabetes, type 2 diabetes and persons without diabetes) and are presented in table 4. For type 1 diabetes, the age was higher in those who responded. Similarly, the age of responders compared with non-responders was higher in both type 2 diabetes and in the group without diabetes. Non-responders were more likely to live alone and a higher fraction among non-responders compared with responders had under 10 years of education in all three groups. The percentage of people with cardiovascular disease at inclusion were similar in responders and non-responders. The percentage of women was higher in responders compared with non-responders among persons with type 1 diabetes and among persons without diabetes, while the percentage of women was highest among non-responders in type 2 diabetes.

Findings to date

The HICD cohort is newly established. Hence, no prior work from the cohort has been published to date.

Strengths and limitations

The HICD cohort is established on questionnaire data and nationwide register data. With more than 50 000 responders, it will be possible to look into rare exposures and outcomes as well as details in specific subgroups of the population.

The identification of persons with diabetes was based on an algorithm including both information on hospital diagnoses, glucose-lowering medication, podiatric services and laboratory values of HbA1c. Although a validation study of the Danish Diabetes Register[29] recommended the use of HbA1c in identification of individuals with diabetes, laboratory data has not been available until recently.[29] We only included the population in the Central Denmark Region as we wanted to include analysis of HbA1c in the algorithm and only had access to laboratory data from this region. The response rate was as expected and similar to other surveys conducted in Denmark,[30] adding up to more than 57% of the invited population, which ensures a high representability of the

Table 2 Characteristics of study population in Health in Central Denmark

	Type 1 diabetes	Missing n (%)	Type 2 diabetes	Missing n (%)	Without diabetes	Missing n (%)	Total
n	2832		21 140		27 882		51 854
Men	1545 (54.6)	0	12 492 (59.1)	0	15 714 (56.4)	0	29 751 (57.4)
Age, years	53 (40; 63)	0	66 (58; 72)	0	65 (56; 71)	0	65 (56; 71)
Living alone	831 (29.3)	0	6140 (29.0)	0	6198 (22.2)	0	13 169 (25.4)
Education, ≤10 years	477 (17.1)	48 (1.7)	5810 (28.1)	437 (2.1)	4981 (18.2)	456 (1.6)	11 268 (22.1)
Height, cm	175 (168; 181)	27 (1.0)	173 (167; 180)	334 (1.6)	174 (168; 180)	397 (1.4)	174 (167; 180)
Weight, kg	78 (68; 89)	10 (0.4)	90 (78; 103)	110 (0.5)	79 (69; 90)	102 (0.4)	83 (72; 95)
Body mass index, kg/m^2	25.3 (22.9; 28.3)	35 (1.2)	29.6 (26.2; 33.6)	393 (1.9)	25.9 (23.4; 28.3)	450 (1.6)	27.1 (24.2; 30.9)
Smoking status							
Daily	379 (13.5)		3249 (15.6)		3262 (11.8)		6890 (13.5)
Weekly	43 (1.5)		271 (1.3)		414 (1.5)		728 (1.4)
Less than weekly	70 (2.5)		193 (0.9)		393 (1.4)		656 (1.3)
Stopped	971 (34.5)		9327 (44.8)		11 095 (40.3)		21 393 (41.8)
Never	1350 (48.0)	19 (0.7)	7802 (37.4)	298 (1.4)	12 401 (45.0)	317 (1.1)	21 553 (42.1)
How often do you drink alcohol?							
Never	283 (10.1)		3771 (18.1)		2074 (7.5)		6128 (11.9)
Monthly or less	578 (20.6)		5547 (26.6)		4278 (15.5)		10 403 (20.3)
2–4 times a month	868 (30.9)		5307 (25.4)		7489 (27.1)		13 664 (26.6)
2–3 times a week	667 (23.7)		3784 (18.1)		8342 (30.2)		12 793 (24.9)
4 or more times a week	414 (14.7)	22 (0.8)	2479 (11.9)	252 (1.2)	5427 (19.7)	272 (0.1)	8320 (16.2)
Leisure time activity level							
Regular hard physical training	92 (3.3)		159 (0.8)		470 (1.7)		721 (1.4)
Regular physical activity	550 (19.5)		2442 (11.6)		6005 (21.7)		8997 (17.5)
Some physical activity	1740 (61.8)		13 178 (62.9)		18 148 (65.5)		33 066 (64.2)
Sedentary	434 (15.4)	16 (0.6)	5184 (24.7)	177 (0.8)	3082 (11.1)	177 (0.6)	8700 (16.9)
Does it happen you are alone even though you want to be with others?							
Often	245 (8.9)		1830 (8.9)		1379 (5.1)		3454 (6.8)
Once in a while	539 (19.5)		4395 (21.3)		4845 (17.8)		9779 (19.3)
Rarely	846 (30.6)		5719 (27.7)		7970 (29.2)		14 535 (28.7)
No	1137 (41.1)	65 (2.3)	8710 (42.2)	486 (2.3)	13 068 (47.9)	620 (2.2)	22 915 (45.2)
SF-12*							
SF-12 physical component score*, median (IQR)	51.2 (42.3; 55.7)	257 (9.1)	45.5 (36.7; 52.6)	2729 (12.9)	52.2 (44.7; 56.2)	3212 (11.5)	49.8 (40.8; 55.5)

