Population Biobanks: scientific platforms for the investigating of the role of lifestyle, nutrition, metabolism and genetics in chronic disease aetiology

Elio Riboli, MD, MPH, ScM, HonFPH, FMedSci
Professor in Cancer Epidemiology and Prevention
Director, Imperial School of Public Health
Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. BiobankQatar
9. Conclusions
The “Causes of the Causes” of Chronic Disease (WHO)

Deaths Attributable to Lifestyle Risk Factors (Millions)

- High Blood Pressure
- Tobacco Use
- High Blood Glucose
- Physical Inactivity
- High Body Mass Index
- High Cholesterol
- Alcohol Use
- Low Fruit & Vegetable Intake

Low- medium income countries

High income countries

Ezzati M & Riboli E
New Engl J Med 2013
Millions of Overweight People by Region 1990-2008

Ezzati M & Riboli E
New Engl J Med 2013
Millions of Obese People by Region 1990-2008

Ezzati M & Riboli E
New Engl J Med 2013
TWO BILLION PEOPLE OVERWEIGHT OR OBESE IN THE WORLD IN 2008

Ezzati M & Riboli E
New Engl J Med 2013
Can Noncommunicable Diseases Be Prevented? Lessons from Studies of Populations and Individuals

Majid Ezzati* and Elio Riboli

Noncommunicable diseases (NCDs)—mainly cancers, cardiovascular diseases, diabetes, and chronic respiratory diseases—are responsible for about two-thirds of deaths worldwide, mostly in low- and middle-income countries. There is an urgent need for policies and strategies that prevent NCDs by reducing their major risk factors. Effective approaches for large-scale NCD prevention include comprehensive tobacco and alcohol control through taxes and regulation of sales and advertising; reducing dietary salt, unhealthy fats, and sugars through regulation and well-designed public education; increasing the consumption of fresh fruits and vegetables, healthy fats, and whole grains by lowering prices and improving availability; and implementing a universal, effective, and equitable primary-care system that reduces NCD risk factors, including cardiometabolic risk factors and infections that are precursors to NCDs, through clinical interventions.

Fig. 1. Tobacco smoking is the most important risk factor for lung cancer and also has harmful effects on other NCDs.
Lung Cancer Death Rate per 100,000 per year 1950-2010
Cardiovascular Disease Death Rate per 100,000/year 1950-2010

Ezzati M and Riboli E 1485-1487

SCIENCE VOL 337 21 SEPTEMBER 2012
Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. BiobankQatar
9. Conclusions
DESIGN of Population Biobank

Prospective cohort study

Exposure → Biomarkers of exposure (incl. –omics) → Intermediate –omics biomarkers of early effects → Disease

Baseline & Follow-up Assessments → “Lifecourse studies”
EPIC: European Prospective Investigation on Cancer and Chronic Diseases

**BASELINE**
- Subjects recruitment
- Questionnaires data on life style, personal history etc.
- Anthropometry data
- Blood/DNA collection
- Data Base & Biorepository

**FOLLOW-UP:**
- Cancer diagnosis
- Vital status
- Causes of death
- Changes in Lifestyle

**AETIOLOGICAL STUDIES**

Development of common/standardized Nutrient and lifestyle Data Bases
Setting up of lab facilities for sample handling / DNA extraction etc
## EPIC: European Prospective Investigation on Cancer and Chronic Diseases

<table>
<thead>
<tr>
<th>Country</th>
<th>Questionnaire</th>
<th>Q + Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>74 524</td>
<td>28 053</td>
</tr>
<tr>
<td>Italy</td>
<td>47 749</td>
<td>47 725</td>
</tr>
<tr>
<td>Spain</td>
<td>41 440</td>
<td>39 579</td>
</tr>
<tr>
<td>U.K.</td>
<td>87 942</td>
<td>43 141</td>
</tr>
<tr>
<td>Netherlands</td>
<td>40 072</td>
<td>36 318</td>
</tr>
<tr>
<td>Greece</td>
<td>28 555</td>
<td>28 483</td>
</tr>
<tr>
<td>Germany</td>
<td>53 091</td>
<td>50 678</td>
</tr>
<tr>
<td>Sweden</td>
<td>53 826</td>
<td>53 781</td>
</tr>
<tr>
<td>Denmark</td>
<td>57 054</td>
<td>56 131</td>
</tr>
<tr>
<td>Norway</td>
<td>37 215</td>
<td>31 000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>521 468</strong></td>
<td><strong>414 889</strong></td>
</tr>
</tbody>
</table>

