

ARIC Manuscript Proposal # 2820

PC Reviewed: 8/9/2016
SC Reviewed: _____

Status: _____
Status: _____

Priority: 2
Priority: _____

1.a. Full Title: Is it possible to optimize the definition of prediabetes?

b. Abbreviated Title (Length 26 characters): Optimal definition of prediabetes

2. Writing Group:

Writing group members: The manuscript will be published under “The Obesity, Diabetes and Cardiovascular Disease Collaboration”. A writing committee will be formed when data collection is complete.

1. Rachel Huxley, Curtin University, Australia, rachel.huxley@curtin.edu.au.

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. Crystal Lee [please confirm with your initials electronically or in writing]

First author: Crystal Lee

Address: School of Public Health, Curtin University, GPO Box U1987, Perth, WA 6845, Australia.

Phone: +618 92662470
E-mail: crystal.lee@curtin.edu.au

Fax:

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

Name: June Stevens

Address: 2207 McGavran-Greenberg Hall, CB #7461, Chapel Hill, NC 27599, USA

Phone: (919) 966-7348
E-mail: june_stevens@unc.edu

Fax:

3. Timeline:

October 2016 – Finalize data collection

March 2017 – Complete data analysis

August 2017 – Complete manuscript preparation and submission

4. Rationale:

The concept of borderline diabetes or ‘pre-diabetes’ has been in use for nearly 50 years (1) but there is still confusion as to whether it is itself a disease condition or simply a risk factor for diabetes. Moreover, there is no global consensus as to the optimal definition for prediabetes with the World Health Organization (WHO) (2) and the American Diabetes Association (ADA) (3) adopting different guidelines. Finally, only a small proportion of individuals with prediabetes actually progress to overt diabetes - although the actual percentage and the time it takes to convert from prediabetes to overt diabetes remains unclear.

The concept of ‘pre-diabetes’ – a term used to indicate a high-risk state for future development of diabetes - has been in medical lexicon for nearly 50 years (1). Its diagnosis is based on one of two measures either fasting blood glucose (FBG) or 2-hour blood glucose following a standardised oral glucose tolerance test. Yet, or perhaps because of its longevity, the definition of prediabetes has changed over time in parallel with the accruing epidemiological evidence of the relation between blood glucose and incident diabetes and other important outcomes such as all-cause mortality.

In 2003 the ADA controversially, and in isolation from other organisations, changed the criteria for prediabetes based on FBG from 6.1-6.9 mmol/L to 5.6-6.9 mmol/L (3). The rationale behind the move – which overnight resulted in an additional 25 million US adults being considered as having pre-diabetes – was based on examination of the predictive relationship between FBG and incident diabetes from four study populations using unpublished data (the Pima Indian, Hoorn, San Antonio and Mauritius cohorts). From these four cohorts the values for FBG which equalled the maximum sum of specificity and sensitivity for discriminating diabetes were deduced. These values ranged from between 5.2–5.7 mmol/L and thus, the ADA subsequently recommended reducing the FBG threshold to 5.6 mmol/L.

In stark contrast, the WHO did not find the evidence sufficiently compelling to warrant a lowering of the threshold. One of the methodological issues concerning a lowering of the lower FBG threshold to 5.6 mmol/L is the assumption that the relationship between FBG and incident diabetes is similar across populations – which judging by our preliminary investigations may not be true. It therefore remains unknown which values of FBG represent the ‘optimal’ range for discriminating those individuals at highest risk of developing diabetes in a broad range of populations and in need of further screening.

As with FBG, there has been a similar reliance on data from select populations to derive cut-points for defining impaired glucose tolerance (IGT). The validity of the current definition for IGT is based on the assumed risk of developing diabetes (or other adverse outcomes) associated with 2-hour plasma glucose levels. The 2 hour post-load plasma glucose cut-point of 7.8 mmol/L for defining IGT was derived based mainly on data collected nearly four decades ago from the Pima Indian study – a cohort with a very high prevalence of diabetes and limited generalizability (4). In that study, the incidence of diabetes increased from less than 2% per year in those with 2-hour plasma glucose levels of <5.6 mmol/L to nearly 7% in those with values for between 7.8-11.0 mmol/L. There has been limited epidemiological investigation for the validity of using the lower cut-point of 7.8mmol/L for defining IGT and most subsequent studies have shown that there is no threshold in the association (5). Despite these obvious limitations there has been little

impetus to investigate whether the current definition for IGT needs to be revisited with commentators suggesting that there “has been no specific interest in changing what is now a well-established diagnostic tool” (6). Indeed, the 2006 WHO report “Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycaemia” admitted that there was a lack of evidence to substantiate the current values for IGT concluding that “although there are limited data to support the current 2hr plasma glucose value used to define IGT, the current cut-point seems to be operationally adequate” (2).

5. Main Hypothesis/Study Questions:

Main study question:

Can the definition of prediabetes be optimized to better predict future diabetes risk?

Hypothesis:

The ‘optimal’ definition of prediabetes will vary depending on what are considered to be acceptable levels of sensitivity and specificity.

