**Protocol title: Global and country-specific rates and trends in the incidence of young-onset type 2 diabetes**

**Authors:** all collaborators.

**Background**

The prevalence of type 2 diabetes in children, adolescents, and young adults is increasing in many countries. Individuals diagnosed with diabetes at a younger age (15-39 years) seemingly have a more rapid deterioration of β-cell function than is seen in those with late-onset type 2 diabetes. Compared with late-onset type 2 diabetes, young-onset type 2 diabetes is associated with longer disease exposure and increased risk for chronic complications. Individuals with young-onset type 2 diabetes also have a greater risk of complications compared with young people with type 1 diabetes, suggesting a more aggressive disease phenotype.

Our knowledge of the incidence trends of young-onset type 2 diabetes is limited to a few studies, which are mainly from high-income countries and much of the current data is based on clinic-based studies with small numbers of cases. Further, investigation of incidence trends across different countries was limited considering that the criteria used for classifying diabetes type and methods of data collection are not always consistent, and many studies of adolescents and/or young adults with type 2 diabetes used variable age cut-off points.

We have established an international diabetes collaboration of large, predominantly national, data sources from 25 countries/regions, which is the first global systematic approach to ascertain multinational trends in diabetes incidence and mortality. We have successfully used this database and published the data in a multi-country analysis of trends in the incidence of diagnosed diabetes.1 This multi-country data analysis showed that the incidence of diagnosed total or type 2 diabetes has been stable or falling since approximately 2010 onwards in both men and women in over 80% of high-income countries/regions. However, there has been little attention focussed on whether declines in the incidence of type 2 diabetes have been shared equally by all age groups. Therefore, we will take advantage of our established international diabetes consortium, to explore the temporal trends in the incidence of young-onset type 2 diabetes and to predict its future direction. We will also explore all-cause mortality trends among individuals with young-onset diabetes.

**Hypothesis:**

* The incidence of young-onset type 2 diabetes (age of diagnosis at 15-39 years) is increasing in most countries.
* Trends in the incidence of young-onset type 2 diabetes may differ by country or region, age group and sex.
* All-cause mortality trends in among younger people (age 15-39 years) with diabetes may be different to the overall declining trends observed among the whole population with diabetes.

**Aims:**

The overall aim of the project is to understand trends of the incidence of young-onset type 2 diabetes and mortality trends in younger people with diabetes in multiple sites around the world.

***Specific aims:***

Aim 1: to assess the regional and country-specific rates and trends in the incidence of young-onset type 2 diabetes and mortality trends among younger people with diabetes from 2000 onwards (or a subset thereof);

Aim 2: to investigate if the changes in the incidence of young-onset type 2 diabetes over time varied by country or region, age group and sex.

**Research Plan:**

***Infrastructure:*** To facilitate this study, a partnership has been established between the Diabetes and Population Health group at Baker Heart and Diabetes Institute and the Centers for Disease Control and Prevention in Atlanta to conduct this study.

***Data sources:*** We aimed to identify all data sources that could potentially report the number of incident cases year by year for at least three years within the years 2000 onwards. Appropriate data sources for measuring the trends in the incidence of young-onset type 2 diabetes: (i) have on-going enrolment of new members (or regular recruitment of new independent cohorts); (ii) record new-onset (incident) diabetes; (iii) record sex- and age-specific data; (iv) include at last 5000 individuals in the population at risk of developing diabetes in each calendar year; (v) have linkage to medication data. We will ask each collaborator to provide aggregate data for each calendar year on the incidence of diabetes stratified by sex and 5-year age group (15-19, 20-24, …, 35-39 years) over the time period from 2000 onwards (or a subset thereof). We will also request data on incident diabetes by diabetes type using the proposed algorithm (see details below). To contribute fully to the project, data sources are also required to have accurate information about the numbers of people in the background population (the denominator), about prevalent diabetes, and about deaths in people with and without diabetes. The most likely sources are diabetes registries, health insurance providers, health maintenance organizations and collections of electronic medical records.

Data sources were identified by three methods. First, we will contact each collaborator of the established diabetes consortium network to determine both interest and capacity to participate. Second, we continue to look for additional data sources from new collaborators, from the clinical and research networks of the investigators. Third, we recently conducted a narrative review of the incidence, prevalence, morbidity and mortality of young-onset type 2 diabetes, and we will further identify all published studies with the potential to contribute to the project.

***Data extraction and definitions:*** Each data source will be requested to provide tabular, summary data, with counts of incident and prevalent diabetes cases, deaths and the background population. No unit record data will be requested.

Data extraction and specific definitions will vary among the data sources. This is at least in part because of issues that are necessarily specific to countries and data sources. Nevertheless, the following definitions and guidelines should apply wherever possible, and exceptions should be noted.