- 30 ml venous blood:
  - 20 ml citrated + 10 ml dry

- 28 aliquots of 500 µl:
  - plasma 12 (red straws)
  - serum 8 (yellow straws)
  - buffy coat 4 (blue straws)
  - RBC 4 (green straws)

28 aliquots x 300,000 subjects = 8.4 Million aliquots stored, half in each EPIC centre, half at IARC-Lyon

Plus: 12 x 110,000 = 1.3 Million in Sweden and Denmark
Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. BiobankQatar
9. Conclusions
Relative Risk of Death among Men and Women in EPIC By BMI, Waist and Waist-to-Hip Ratio

Relative Risk of Death among Men and Women in EPIC
By Waist and Waist-to-Hip Ratio after Adjustment for BMI

Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. BiobankQatar
9. Conclusions
Postmenopausal BMI and Breast Cancer risk

- Sonnenschein 1999: 1.20 (1.08–1.34)
- Toniolo 1994: 1.20 (1.05–1.39)
- Galans 1998: 1.09 (1.01–1.16)
- Gapstur 1992: 1.08 (1.04–1.13)
- Wirfalt 2004: 1.07 (1.00–1.15)
- Saadatian-Elahi 2002: 1.07 (0.95–1.19)
- Barrett-Connor 1993: 1.06 (0.80–1.41)
- Tornberg 1994: 1.05 (1.01–1.09)
- Tullius 1997: 1.04 (0.98–1.10)
- Tehard 2004: 1.03 (1.00–1.07)
- Van den Brandt: 1.03 (0.97–1.09)
- Huang 1997: 1.02 (0.99–1.05)
- Rissanen 2003: 1.01 (0.85–1.21)
- Manjer 2001: 0.97 (0.88–1.07)
- Kaaks 1998: 0.97 (0.86–1.09)
- Jumaan 1999: 0.95 (0.87–1.03)

Summary estimate: 1.03 (1.01–1.04)

Relative risk, per 2 kg/m²
Postmenopausal BMI and Breast Cancer risk

Summary estimate:
3% increase in Breast Cancer risk per BMI unit
~ 15 - 30 % increase in incidence in obese compared to slim women
RR of breast cancer by weight gain and hormone use postmenopausal women in EPIC

Relative Risk

Weight gain (kg)

Current hormone use

+/-2 kg 2-15 kg >15 kg

no yes

1.00 1.20 1.49

3.06 2.50 2.67

1.00 1.20 1.49

3.06 2.50 2.67

Relative Risk

Weight gain (kg)
• WEIGHT GAIN SINCE AGE 20 IS STRONGER PREDICTOR OF BREAST CANCER RISK THAN ADULT WEIGHT
• HRT INCREASE BREAST CANCER RISK MORE IN LEAN THAN IN OBESE WOMEN
Serum estrone by BMI
EPIC postmenopausal women (n= 1171)
Serum SHBG by BMI
EPIC postmenopausal women (n = 1210)
Serum free estradiol by BMI
EPIC postmenopausal women (n= 1171)
Serum free testosterone by BMI level
EPIC postmenopausal women (n=1171)
Postmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 300,000 postmenopausal women, average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)
Postmenopausal Serum Sex Steroids and Breast Cancer Risk

EPIC, 300,000 postmenopausal women, average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)
Postmenopausal Serum Sex Steroids and Breast Cancer Risk

EPIC, 300,000 postmenopausal women, average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)
Postmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 300,000 postmenopausal women, average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)
Steroidogenesis pathway

- Cholesterol → CYP11A1
- Pregnenolone
- Progesterone → 3βHSD
- 17-hydroxy-progesterone → CYP17
- 17βHSD
- Androstenedione → CYP17

Female specific:
- CYP19
- Estrone
- Estradiol

Male specific:
- Testosterone
- Dihydrotestosterone
- 5α-reductase
- Estradiol

Inactive form in the circulation
- Estradiol SHBG

Active form in the cell
- Testosterone SHBG
- Androgen receptor

Active form in the nucleus
- Estrogen receptor
Steroidogenesis pathway

Cholesterol $\rightarrow$ CYP11A1
Pregnenolone $\rightarrow$ 3βHSD
Progesterone $\rightarrow$ CYP17
17-hydroxy-progesterone $\rightarrow$ CYP17
Androstenedione $\rightarrow$ CYP17