Specific research aim:

To determine the optimal definition for prediabetes by 1) comparing the ability of the ADA and WHO definitions of prediabetes based on impaired fasting glucose (IFG) at baseline to discriminate individuals who subsequently develop diabetes; 2) examining whether the lower threshold (7.8 mmol/L) for defining IGT is optimal for the discrimination of future diabetes overall and across populations.

6. Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).

Study design

Individual participant data analysis of longitudinal studies included in an international data pooling collaboration

Data source

This study will utilize data from the Obesity, Diabetes and Cardiovascular Disease Collaboration (ODCDC). The ODCDC, based at Curtin University, Australia, is an international data pooling collaboration of longitudinal studies established to address outstanding issues of epidemiological and clinical importance regarding obesity, diabetes and cardiovascular disease in diverse populations. The ODCDC database was developed from a cleaned and coded dataset provided by investigators of the Collaborative Study of Obesity and Diabetes in Adults (CODA) after obtaining permission of data use from investigators of each of the prospective studies included in CODA (7). With funding available through a National Health and Medical Research Council project grant (2016-2018), we have begun contacting investigators of existing and potential studies to provide data for all available study visits in order to develop a more comprehensive and informative dataset for the Collaboration. To date, 23 studies with more than 150,000 individuals from 11 countries have either provided or agreed to provide data to the collaboration.

Variables of interest

Baseline – subject ID, age, sex, date of birth, exam date, education, race, systolic blood pressure, diastolic blood pressure, antihypertensive medication, diabetes treatment type, triglycerides, total cholesterol, HDL cholesterol, fasting insulin, fasting time, fasting plasma glucose, 2 hour plasma glucose, height, weight, waist circumference, hip circumference, smoking status, physical activity, self-report/known diabetes, family history of diabetes, cigarettes smoked per day, coronary heart disease status, gestational diabetes, age when first diagnosed with diabetes.

All follow-up visits – variables that can be used to determine diabetes incidence e.g. fasting plasma glucose, 2 hour plasma glucose, self-report/known diabetes, diabetes treatment type, exam date, follow up time for diabetes, age when first diagnosed with diabetes.

Statistical analysis

Individuals will be classified as having prediabetes based on FBG at baseline using the two different cut-off values recommended by the ADA and WHO to indicate IFG. The risk of incident diabetes in these individuals using these definitions will be estimated using Cox regression models adjusting for other covariates. Model performance measures such as Harrell's C that measures the concordance between predicted and the actual survival will be used to examine the model discrimination power. The differences in the model performance will also be examined across different populations.

To determine whether there is a more optimal FBG value for defining pre-diabetes than those specified by the WHO and ADA we will then examine the continuous nature of the relationship. Continuous measures of FBG will be modelled to examine its association with incident diabetes using a survival analysis model framework. With the use of smoothing splines for modelling the dose-response trend in time to diabetes, we will explore if there is a threshold effect between FBG and incident diabetes. If so, the threshold level will be estimated using parametric non-linear models and used to define a new cut-point for FBG and subsequently report on its predictive capability and compare with the results from the above models using standards methods such as comparison of area under the curve and Goodness of Fit. We will use the same methodology to examine the optimal cut-point for IGT. The analysis will be conducted in the overall population and then separately by subgroups defined by age, gender and race/ethnicity/region to determine if and how the cut-point values for FBG and 2-hour plasma glucose vary.

7.a. Will the data be used for non-CVD analysis in this manuscript? Yes No

b. If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = "CVD Research" for non-DNA analysis, and for DNA analysis RES_DNA = "CVD Research" would be used? Yes No

(This file ICTDER has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? Yes No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ___ Yes ___ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at: <http://www.csc.unc.edu/ARIC/search.php>

Yes No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?

According to the ARIC proposals on diabetes and body mass index, we have identified the following manuscript proposals to be partly related to our proposed manuscript

1. MS proposal by A. McNeill – Evaluation of the FINDRISC diabetes score to identify individuals at high risk for diabetes among middle-aged Caucasian and African.
2. MS proposal by M.I. Schmidt – Prediction of diabetes mellitus and impaired glucose tolerance in middle-aged adults: The Atherosclerosis Risk in Communities study
3. MS proposal by L. Zimmermann – Modifiable risk factors associated with progression from incident impaired fasting glucose to diabetes mellitus type 2: A pooled
4. MS1431 - Hemoglobin A1c, glucose and incident diabetes: the ARIC study
5. MS1271 – Does the diabetes case definition affect the relationship of incident diabetes and traditional and novel risk factors?

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data? ___ Yes No

11.b. If yes, is the proposal

- A. primarily the result of an ancillary study (list number* _____)**
 B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* _____)

*ancillary studies are listed by number at <http://www.csc.unc.edu/aric/forms/>

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is your responsibility to upload manuscripts to PubMed Central whenever the journal does not and be in compliance with this

policy. Four files about the public access policy from <http://publicaccess.nih.gov/> are posted in <http://www.csc.unc.edu/aric/index.php>, under Publications, Policies & Forms. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

13. Per Data Use Agreement Addendum, approved manuscripts using CMS data shall be submitted by the Coordinating Center to CMS for informational purposes prior to publication. Approved manuscripts should be sent to Pingping Wu at CC, at pingping_wu@unc.edu. I will be using CMS data in my manuscript ____ Yes No.