1. Diabetes will be ascertained and defined on the basis of a diagnosis or diagnostic code provided by a relevant healthcare professional, or at least two of:
   1. the presence of two or more blood glucose or HbA1c values within the diagnostic ranges for diabetes (fasting plasma glucose ≥7.0 mmol/l (126 mg/dl), random or 2-hour plasma glucose ≥11.1 mmol/l (200 mg/dl) or HbA1c ≥6.5% (48 mmol/mol)), within a 6-month period;
   2. prescription of glucose-lowering medication for at least 3 months;
   3. the provision of a service that is unique to people with diabetes.
2. Those with gestational diabetes, secondary diabetes (e.g., drug-induced, chemical-induced, exocrine pancreatic insufficiency, and genetic defects), maturity-onset diabetes of the young, and rare forms of diabetes will be excluded.
3. People **with incident diabetes** will be classified into three groups based on age of diagnosis, time to insulin therapy and use of non-insulin glucose lowering drugs other than metformin.
4. People are classified as having certain/definite type 1 diabetes if they were diagnosed as diabetes **before 35 years of age**, the time between date of diagnosis and date of first insulin prescription was less than one year, and had never been treated with non-insulin glucose lowering drugs other than metformin.
5. People are classified as having certain/definite type 2 diabetes if they did not have any insulin therapy **within the first two years after diagnosis**.
6. All other people are classified as being of uncertain diabetes type.
7. An incident case of diabetes for a particular year is defined as a person who, between 1 January and 31 December of that year, is either:
   1. listed on a diabetes register as having diabetes onset in that calendar year; or
   2. newly diagnosed in a medical record or claims database, and was registered in the medical record or claims database for the previous 12 months and was not identified as having diabetes during that time.
8. The date of diagnosis of an incident case is the date on which the earliest diagnostic criterion (listed in (3) above) is satisfied. Thus, a person whose first criterion is satisfied in one year, and their second criterion is satisfied the next year is deemed to be an incident case in the first year.
9. Individuals with a date of diagnosis of diabetes earlier than their date of registration with the database (e.g. a person with established diabetes joining an HMO or medical practice) should not be included as an incident case (to comply with 4b), when the medical record or a claims database is the direct source of information (which may also be the case for virtual registers). This does not generally apply to stand-alone registries, in which such cases can be included, unless such cases are likely to be new immigrants. When such cases are included, it is assumed that they are represented in the population counts of the denominator in the years of and preceding their year of diagnosis. Their year of diagnosis, rather than year of registration, should usually be taken as the year in which they became an incident case. Date of registration may be used a proxy for the diagnosis date, if date of diagnosis is missing, but this needs to be recorded in the additional information of the data source.
10. A prevalent case of diabetes in any given year is someone who has diabetes, as established by the above criteria, on January 1 that year. Incident cases are not considered as prevalent cases in the calendar year in which they became an incident case.
11. Death will be ascertained either from national death registers or from death data held in the database being used.
12. The denominator for calculating incidence of diabetes each year and in each age-group is the number of people without clinically diagnosed diabetes on 1 January of each year in the population from which the incident cases are drawn – this might be the national or regional population (derived from census data), or the population in the electronic medical record or claims database. Where possible, person-years are preferred to simple counts, in order to fully account for the precise amount time an individual spends in each state.

***Analysis Plan:***

We will analyse the incidence of young-onset type 2 diabetes by age and calendar year to determine trends over calendar time and compare these across data sources. All analyses will be conducted separately for males and females. We will model the incidence rates using age and calendar time as quantitative variables. We will use Poisson likelihood for multiplicative models with events as outcome and log person-years as offset. Specifically we will fit age-period, age-period-cohort2 and Lee-Carter models3 for incidence counts using Poisson regression with person-years as denominators and provide smoothed time trends for each data source, enabling quantitative display of trends as well as formal comparison of trends between data sources. We will conduct similar analysis of mortality trends among younger people with total diabetes (type 1 and type 2 diabetes).

We have designed a purpose-built data collection form (see attachment 1 - Data Extraction Template) and a secure data-uploading facility to enable timely and secure data collection. Data preparation and management will be undertaken in the coordinating centre (Baker Heart and Diabetes Institute, Australia). The assembled database will be converted to a CSV containing all of the data. There is an additional document which asks some questions about the data source (see attachment 2).

***Publications and Disseminations:***

The results of this research will be published in peer-reviewed journals. Up to two authors will be included on each publication from each collaborating centre as well as the authors from the coordinating centre (Baker Heart and Diabetes Institute, Australia). Data obtained from each centre are aggregated data only.

