ARIC Manuscript Proposal #4002

PC Reviewed: 2/8/22 **Status: _____** **Priority: 2**
SC Reviewed: _____ **Status: _____** **Priority: ____**

1.a. Full Title: **Adolescent and adulthood BMI and cancer risk using obese-year metrics (ABACus 2)** [PROSPERO CRD42021238270)] **and waist-circumference year metrics (ABACus 3)** [PROSPERO CRD42021287001)]

b. Abbreviated Title (Length 26 characters): **Body fatness metrics and cancer**

2. Writing Group:

Writing group members: Nadin Hawwash, Lucy Osborne, Glen P. Martin, Matthew Sperrin, Andrew G. Renehan, Elizabeth A. Platz, Corinne E. Joshu and up to 2 other interested ARIC investigators

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

First author: **Nadin Hawwash**
Address: 555 Wilmslow Road, Manchester M20 4GJ

Phone: 07437923175 Fax:
E-mail: nadine.hawwash@postgrad.manchester.ac.uk

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: **Professor Elizabeth Platz**
Address: Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health 615 N. Wolfe St., Room E6132, Baltimore, MD 21205

Phone: 410.614.9674 Fax: 410.614.2632
E-mail: eplatz1@jhu.ac.uk

3. Timeline:

4. Rationale:

Body fatness, commonly approximated by body mass index (BMI), is an established risk factor for 13 obesity-related cancers (1)**. Excess body fatness is the second commonest cause of cancer after smoking in western countries** (2)**. Several biological mechanisms may explain these associations including insulin resistance, altered sex hormone balance, subclinical inflammation** (3) **and obesity associated gastroesophageal reflux disease** (4)**. But much of the current evidence linking excess body fatness and cancer risk is based on time-updated BMI measures. One unanswered question is whether cumulative exposure to body fatness through individual obese/overweight year metrics provide a more informative measure of body fatness exposure than a current BMI reading on the related cancer risk. This methodology is comparable to tobacco smoking pack-years and is familiar in the cardio-metabolic literature** (5) **but scarce in cancer epidemiology** (6)**. Another unanswered question is whether the waist circumference (WC) years metric may provide a more precise indicator of body fatness exposure. To address these unanswered questions, we will determine individual obese-year, overweight-year and waist circumference year metrics and the related cancer risk. Our primary focus is to optimize the expression of body fatness by accounting for the life-course exposure so that these optimized measures can be used in risk prediction models for clinical use. We aim to complete two inter-related projects which look at two distinct but related research questions as shown in Figure 1.**

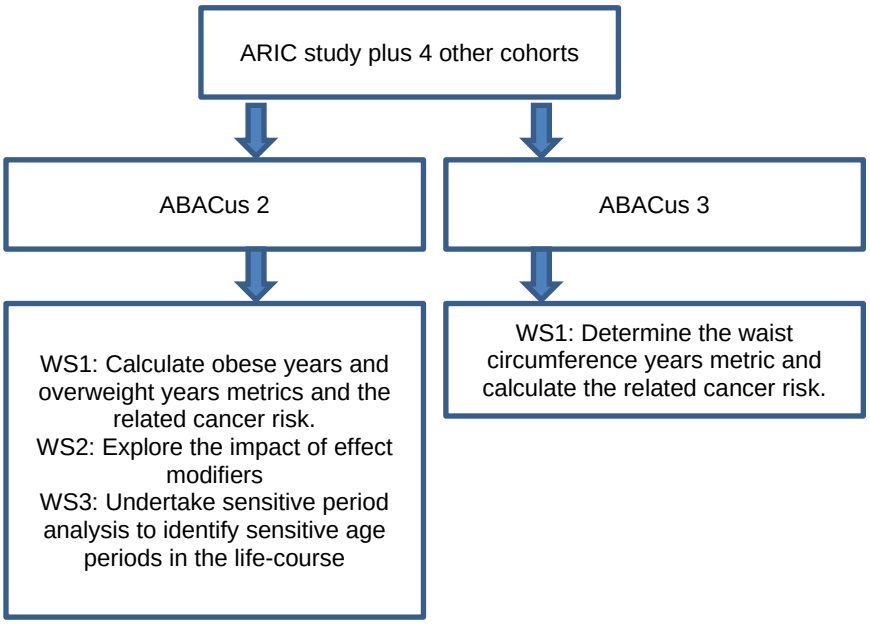

Figure 1: Study design flow diagram.