Continued

Table 2 Continued

	Type 1 diabetes	Missing n (%)	Type 2 diabetes	Missing n (%)	Without diabetes	Missing n (%)	Total
SF-12 mental component*, median (IQR)	50.2 (42.3; 55.2)	257 (9.1)	48.8 (40.6; 54.8)	2729 (12.9)	52,8 (46.2; 56.7)	3212 (11.5)	51.2 (43.5; 56.1)
WHO5†, risk of depression or severe stress							
High risk	294 (10.9)		2380 (12.0)		1762 (6.6)		4436 (9.0)
Risk	420 (15.6)		2627 (13.2)		2710 (10.2)		5757 (11.7)
Low risk	1984 (73.5)	134 (4.7)	14901 (74.8)	1232 (5.8)	22086 (83.2)	1324 (4.8)	38971 (79.3)
Insomnia severity index level‡							
No clinically significant insomnia	1540 (55.8)		10857 (52.7)		17008 (62.4)		29405 (58.1)
Subthreshold insomnia	827 (30.0)		6283 (30.5)		7431 (27.3)		14541 (28.7)
Clinical insomnia (moderate severity)	320 (11.6)		2889 (14.0)		2444 (9.0)		5653 (11.2)
Clinical insomnia (severe)	71 (2.6)	74 (2.6)	578 (2.8)	533 (2.5)	355 (1.3)	644 (2.3)	1004 (2.0)
Are you satisfied with your sexual function§							
No—women	252 (20.3)		1617 (19.4)		1897 (16.2)		3766 (17.7)
Yes—women	661 (53.1)		3634 (43.6)		6183 (52.7)		10478 (49.2)
Do not know—women	331 (26.6)	43 (3.3)	3086 (37.0)	311 (3.6)	3654 (31.1)	434 (3.6)	7071 (33.2)
No—men	542 (36.2)		4759 (39.4)		4442 (29.0)		9743 (33.7)
Yes—men	664 (44.3)		4124 (34.2)		7303 (47.7)		12091 (41.9)
Do not know—men	293 (19.5)	46 (3.0)	3193 (26.4)	416 (3.3)	3561 (23.3)	408 (2.6)	7047 (24.4)
Blood pressure (self-reported)							
Systolic blood pressure	130 (122; 135)	1222 (43.2)	130 (125; 140)	8223 (38.9)	130 (123; 140)	15679 (56.2)	130 (125; 140)
Diastolic blood pressure	78 (70; 82)	1230 (43.2)	80 (75; 85)	8334 (39.4)	80 (75; 85)	15764 (56.5)	80 (75; 85)
Current or previous diseases (self-reported)							
Hypertension	1325 (49.1)		14943 (73.5)		9640 (36.0)		25908 (52.0)
Acute myocardial infarction	96 (3.6)		1824 (9.0)		900 (3.4)		2820 (5.7)
Angina	90 (3.3)		1477 (7.3)		807 (3.0)		2374 (4.8)
Stroke	119 (4.4)		1351 (6.6)		1026 (3.8)		2496 (5.0)
Cancer	168 (6.2)		2439 (12.0)		2865 (10.7)		5472 (11.0)
Psychiatric disease less than 6 months	155 (5.7)		994 (4.9)		1119 (4.2)		2268 (4.6)
Psychiatric disease more than 6 months	283 (10.5)	133 (4.7)	2275 (11.2)	802 (3.8)	1929 (7.2)	113 (4.1)	4487 (9.0)

Categorical data are expressed as n (%) and continuous as medians (p25; p75).
SF-12-DK physical component score and mental component score weights, Defactum (2018).
*Short Form Health Survey (SF-12).[20]
†Well-Being Index 5-items (WHO-5).[19]
‡Insomnia severity index.[21 22]
§Brief Sexual Symptoms Checklist (BSSC-men and BSSC-women).[23]