**Outcomes and Significance:**

The increasing prevalence of young-onset type 2 diabetes and its aggressive trajectory heralds a resurgence of diabetes-related complications at much younger ages. Such complications are life threatening and carry a large burden for patients, families and government. The successful completion of this project will provide vital data to understand the current trends in the incidence of young-onset type 2 diabetes which is crucial for health care planning and risk factor management of those vulnerable populations with diabetes.

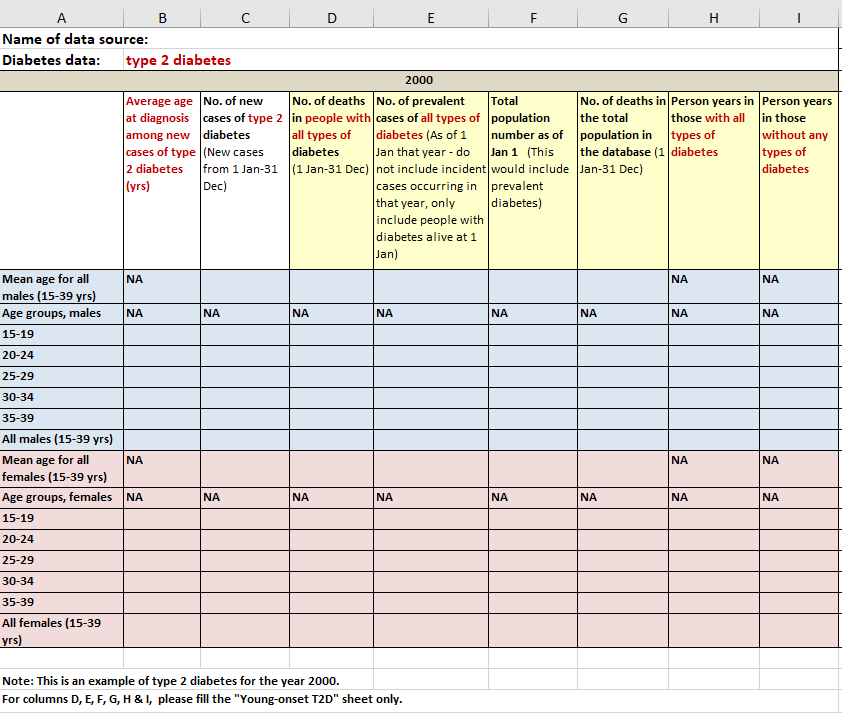
**References**

1. Magliano DJ, Chen L, Islam RM, et al. Trends in the incidence of diagnosed diabetes: a multicountry analysis of aggregate data from 22 million diagnoses in high-income and middle-income settings. *Lancet Diabetes Endocrinol* 2021; **9**(4): 203-11.

2. Carstensen B. Age-period-cohort models for the Lexis diagram. *Stat Med* 2007; **26**(15): 3018-45.

3. Lee RD, Carter LR. Modeling and Forecasting U.S. Mortality. *J Am Stat Assoc* 1992; **87**(419): 659-71.

**Attachment 1. Data extraction Template**



Notes: This is an example of data extraction required for type 2 diabetes for a single year of 2000.

**Attachment 2. Data source description for the incidence trends of young-onset type 2 diabetes project**

Please answer the questions below, so that we have an accurate description of each of the data sources used for this project. These answers are very important in order to fully understand any biases that may exist. If you have previously provided us with all of this information, there is no need to repeat it here. If there has been some update on the last data source that you provided to us, we would like you to update the relevant information if needed.

1. Please provide the name and description (e.g. national register, insurance).
2. Please describe the nature of the population that the people with diabetes come from (e.g. national population, people in employment, national insurance scheme, and if there are any geographic restrictions).
3. Please indicate if, within the geographic area that that the data source serves, there are groups of people who are systematically missed from the data source (e.g. uninsured, insured with a different insurer, users of private health care, veterans).
4. Please give a description of how diabetes is defined (e.g. diagnostic code, use of glucose-lowering drugs, blood glucose/HbA1c, or combination of fields).
5. Please explain what criteria are used to determine diabetes type (i.e. T1DM and T2DM) in your database.
6. Please explain if and how you have excluded or identified gestational diabetes for this data extraction.
7. Please explain how you derive the number of deaths (mortality counts) in your database, e.g. via linkage to national death registries?
8. What year was the database first established?
9. Please estimate the percentage (or percentage range) of completeness of the capture of diabetes cases. If this differs between type 1 and type 2 diabetes as recorded in your database, please provide two estimates.
10. Have you conducted any studies to assess the completeness of diabetes capture? If so, can you please provide the reference or the document?
11. Other relevant information about the database, if any.