ABACus (Adulthood BMI exposure And Cancer risk using clustering methodologies) is a set of projects out if the University of Manchester, United Kingdom. ABACus (7) **included ARIC data along with other cohort data and performed latent class trajectory modelling**

(LCTM) with cancer incidence as endpoints. LCTM was resource and computationally intensive and these models performed poorly, such that no publication is planned from these data. In ABACus 2 and 3, we are returning to mathematically simpler models of obesity exposure.

In ABACus 2, we will primarily focus on using the obese-years and overweight years metric and calculating the related cancer risk. In ABACus 3, we will focus on calculating the waist circumference years metric using repeated waist circumference measures and calculating the related cancer risk as well as other major diseases including cardiovascular disease, diabetes, mental health problems and rheumatoid arthritis. Cohorts with repeated waist circumference data and the outcomes of interest will be included. We will then analyse the time-to-incidence related cancer risk using a Cox proportional hazards regression model. In ABACus 2, as a separate analysis, we will investigate the possibility of effect modification by factors including age, gender, smoking, HRT, baseline waist circumference, use on the body fatness and cancer link. As another separate analysis in ABACus 2, we will assess for 'sensitive age' periods for exposure using sensitive period analysis (SPA). Our work aims to inform the design, implementation, and optimisation of targeted prevention strategies, establish changes in public health policy and enable translation to clinical practice by the improvement of risk prediction models.

5. Main Hypothesis/Study Questions:

Overall aim:
In ABACus 2, we aim to investigate the use of obese-years and overweight-years, analogous to cigarette pack-years, as a cumulative exposure to body fatness with cancer incidence as the primary outcome measure. To do this, first we will analyse each cohort separately from the ABACus 2 consortium (currently consisting of NIH-AARP Diet and Health Study (NIH-AARP), European Prospective Investigation into Cancer and Nutrition (EPIC), Prostate, Lung, Colorectal, Ovarian Cancer Screening Trial (PLCO), Women's Health Initiative (WHI) and Atherosclerosis Risk in Communities study (ARIC). Then, we will perform an Individual Participant Data (IPD) meta-analysis across cohort studies. In ABACus 3, we aim to investigate the use of the WC-year metric as an exposure with cancer incidence and cancer mortality as the primary outcome measure. Secondary outcome measures for ABACus3 include other major diseases such as cardiovascular disease, diabetes, mental health problems and rheumatoid arthritis. We will undertake the analysis in ABACus 3 using the ABACus 3 consortium (currently consisting of ARIC and EPIC).

Including data from ARIC in the ABACus 2 consortium, we aim to:
Aim 1: Determine individual time-updated obese-year metrics, for instance, duration of obesity, degree of obesity, obese-years and cumulative obese-years using repeated BMI measurement, for instance for each cohort in the ABACus 2 consortium and estimate the association between these metrics and risk of known obesity-associated cancers and risk of cancers not thought to be obesity associated. Both relative and absolute risks will be estimated. Metrics may be further modified to include BMI-years using continuous BMI rather than categorising BMI and then calculating exposure using a metric.

Aim 2: Cumulative obesity degree and duration and their related cancer risk will be assessed.
Aim 3: Explore the impact of effect modifiers on the body fatness and cancer link.
Aim 4: Use Sensitive Period Analysis (SPA) approaches to fit separate models to assess for 'sensitive age periods' during an individual's lifetime – for example, early adulthood.
Aim 5: Pool relevant effect estimates from aim 1-3 across cohorts using a 2 stage individual participant data (IPD) meta-analysis to determine the causal effects to the cancer risk.