Female specific: CYP19
Male specific: CYP19

17βHSD

Estrone $\rightarrow$ CYP19
Testosterone $\rightarrow$ CYP19

Inactive form in the circulation

Active form in the cell

Active form in the nucleus

Testosterone

SHBG

Estrogen receptor

Androgen receptor

...it was a nice dream!
The BPC3 – NIH-NCI Cohort Consortium: (EPIC, ACS, Harvard Cohorts, MEC)

- Only < 10% of between subjects variability in hormone level was genetically regulated.

- 90% or more of the variance was due to known and unknown lifestyle and metabolic factors.

- None of the SNPs in genes involved in steroidogenesis and sex hormones bioavailability was associated with breast cancer risk.

Beckman L J Clin Endocr Metabolism 2011
BMI and Breast Cancer

BEFORE MENOPAUSE

Galans 1998
Tonilo 1994
Manjer 2001
Tuilnus 1997
Karaks 1998
Sonenschein 1999
Huang 1997
Saadatian-Elahi 2002
Vatten 1992
Tehard 2004
Welderpass 2004
Ahlgren 2004
Tomberg 1994
Rissanen 2003
Summary estimate

Relative risk, per 2 kg/m²

BEFORE MENOPAUSE

Sonenschein 1999
Tonilo 1994
Galans 1998
Gapsur 1992
Wirfalt 2004
Saadatian-Elahi 2002
Barrett-Connor 1993
Tornberg 1994
Tuilnus 1997
Tehard 2004
Van den Brandt
Huang 1997
Rissanen 2003
Manjer 2001
Kaaks 1998
Jumaan 1999
Summary estimate

Relative risk, per 2 kg/m²
BMI and Breast Cancer

BEFORE MENOPAUSE

AFTER MENOPAUSE

Relative risk, per 2 kg/m²
PREmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 90,000 women average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)

(Kaaks et al. JNCI, 2005 May 18;97(10):755-65)
PREmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 90,000 women average follow-up 8 yeras

Quintile of Serum levels: from LOW (1) to HIGH (5)

(Kaaks et al. JNCI, 2005 May 18;97(10):755-65)
PREmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 90,000 women average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)

(Kaaks et al. JNCI, 2005 May 18;97(10):755-65)
PREmenopausal Serum Sex Steroids and Breast Cancer Risk
EPIC, 90,000 women average follow-up 8 years

Quintile of Serum levels: from LOW (1) to HIGH (5)

(Kaaks et al. JNCI, 2005 May 18;97(10):755-65)
EPIC: Colorectal Cancer
Combined effect of consumption of red meat and fibre

(Norat T. et al. JNCI, 2005)
EPIC: Colorectal Cancer
Combined effect of consumption of fish and fibre

Norat T. et al. (JNCI, 2005)
Obesity and Colorectal cancer
Men and Women, EPIC

(Pischon et al. JNCI 2006)
Obesity and Colorectal cancer
Men and Women, EPIC

(Pischon et al. JNCI 2006)
Insulin Resistance and risk of colon cancer by serum levels of C-Peptide, IGFBP-1 and IGFBP-2.

NYU Women Health Study

Kaaks et al. JNCI, 2001
Insulin Resistance and risk of colon cancer by serum levels of C-Peptide, IGFBP-1 and IGFBP-2.
NYU Women Health Study and EPIC

Kaaks et al. JNCI, 2001
Interaction of Serum 25OHD Concentration and Dietary Calcium Intake Level on Risk of CRC

p interaction = 0.154

Tertile of Dietary Calcium Intake (mg/day)

Odds Ratio of CRC Risk

p<0.05

Jenab et al. – in preparation
Colorectal ca: a multifactorial causal model

Diet

- Fruit
- Vegetables
- Whole cereals
- Fibre
- Vid D
- Calcium
- B Vit / Folates?
- Red Meat
- Processed meat
- Eme-Iron ?

