COHORT PROFILE

Cohort Profile: The Boston Area Community Health (BACH) survey

Rebecca S Piccolo,1* Andre B Araujo,1 Neil Pearce2 and John B McKinlay1

1Institute for Health Services and Disparities Research, New England Research Institutes, Watertown, MA, USA and 2Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK

*Corresponding author. New England Research Institutes Inc., 9 Galen Street, Watertown, MA 02472, USA. E-mail: rpiccolo@neriscience.com

Accepted 11 October 2012

The Boston Area Community Health (BACH) Survey is a community-based, random sample, epidemiologic cohort of \( n = 5502 \) Boston (MA) residents. The baseline BACH Survey (2002–05) was designed to explore the mechanisms conferring increased health risks on minority populations with a particular focus on urologic signs/symptoms and type 2 diabetes. To this end, the cohort was designed to include adequate numbers of US racial/ethnic minorities (Black, Hispanic, White), both men and women, across a broad age of distribution. Follow-up surveys were conducted \( \sim 5 \) (BACH II, 2008) and 7 (BACH III, 2010) years later, which allows for both within- and between-person comparisons over time. The BACH Survey’s measures were designed to cover the following seven broad categories: socio-demographics, health care access/utilization, lifestyles, psychosocial factors, health status, physical measures and biochemical parameters. The breadth of measures has allowed BACH researchers to identify disparities and quantify contributions to social disparities in a number of health conditions including urologic conditions (e.g. nocturia, lower urinary tract symptoms, prostatitis), type 2 diabetes, obesity, bone mineral content and density, and physical function. BACH I data are available through the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Central Repositories (www.niddkrepository.org). Further inquiries can be made through the New England Research Institutes Inc. website (www.neriscience.com/epidemiology).

Why was the cohort set up?

Despite steady improvement in the overall longevity of the US population,1 racial and ethnic minorities, with few exceptions, experience higher rates of morbidity and mortality than non-minorities.2,3 The reasons for these health disparities are multifactorial and poorly understood, but are hypothesized to reflect differences in socio-economic status (SES), lifestyle and behavioural risk factors, environmental effects, genetic influences and access to health care. Given these competing and interrelated potential explanations for health disparities, there was a compelling need for research that simultaneously examined and measured these multiple potential explanations using a multidisciplinary approach.

The Boston Area Community Health (BACH) Survey was designed to explore these relative contributions conferring increased health risks on minority populations.2 In addition to the primary research interests
in the effects of age, sex and race/ethnicity, the BACH Survey was also concerned with lack of adequate health insurance, lack of access to adequate medical care and how these problems influence patterns of disease. The baseline BACH Survey was initiated in 2002 in response to a National Institutes of Health (NIH) consensus panel recommendation that research on urologic and gynaecologic conditions in racial/ethnic minorities be prioritized. At that time, epidemiologic studies in the field of urology were limited by three key factors that the BACH Survey was designed to address: (i) lack of representation of racial/ethnic minorities; (ii) cohorts of patients who access medical care and receive a diagnosis (i.e. non-population-based studies); and (iii) reliance on variably defined and diagnosed medical conditions. Before the BACH Survey, little was known about the basic descriptive epidemiology (i.e. prevalence, incidence) of urologic symptoms in the general population or about how they vary by major social determinants, such as race/ethnicity. The goal of the baseline BACH Survey (BACH I: 2002–05) was to measure the prevalence of urologic symptoms/conditions by race/ethnicity, age, sex and SES. To this end, the BACH Survey used a random community-based sample of racially/ethnically diverse men and women across a broad age range (30–79 years) from the Boston (MA), population. From the outset, this initial survey was intended to provide the baseline data for a longitudinal study, and in 2008, enrolment began for the first follow-up survey (BACH II: 2008–10). Enrolment in a third wave (BACH III: 2010–12) has been recently completed. All three waves of the BACH Survey were funded by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

Who is in the cohort?
The BACH Survey was designed to include adequate numbers of US racial/ethnic minorities (Black, Hispanic and White participants) and sufficient numbers of both men and women, and to balance across a broad age distribution (30–79 years, by 10-year age groups). These requirements were intended to permit examination of rare conditions across major population sub-groups of interest. The final baseline sample, by design cell, is provided in Table 1. A total of 5502 participants were recruited with similar numbers across the three racial/ethnic groups considered (1767 Black, 1876 Hispanic and 1859 White).