Including data from ARIC in the ABACus 3 consortium, we aim to:
Aim 6: Determine individual WC-year metrics using binned WC measures and estimate the association between these metrics and cancer risk as well as other major diseases including cardiovascular disease, diabetes, mental health problems and rheumatoid arthritis]

Aim 7 Cumulative measures in ABACus 2 and 3 will be compared with time-updated BMI to establish whether these measures provide more information on cancer risk.

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).

Data analysis strategy:
Within each cohort of the ABACus 2 consortium, we will apply the following exclusion criteria: (i) exclude individuals over 80 years of age at study entry, (ii) exclude individuals with prevalent cancers at baseline, and (iii) exclude individual BMI observations that are outside the limit of <15 kg/m^2 and >60 kg/m^2 given these are clinically implausible values. Following this, we will undertake two separate analyses: (1) include those individuals that have three or more BMI measurements at three time points throughout that individual's lifespan and (2) include those individuals with at least one BMI measurement and use a mixed effect model to predict BMI. Repeated BMI measures are required to calculate the obese-years.

In ABACus 3, repeated WC measures are required to calculate the waist circumference years metric. Our scoping search suggests that there are no studies of this type with cancer as the primary outcome measure – so this is a novel question. Unlike BMI, which is recorded in some cohorts by recall, and in other cohorts prospectively or both, we anticipate that repeated WC measures will be measured by trained staff in ARIC.

In ABACus 2 and ABACus 3, our primary outcome will be incidence of obesity-related cancer, and as a secondary outcome we will consider the incidence of non-obesity related cancer as a secondary outcome. Obesity-related cancers will be identified using the corresponding ICD-10 codes: C15.5, C15.8 (Oesophagus – lower third), C18.0-18.9, C19.9, C20.0 (Colorectal), C22.0 (Liver), C23.9 (Gallbladder), C25.0-25.9 (Pancreas), C50.0-50.9 (Breast*), C54.0-54.9, C55.9 (Corpus Uteri/Endometrial), C56.9 (Ovary), C64.9 (Kidney), C16.0 (Gastric cardia), C70.0, C70.1, C70.9 (Malignant meningioma), C73.0, C73.9 (Thyroid), C61 fatal prostate cancer, and multiple myeloma using morphology code C90.0.

Outcomes may be refined further in a sensitivity analysis e.g., when identifying hormone-related obesity-associated cancers oesophageal cancers will be excluded.

In all analyses, we will start by performing exploratory analyses of the data to describe the population in each cohort; for instance, the age ranges (and mean/median), proportion of men and women, number of HRT users and number of smokers. The data from all cohorts will be cleaned based on pre-specified steps previously considered (in the previous ABACus project) and filtered to remove any outcomes that are implausible such as the presence of prostate cancer in women (along with other data cleaning steps, such as variable coding and missing data handling).

Using codes written within a preliminary analysis of English Longitudinal Study of Ageing (ELSA), we will then calculate the obese-years metrics across each individual cohort. We will also calculate the waist circumference (WC) years metric. After calculating the obese/overweight/WC-years metrics, we will fit Cox proportional hazards regression models (within each cohort separately) to model the association between the cumulative exposure to body fatness and cancer risk. Cox proportional hazards regression models from ABACus 2 will also be used to identify effect modifications by prespecified factors: age, race, smoking and HRT and waist circumference on the association between body fatness and cancer. We will examine the proportional hazards assumption and include time-varying coefficients where needed (e.g., through splines, or stratification). To examine whether the association between the obesity duration and cancer risk may be affected by imputation for missing BMI, a sensitivity analysis will be performed. In particular, an analysis will be undertaken only using participants that had no missing BMI values in any examinations.

In ABACus 2, single estimates of the exposure-outcome association will be derived from each cohort, and we will apply random-effect meta-analysis to combine these estimates (i.e., 2-stage meta-analysis). The ELSA dataset does not include sufficient data on cancer outcomes to perform this analysis, for example, a cancer outcome is included but on initial search there is no reference to the type of first cancer diagnosed. The rest of the datasets in the ABACus2 consortium (NIH-AARP, WHI, PLCO, ARIC and EPIC) have sufficient data on cancer incidences and outcomes. We will also work on adapting the obese year metric to create further measures that overcome limitations observed from the metric such as overcoming the limitation of using an arbitrary threshold for obese BMI category of ≥ 29.9 kg/m^2 by using overweight year metrics.