Normal Mucosa

Polyps

- Small
- Medium
- Large

Cancer

Metabolic factors

- Physical Activity
- Early life growth
- Obesity
- Insulin resistance
- Chronic inflammation

Alcohol
Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. Biobank Qatar
9. Conclusions
EPIC: Relative risk estimates for …CVD? Diabetes?

Cholesterol         HDL- C            LDL- C             TRIGLYC.         GLYCEMIA

LOW > HIGH          LOW > HIGH        LOW > HIGH         LOW > HIGH        LOW > HIGH

p= 0.45            p= 0.0003         p= 0.75             p= 0.04            p= 0.005
EPIC: Relative risk estimates for ...

Endometrial Cancer

Adjusted for: menopausal status, age at blood collection, time of day of blood collection, fasting status, and in premenopausal women, phase of menstrual cycle

- Cholesterol
- HDL-C
- LDL-C
- TRIGLYC.
- GLYCEMIA

RR

LOW > HIGH

p = 0.45

p = 0.0003

p = 0.75

p = 0.005

p = 0.04

LOW > HIGH

LOW > HIGH

LOW > HIGH

LOW > HIGH

LOW > HIGH

p = 0.45

p = 0.0003

p = 0.75

p = 0.005

p = 0.04

EPIC: Relative risk estimates for ... Endometrial Cancer

Adjusted for: menopausal status, age at blood collection, time of day of blood collection, fasting status, and in premenopausal women, phase of menstrual cycle.

POSSIBLY THESE BIOMARKERS ARE "CORRELATED METABOLIC CHANGES", UNLIKELY TO BE PART OF THE AETIO- PATHOGENETIC PATHWAY TO CANCER
Endometrial cancer risk by plasma level of

C Peptide

Free estradiol

Adjusted for BMI, OC and HRT use
Endometrial cancer risk by plasma level of

MORE LIKELY TO BE CAUSALLY INVOLVED

C Peptide

Free estradiol

Adjusted for BMI, OC and HRT use
Outline of the presentation

1. Rational for population biobanks. Major lifestyle and anthropometric changes over recent times
2. EPIC rational and design
3. Abdominal obesity and mortality
4. Breast cancer
5. Colorectal cancer
6. Disease X
7. Disease Y
8. BiobankQatar
9. Conclusions
Relative risk of ........................................... according to systolic (A) and diastolic (B) blood pressure, EPIC, 1992–2004

A) Systolic blood pressure (mmHg)

B) Diastolic blood pressure (mmHg)
Relative risk of **Kidney Cancer** according to systolic (A) and diastolic (B) blood pressure, EPIC, 1992–2004
Relative Risk of Kidney cancer associated with Systolic Blood Pressure and BMI

EPIC: 296,638 participants, median follow-up 6 years
UK Biobank

• Prospective cohort with follow up over many years
• Data and samples on 500,000 men and women ages 40-69

<table>
<thead>
<tr>
<th>Vacutainer tube</th>
<th>Fractions</th>
<th>Number of aliquots</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDTA (9ml × 2)</td>
<td>Plasma</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Buffy coat</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Red cells</td>
<td>1</td>
</tr>
<tr>
<td>LH (PST)</td>
<td>Plasma</td>
<td>3</td>
</tr>
<tr>
<td>Clot activator (SST)</td>
<td>Serum</td>
<td>3</td>
</tr>
<tr>
<td>ACD</td>
<td>DMSO blood</td>
<td>–</td>
</tr>
<tr>
<td>EDTA (4ml)</td>
<td>Haematology</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(immediate)</td>
<td>–</td>
</tr>
<tr>
<td>Urine</td>
<td>Urine</td>
<td>4</td>
</tr>
<tr>
<td>Total Aliquots</td>
<td></td>
<td>19</td>
</tr>
</tbody>
</table>

PST, plasma separation tube; SST, serum separation tube; LH, lithium heparin.