Table 3 Self-reported diabetes-specific characteristics in the Health in Central Denmark cohort

	Type 1 diabetes	Missings n (%)	Type 2 diabetes	Missings n (%)	Total
n	2832	0.0	21 140	0.0	23.972
Age at diagnosis onset, years	21 (12; 33)	132 (4.7)	54 (45; 60)	2383 (11.3)	51 (41; 60)
Diabetic neuropathy, yes	400 (14.9)	149 (5.3)	2481 (13.3)	2532 (12.0)	2881 (13.5)
Diabetic retinopathy, yes	774 (28.8)	142 (5.0)	2481 (13.2)	2330 (11.0)	3255 (15.1)
Parental diabetes, yes	611 (22.3)	95 (3.4)	8347 (40.9)	742 (3.5)	8958 (38.7)
Sibling with diabetes, yes	455 (16.7)	105 (3.7)	5190 (25.6)	890 (4.2)	5645 (24.6)
PAID5*, sum score	5.1 (4.5)	199 (7.0)	3.0 (0.0; 6.0)	2918 (13.8)	3.0 (0.0; 6.0)
Diabetic distress, yes	688 (26.1)	199 (7.0)	3207 (17.6)	2918 (13.8)	3895 (18.7)
PAM-13†, score	63.2 (52.9; 75.3)	293 (10.4)	60.0 (49.9; 70.8)	3260 (15.4)	60.0 (49.9; 70.8)
Patient activation level					
Not important	377 (14.8)		3136 (17.5)		3513 (17.2)
Lack of knowledge	377 (14.8)		3140 (17.6)		3517 (17.2)
Beginning, action	698 (27.5)		5801 (32.4)		6499 (31.8)
Taking action	1087 (42.8)	293 (10.4)	5803 (32.5)	3260 (15.4)	6890 (33.7)

Categorical data are expressed as n (%) and continuous as medians (p25; p75)
*Problem Areas in Diabetes, 5-items scale (PAID5).[24]
†Patient Activation Measure (PAM-13).[25]

general population. To our knowledge, this is the largest population-wide cohort with both questionnaire data and register data that uses HbA1c in the identification of persons with diabetes in Denmark.

An important strength of this study is the possibility to link questionnaire data to register data. This does not only open the opportunity for collecting comprehensive health and administrative data at an individual level at inclusion, it also provides a source for repeating measures both prior to and after inclusion in the cohort. Due to the historical and continuing data collection in the Danish registers we are able to perform both retrospective studies and long-term prospective studies in a cost-effective way. In addition, the nationwide registers enable us to follow the group of non-responders.

LIMITATIONS

The non-responder analysis shows some differences between responders and non-responders, which could limit the external validity of the cohort. However, we expect any selection bias to be small, as the differences were limited and had similar patterns among persons with and without diabetes. Also, we expect any investigator bias to be minimal as the information in the Danish health registers is collected for administrative reasons

Table 4 Characteristics of responders and non-responders to the Health in Central Denmark questionnaire

	Type 1 diabetes		Type 2 diabetes		Without diabetes		Missing
	Responder	Non-responder	Responder	Non-responder	Responder	Non-responder	%
n (%)	2832 (55.9)	2061 (44.9)	21 140 (53.2)	18 626 (46.8)	27 882 (6.4)	18 313 (39.6)	0.0
Sex, n (%)							
Men	1545 (54.6)	1324 (64.2)	12 492 (59.1)	10 506 (56.4)	15 714 (56.4)	10 873 (59.4)	0.0
Women	1287 (45.4)	737 (35.8)	8648 (40.9)	8120 (43.6)	12 168 (43.6)	7440 (40.6)	0.0
Age, median (IQR)	53 (40; 63)	46 (33; 57)	66 (58; 72)	64 (55; 71)	65 (56; 71)	60 (49; 70)	0.0
Cohabitant, n (%)							
Living alone	831 (29.3)	879 (42.6)	6140 (29.0)	7926 (42.6)	6198 (22.2)	6428 (35.1)	0.0
Education, n (%)							
<10 years	477 (17.1)	581 (29.2)	5810 (28.1)	7945 (45.1)	4981 (18.2)	5327 (30.3)	3.1
Cardiovascular disease, n (%)	297 (10.5)	172 (8.3)	3630 (17.2)	3356 (18.0)	2173 (7.8)	1386 (7.6)	0.0