How was this sample attained?
The sampling strategy for the BACH Survey has been published previously. Briefly, to ensure a representative sample, a stratified, two-stage cluster sampling design was used (Figure 1), with census blocks as the primary sampling units and households as the secondary sampling units. Census blocks were stratified by minority density, and high minority strata were over sampled to attain a sample with roughly 1/3 Black, 1/3 Hispanic and 1/3 White participants. The individual response rate, which was calculated as the number of participants interviewed divided by the number of participants for whom contact was attempted, was 57.3%.

How often have they been followed up?
Two follow-up surveys to BACH have been completed. BACH II was initiated in 2008 with \( n = 4145 \) participants. BACH III was initiated in 2010 with \( n = 3150 \) participants.

BACH II (2008–10)
Approximately 5 years after the initial BACH Survey, 4145 participants completed the BACH II survey representing an 80.5% retention rate (Table 2). The average length of time between the baseline and follow-up interviews was 4.8 years. Attrition between BACH I and BACH II was highest among racial/ethnic minorities and men. Retention rates for men were 77.3, 68.1 and 82.6% among Black, Hispanic and White men, respectively. Retention rates for women were 84.1, 79.1 and 88.1% among Black, Hispanic and White women, respectively. Retention rates were higher with increasing age, with the exception of the oldest age group (70–79 years at baseline). Lower retention rates were observed among lower SES participants.

BACH III (2010–12)
Participants were approached in 2010 to participate in BACH III (2010–12) achieving an 81.4% retention rate (of those completing BACH II). Overall, 65.2% of eligible BACH I participants were retained through BACH III. Eleven participants participated in BACH III, but not BACH II. The average length of time between BACH II and BACH III was 2.5 years. Retention rates were lowest among men from BACH II and BACH III. Retention rates (conditional on BACH II participation) were 81.0, 77.8 and 81.6% among Black, Hispanic and White men, respectively, and 83.5, 83.5 and 82.4% among Black, Hispanic and White women, respectively. Retention rates increased slightly with older age. Retention was not significantly related to SES.

Sub-studies
In addition to the three waves of the BACH Survey, a number of sub-studies have used in the BACH cohort.

The BACH/Bone Survey is an observational research study of musculoskeletal health in 1219 men recruited from the parent study, BACH. The baseline
A follow-up survey (BACH/Bone II) is currently recruiting men from the original cohort to examine longitudinal changes in fall risk and bone density.

Endothelial Function and Erectile Dysfunction (ED/EnD) is an observational research study conducted among 400 men participating in the BACH/Bone Survey. This study investigates the association between endothelial function and ED and is designed to establish the role of endothelial dysfunction in the aetiology and natural history of ED. The study was initiated in January 2010.

Beneath the Urologic Iceberg is a qualitative study linked to the BACH Survey, consisting of focus groups and in-depth interviews. A primary objective was to explore factors underlying the care-seeking process for urinary symptoms. Participants were randomly sampled from each of the six sub-groups of the BACH sample and included individuals who reported one or more lower urinary tract symptoms (LUTS) on the survey. Fifty-eight participants participated in a total of eight focus groups. In-depth interviews were conducted with 151 participants.

### Table 1 BACH study design (age, sex and racial/ethnic composition of the BACH sample)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>614 661 509 329 188 2301</td>
<td>403 480 381 245 101 1610</td>
<td>265 350 306 188 75 1184</td>
</tr>
<tr>
<td>Black</td>
<td>164 224 156 103 53 700</td>
<td>105 168 120 71 22 486</td>
<td>76 129 98 51 16 370</td>
</tr>
<tr>
<td>Hispanic</td>
<td>249 229 156 92 40 766</td>
<td>150 147 105 67 22 491</td>
<td>96 107 85 51 15 354</td>
</tr>
<tr>
<td>White</td>
<td>201 208 197 134 95 835</td>
<td>148 165 156 107 57 633</td>
<td>93 114 123 86 44 460</td>
</tr>
<tr>
<td>Women</td>
<td>793 835 776 517 280 3201</td>
<td>610 660 643 434 188 2535</td>
<td>460 529 514 332 132 1967</td>
</tr>
<tr>
<td>Black</td>
<td>259 284 249 179 96 1067</td>
<td>196 229 207 143 66 841</td>
<td>156 187 163 108 43 657</td>
</tr>
<tr>
<td>Hispanic</td>
<td>337 319 256 138 60 1107</td>
<td>249 240 212 112 37 850</td>
<td>192 193 180 85 29 679</td>
</tr>
<tr>
<td>White</td>
<td>197 232 271 200 124 1024</td>
<td>165 191 224 179 85 844</td>
<td>112 149 171 139 60 631</td>
</tr>
<tr>
<td>Total</td>
<td>1407 1496 1285 846 468 5502</td>
<td>1013 1140 1024 679 289 4145</td>
<td>724 879 820 520 207 3151</td>
</tr>
</tbody>
</table>