Plus RNA and saliva

Qatar Biobank

- Established in 2010 by Qatar Foundation, Hamad Medical Corporation and the Supreme Council for Health in collaboration with Imperial College London.
- Qatar Biobank (QBB) is Qatar’s first long-term and large scale medical research initiative for Qatar’s population.
Knowledge and understanding

New knowledge on the causes and prevention of diseases

Investigation of disease pathways in patient sub-cohorts

Improved public health & health care provision

Personalised healthcare

The Qatar Biobank

Large prospective study with storage of baseline data & biological samples

Long-term follow-up with linkage to future health events

Qatar Foundation + SCH + Imperial College
TARGET POPULATION of Qatar Biobank

From: Qatar Statistics Authority (SA)
TARGET POPULATION of Qatar Biobank

AGE 14-17 to be included at later stage

Total Population by Age Groups and Sex
Qatar 2009

From: Qatar Statistics Authority (SA)
QatarBiobank: Clinic Visit and Data Collection

**Blood Pressure, ECG, Arterial Stiffness**

**Spirometry**

**Imaging**

**Urine & saliva**

**Diet**

**Building 29**

**Blood samples**

**Questionnaire**

**Hand Grip, Anthropometry**

**Cognitive test**
Pre-pilot participants (N=100)

45 Qatari nationals and 55 long term residents

**Distribution by Age and Gender:**
- **18-25 years:**
  - Males: 30
  - Females: 25
- **25-30 years:**
  - Males: 15
  - Females: 15
- **30-40 years:**
  - Males: 10
  - Females: 10
- **40-50 years:**
  - Males: 5
  - Females: 5
- **>50 years:**
  - Males: 10
  - Females: 5

Total:
- **Males:** 80
- **Females:** 50

**Note:**
- The graph shows the number of participants in different age groups, with a breakdown by gender. The total number of participants is 100, with 45 Qatari nationals and 55 long term residents.
Self-reported health conditions

N for diabetes (28 M, 67 F), high cholesterol (28 M, 68 F), high blood pressure (27 M, 68 F), asthma (27 M, 68 F), weight gain (28 M, 68 F), and chest wheeze (27 M, 63 F) calculations.
Total n does not include those who selected ‘do not know’, ‘Prefer not to answer’ or were missing data.
Leisure-time physical activity levels by gender
(Percentages calculated on 71 women and 29 men)
Leisure-time physical activity levels by gender (Percentages calculated on 71 women and 29 men)
Obesity and diabetes

BMI

Waist Circumference

<table>
<thead>
<tr>
<th>BMI</th>
<th>Diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Waist Circumference</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
</tr>
</tbody>
</table>
Systolic blood pressure in relation with abdominal obesity and sleep habits

<table>
<thead>
<tr>
<th>Usual Hours of Sleep</th>
<th>Systolic Blood Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;5</td>
<td>Large Waist Circumference</td>
</tr>
<tr>
<td>6 - &lt;7</td>
<td>Normal Waist Circumference</td>
</tr>
<tr>
<td>7 - &lt;8</td>
<td>Large Waist Circumference</td>
</tr>
<tr>
<td>&gt; 8</td>
<td>Normal Waist Circumference</td>
</tr>
</tbody>
</table>
Physical activity and blood pressure

BMI

Sex
Female
Male

Physical activity and abdominal obesity

Waist Circumference

None 1-7 > 7
Hours of Leisure Physical Activity

None 1-7 > 7
Hours of Leisure Physical Activity
Lung function (Forced Expiratory Volume per sec), In relation with smoking and asthma
Snoring, Obesity and lung function

- **Waist Circumference**
  - No
  - Yes
  - Your spouse/relative mentioned your snoring

- **FEV 1**
  - No
  - Yes
  - Your spouse/relative mentioned your snoring
Advantages of Prospective Cohort Studies

- Risk factors and biomarkers can be measured before the disease develops (helping to avoid “reverse causality”)
- Associations can be assessed with a range of common diseases
- Appropriate controls can be selected from within the same population as the disease cases
- Favorable for systems biology approach
### Examples of Large Prospective Cohort Studies with Biobanks

<table>
<thead>
<tr>
<th>Study</th>
<th>N of Participants:</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPIC (Europe)</td>
<td>520,000</td>
</tr>
<tr>
<td>UK Biobank</td>
<td>500,000</td>
</tr>
<tr>
<td>Harvard Cohorts</td>
<td>300,000</td>
</tr>
<tr>
<td>NIH-NCI PLCO</td>
<td>90,000</td>
</tr>
<tr>
<td>Qatar Biobank</td>
<td>up to 60,000</td>
</tr>
</tbody>
</table>

Studies in different populations allow investigation of a wide range of exposures and diseases.
Global challenges in public health