Categorical data are expressed as n (%) and continuous as medians (p25; p75)

by clinicians and health personnel during diagnosis and treatment. Lastly, there is strong tradition for validating the data available in the Danish health and administrative registries,that is, the validity of the nationwide registries and the regional laboratory data are high.[14–16 31]

For some variables, missing data are considerable. Among those who answered the questionnaire, we did not expect persons with diabetes to be more or less likely to answer specific questions compared with persons without diabetes. Therefore, any introduced information bias due to missing is expected to be non-differential, hence, associations would be biased toward the null. Most of the included questions focus on concurrent data, limiting the chance of recall bias. However, for more historical data the chance of recall bias is higher. Also, the survey mainly included validated questionnaires, but the validity of each single questionnaire scale does differ and therefore this should be taken into account in future studies. Register data, such as education, occupation and migration, may be missing or misclassified in the registries. Therefore, including self-reported demographic and socioeconomic data in the survey could add valuable knowledge.

As we use an epidemiological approach for identification of persons with diabetes, there may be some misclassification. Presumably, this is especially true for newly diagnosed persons, as the cohort was defined based on register data on 31 December 2018 while the questionnaire was distributed almost 2 years later. As we will follow this cohort for several years, we expect individuals to develop diabetes during follow-up or change status according to the algorithm as additional data are added. The date of diagnosis and type of diabetes will be included each time the data are updated with new registry data, so the information can be used in future research questions.

The large cohort size also introduces some limitations, as the cohort is heterogenic, for example, when it comes to diabetes duration and diabetes treatment. Some persons have had diabetes for decades while others are newly diagnosed, and some are well-regulated in terms of cardiovascular risk factors while others are not. Lastly, the catchment area for the cohort was the Central Denmark Region which includes both rural and urban municipalities. The generalisability of the cohort may be limited if the proportion of responders differed between urban and rural areas. However, this was not the case (data not shown). In addition, the matched group without diabetes was drawn from the entire general population limiting selection bias.

To sum up, the overarching goal of the HICD cohort is to provide a data-rich resource for diabetes research, open to both national and international researchers. Inequality in diabetes is increasing and the combination of large-scale questionnaire data and nationwide register data provides new opportunities to explore long-term associations of patient-related information and classical and non-classical diabetes-related complication, for example, in smaller sociodemographic subgroups.

Collaboration

Information on the HICD can be found on the website: www.hicd.rm.dk. The HICD is managed by Steno Diabetes Center Aarhus (SDCA), Aarhus University Hospital and has a steering committee with members from both epidemiological and clinical diabetes research. The steering committee strongly encourages interested researchers across disciplines to contact Director Annelli Sandbæk at ANESND@rm.dk. SDCA will provide applicants with detailed information on the available data and the application procedure. For more information on how to get access to data and how to apply to the steering committee, please contact auh.helbredimidt@rm.dk. All research on HICD data are based on questionnaires and register data which is why no ethical approval of research projects is required in Denmark. All data analyses are to be performed on SDCA's remote access to Statistics Denmark's server. A senior researcher affiliated with SDCA will be designated contact person to potential research projects.

Contributors LB, E-MD, KN and AS designed the HICD cohort. AAI and LB performed data management. KN and E-MD performed data management and performed all statistical analyses. LB wrote the first draft of the manuscript. All authors revised and edited the manuscript and approved the final version before submission. LB is the guarantor of this work and, as such, had full access to all the data in the study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. All data analyses are to be performed on SDCA's remote access to Statistics Denmark's server. A senior researcher affiliated with SDCA will be designated contact person to any potential research projects.

ORCID iD

Lasse Bjerg http://orcid.org/0000-0002-4724-7276

REFERENCES

1 Cho NH, Shaw JE, Karuranga S, *et al*. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res Clin Pract* 2018;138:271–81.
2 Harding JL, Pavkov ME, Magliano DJ, *et al*. Global trends in diabetes complications: a review of current evidence. *Diabetologia* 2019;62:3-16.