We will then focus in ABACus 2 on the impact of effect modifiers such as age, gender, smoking and hormone replacement therapy (HRT) on the body fatness and cancer link to outline modifiable risks. Other variables of interest include dietary habits, physical activity, alcohol consumption, aspirin use, social and educational levels and other medications which may alter body fatness such as statins and diabetic medications such as metformin. Effect modification will be tested for by stratifying individuals in each potential effect modifier of interest and analysing the related cancer risk using Cox proportional hazards models.

As a separate workstream, in ABACus 2, we will undertake SPA to identify sensitive periods in the life-course where body fatness exposure at certain periods is related to a stronger association with cancer incidence. Initial careful planning of study methodology will be necessary to avoid incorporation of immortal time bias. The main modelling steps to be used will be discussed. First, the index date of the study must be determined as the methods of analyses used in SPA vary depending on where the time point of interest (TOI) at which the exposure is quantified is relative to the index date, as the nature of the immortal time varies. The index date will be defined as the period at which individuals are considered to be at risk of disease incidence which is the date of study entry and also the start of study follow-up. Second, fixed time periods of interest (POI) will be selected. The POIs will be predefined by clinically meaningful ages (for example, menopause in women) or data driven POI and analysed across the life-course to compare the association between the exposure quantified at the POI on the outcome calculated at any time after the index date. Third, the cohort will be subset to only include individuals that survived till the POI and the start of follow-up. Sub-setting the population at the TOI will allow the exposure to be quantified at the TOI. Fourth, left truncation will be included to avoid immortal time bias if the POI is prior to the index date. Fifth, Cox regression models will be used to model the association between the exposure and outcome. Sensitive period analysis will help focus research on specific periods in the life-course where the association between body fatness and cancer incidence is stronger to identify causal determinants and underlying biological mechanisms involved. For WS3, SPA for each cohort in the ABACus 2 consortium will be performed.

Typical required variables

- **ID**
- **Age at baseline (study entry)**
- **Age**
- **Gender**
- **Weight kg**
- **Height cm**
- **Baseline hip circumference**
- **Baseline waist circumference**
- **Baseline waist to hip ratio**
- **BMI at baseline, kg/m^2**
- **Repeated waist circumference measures by visit**
- **Repeated body mass index (or weight and height) measures by visit**
- **Ages when repeated BMI measures were taken**
- **Ages when repeated waist circumference measures were taken**
- **Ever diagnosed with diabetes at study entry**
- **Ever diagnosed with cardiovascular disease at study entry**
- **Time-updated and baseline smoking status (current, former, never)**
- **Time-updated and baseline postmenopausal hormone therapy status (current, former, ever)**
- **Race/Ethnicity**
- **Time-updated and baseline alcohol units consumed per week (g/d)**
- **Baseline highest educational level**

- **Baseline key baseline dietary components e.g., meat consumption from FFQ**
- **Baseline Cancer family history (main sites; site-specific)**
- **Cancer diagnosis (Yes/No)**
- **Cancer incidence at baseline (study entry)**
- **First cancer site (ICD-10 codes)**
- **Age of first cancer diagnosis**
- **Age of first obesity related cancer diagnosis (if available but can be derived from ICD-10 codes)**
- **Age of first non-obesity-related cancer diagnosis (if available but can be derived from ICD-10 codes)**
- **Age of end of follow-up**
- **Obesity-related cancer diagnosis (if available but can be derived from ICD-10 codes)**
- **Non-obesity related cancer diagnosis (if available but can be derived from ICD-10 codes)**
- **Prevalent cancers at baseline**
- **Baseline age at menarche**
- **Baseline and time-updated recreational physical activity**
- **Baseline and time-updated hip circumference (if available)**
- **Baseline and time-updated waist to hip ratio (if available)**
- **Treatments that may be associated with cancer risk such as statins, aspirin use, NSAID use and diabetes medication e.g., metformin.**
- **Age of death**
- **Cause of death**
- **Age of menopause**
- **Time-updated and baseline recreational physical activity**
- **Cancer mortality**
- **Age of cancer mortality**
- **Age at menarche/puberty**
- **Regular Ibuprofen user in the last 12 months**
- **Regular aspirin user in the last 12 months**
- **Highest educational level**
- **Family history of cancer**
- **Reproductive history: parity and age of first pregnancy**
- **Martial status**
- **Duration of hormone replacement therapy use**
- **Diet: total energy kcal/d; fiber g/d; folate mcg/d; calcium mg/d; fruit and vegetables servings/d; red meat g/d**