3 Rawshani A, Rawshani A, Franzén S, *et al*. Mortality and cardiovascular disease in type 1 and type 2 diabetes. *N Engl J Med* 2017;376:1407–18.
4 Feng X, Astell-Burt T. Impact of a type 2 diabetes diagnosis on mental health, quality of life, and social contacts: a longitudinal study. *BMJ Open Diabetes Res Care* 2017;5:e000198.
5 Gregg EW, Cheng YJ, Saydah S, *et al*. Trends in death rates among U.S. adults with and without diabetes between 1997 and 2006: findings from the national health interview survey. *Diabetes Care* 2012;35:1252–7.
6 , Orchard TJ, Nathan DM, *et al*, Writing Group for the DCCT/EDIC Research Group. Association between 7 years of intensive treatment of type 1 diabetes and long-term mortality. *JAMA* 2015;313:45–53.
7 Holman RR, Paul SK, Bethel MA, *et al*. 10-Year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med* 2008;359:1577–89.
8 Gæde P, Oellgaard J, Carstensen B, *et al*. Years of life gained by multifactorial intervention in patients with type 2 diabetes mellitus and microalbuminuria: 21 years follow-up on the steno-2 randomised trial. *Diabetologia* 2016;59:2298–307.
9 Scott A, Chambers D, Goyder E, *et al*. Socioeconomic inequalities in mortality, morbidity and diabetes management for adults with type 1 diabetes: a systematic review. *PLoS One* 2017;12:e0177210.
10 Espelt A, Arriola L, Borrell C, *et al*. Socioeconomic position and type 2 diabetes mellitus in Europe 1999-2009: a panorama of inequalities. *Curr Diabetes Rev* 2011;7:148–58.
11 Rawshani A, Svensson A-M, Zethelius B, *et al*. Association between socioeconomic status and mortality, cardiovascular disease, and cancer in patients with type 2 diabetes. *JAMA Intern Med* 2016;176:1146–54.
12 Schmidt M, Pedersen L, Sørensen HT. The Danish civil registration system as a tool in epidemiology. *Eur J Epidemiol* 2014;29:541–9.
13 Pedersen CB. The Danish civil registration system. *Scand J Public Health* 2011;39:22–5.
14 Lynge E, Sandegaard JL, Rebolj M. The Danish national patient register. *Scand J Public Health* 2011;39:30–3.
15 Schmidt M, Schmidt SAJ, Sandegaard JL, *et al*. The Danish national patient registry: a review of content, data quality, and research potential. *Clin Epidemiol* 2015;7:449–90.
16 Kildemoes HW, Sørensen HT, Hallas J. The Danish national prescription registry. *Scand J Public Health* 2011;39:38–41.
17 Andersen JS, Olivarius NDF, Krasnik A. The Danish national health service register. *Scand J Public Health* 2011;39:34–7.
18 Grann AF, Erichsen R, Nielsen AG. Existing data sources for clinical epidemiology: The clinical laboratory information system (LABKA) research database at Aarhus University. In: *Clinical epidemiology*. Denmark, 2011: 3. 133–8.
19 Topp CW, Østergaard SD, Søndergaard S, *et al*. The WHO-5 well-being index: a systematic review of the literature. *Psychother Psychosom* 2015;84:167–76.
20 Ware J, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220–33.
21 Dieperink KB, Elnegaard CM, Winther B, *et al*. Preliminary validation of the insomnia severity index in Danish outpatients with a medical condition. *J Patient Rep Outcomes* 2020;4:18.
22 Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297–307.
23 Romero Otero J, García Gómez B, Medina Polo J, *et al*. The brief sexual symptom checklist as a screening tool for sex-related problems in the primary care setting: focus on erectile dysfunction. *Urol Int* 2018;101:98–105.
24 McGuire BE, Morrison TG, Hermanns N, *et al*. Short-form measures of diabetes-related emotional distress: the problem areas in diabetes scale (PAID)-5 and PAID-1. *Diabetologia* 2010;53:66–9.
25 Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form patient activation measure (PAM13) in a Danish version. *BMC Public Health* 2009;9:209.
26 Helweg-Larsen K. The Danish register of causes of death. *Scand J Public Health* 2011;39:26–9.
27 Arendt JFH, Hansen AT, Ladefoged SA, *et al*. Existing data sources in clinical epidemiology: laboratory information system databases in Denmark. *Clin Epidemiol* 2020;12:469–75.
28 Statistics Denmark. Available: www.dst.dk/da/TilSalg/Forskningsservice [Accessed 16 Dec 2021].
29 Green A, Sortsø C, Jensen PB, *et al*. Validation of the Danish national diabetes register. *Clin Epidemiol* 2015;7:5–15.
30 Schiøtz M, Bøgelund M, Willaing I. Challenges using online surveys in a Danish population of people with type 2 diabetes. *Chronic Illn* 2012;8:56–63.
31 Sundbøll J, Adelborg K, Munch T, *et al*. Positive predictive value of cardiovascular diagnoses in the Danish national patient registry: a validation study. *BMJ Open* 2016;6:e012832.