Although the systematic review in ABACus 3 looks into mental health, rheumatoid arthritis and CVD no data on these are required apart from those listed above.

Baseline variables will be used to allow for harmonisation across cohorts and undertake an IPD meta-analysis. Also the aim is to build a simple model which may in the future be used in a clinical setting.

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? ____ Yes _✓___ No

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ____ Yes ____ No
(The 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 current derived consent file ICTDER05 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.cscc.unc.edu/aricproposals/dtSearch.html

__✓____ 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)?

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

Adulthood BMI latent class trajectory modelling (LCTM) and obesity- related cancers in the ARIC cohort study
#1792 (The influence of obesity, diabetes, and associated metabolic perturbations on cancer risk, Joshu CE)
#1766 (Weight change and cancer risk: The Atherosclerosis Risk in Communities Study (published: Han X, Stevens J, Truesdale KP, Bradshaw PT, Kucharska-Newton A, Prizment AE, Platz EA, Joshu CE. Body mass index at early adulthood, subsequent weight change and cancer incidence and mortality. Int J Cancer. 2014 Dec 15;135(12):2900-9. doi: 10.1002/ijc.28930. Epub 2014 May 8. PubMed PMID: 24771654; PubMed Central PMCID: PMC4192093)
Enhancing the Infrastructure of the Atherosclerosis Risk in Communities (ARIC) Study for Cancer Epidemiology Research: ARIC Cancer

11.b. If yes, is the proposal
__✓_ A. primarily the result of an ancillary study (list number*2011.07 & 1995.04)

___ 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 https://sites.cscc.unc.edu/aric/approved-ancillary-studies

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.cscc.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.

References

1. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness and Cancer — Viewpoint of the IARC Working Group. N Engl J Med. 2016 Aug 25;375(8):794–8. Available from: http://www.nejm.org/doi/10.1056/NEJMsr1606602
2. Arnold M, Pandeya N, Byrnes G, Renehan AG, Stevens GA, Ezzati M, et al. Global burden of cancer attributable to high body-mass index in 2012: A population-based study. Lancet Oncol. 2015 Jan 1;16(1):36–46. Available from: www.thelancet.com/oncology
3. Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: New mechanistic insights from epidemiology [Internet]. Vol. 15, Nature Reviews Cancer. Nature Publishing Group; 2015. p. 484–98. Available from: https://pubmed.ncbi.nlm.nih.gov/26205341/
4. El-Serag H. The association between obesity and GERD: A review of the epidemiological evidence [Internet]. Vol. 53, Digestive Diseases and Sciences. NIH Public Access; 2008. p. 2307–12. Available from: /pmc/articles/PMC2827866/
5. Abdullah A, Amin FA, Stoelwinder J, Tanamas SK, Wolfe R, Barendregt J, et al. Estimating the risk of cardiovascular disease using an obese-years metric. BMJ Open. 2014;4(9). Available from: https://pubmed.ncbi.nlm.nih.gov/25231490/
6. Arnold M, Charvat H, Freisling H, Noh H, Adami HO, Soerjomataram I, et al. Adult overweight and survival from breast and colorectal cancer in swedish women. Cancer Epidemiol Biomarkers Prev. 2019 Sep 1;28(9):1518–24. Available from: http://cebp.aacrjournals.org/
7. Adulthood BMI trajectories and cancer using clusters; a pooled analysis of cohorts with repeated BMI measures [ABACus project] [Internet]. Available from: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=79621