examination occurred between 2002 and 2005.
The Intra-Subject Hormone Variation Study was designed to measure intraindividual variation in hormones among men. Male participants \( n = 134 \) were randomly selected from the BACH Survey’s study strata. Two blood samples (drawn 20 min apart) were obtained from two study visits (1–3 days apart) at study entry and again 3 and 6 months later.

**What has been measured?**

The main outcomes of interest in the first two waves were urologic symptoms and conditions. Extending beyond the initial outcomes of interest, the third wave of the BACH Survey focused on type 2 diabetes, pre-diabetes and metabolic syndrome risk assessment. The characterization and explanation of social disparities (by age, racial/ethnicity and sex) in the prevalence of disease has been the central focus of the BACH Survey through all three waves of the study. All three waves measured a number of other factors thought to contribute to the aetiology of disease or to mediate the relationship between social disparities (according to racial/ethnicity, SES, age, sex) and health outcomes (Figure 2). These variables can be categorized into seven groups: (i) socio-demographic characteristics, (ii) health care access/utilization, (iii) lifestyles, (iv) psychosocial factors, (v) health status, (vi) anthropomorphic measurements and (vii) biochemical parameters. Table 3 gives details of the types of information collected in each wave of the BACH Survey.

When possible, previously validated questionnaires were used in the BACH Survey. Specifically, measures that were previously published in a peer-reviewed journal, had reported metric properties, were available in English and Spanish and were already used in field epidemiology settings were preferred. A National Institutes of Health (NIH) scientific advisory committee offered recommendations on validated scales. The BACH questionnaires and project correspondence were translated into Spanish and then back-translated to ensure cross-cultural equivalence of meaning; 26% of the BACH interviews were conducted in Spanish (76% of interviews among Hispanics were conducted in Spanish). All protocols, questionnaires and forms used in the BACH Survey were annually reviewed and approved by the New England Research Institutes’ Institutional Review Board.

An interviewer-administered questionnaire and anthropomorphic measures were included as a part of the BACH Survey at all three time points. For BACH I and BACH II, sensitive questions, such as sexual functioning and abuse history, were ascertained through a self-administered questionnaire. Blood samples were taken at the first and third study waves, with serum aliquots stored at \(-80°C\) for future use. DNA samples were isolated from the BACH III blood samples, and ancestry informative markers were collected. The ancestry informative markers are a panel of markers informative for geographic ancestry that can identify a participant’s proportion of European, West African or Native American ancestry. DNA samples are stored for future use.

**What has it found? Key findings and publications**

The BACH Survey’s design and the breadth of measurements have allowed researchers to identify disparities and quantify contributions to social disparities
in a number of health conditions; these have included racial/ethnic disparities in obesity, exposure to prescription medications, variation in markers of bone turnover and bone mineral content and density, higher rates of vitamin D deficiency and physical function. The study has also explored potential explanations for these racial/ethnic disparities with a particular focus on socio-economic status. Recent publications from BACH demonstrate that socio-economic status accounts for much of the racial/ethnic disparities seen in the rates of erectile dysfunction, nocturia and diabetes. These findings are of critical importance for informing prevention and treatment strategies.

BACH has also contributed to the literature on gender disparities. BACH findings have suggested sex-specific effects in several health conditions, explored previously unresearched areas of women’s sexual health and have contributed significantly to the literature on the effects of abuse.

The BACH Survey’s novel ‘upstream’ focus has led to new estimates on the magnitude of unmet need for drug treatment of urologic symptoms and has identified populations with unmet health insurance needs by studying both the uninsured and the underinsured.

Basic epidemiologic data on health disparities (e.g. racial/ethnic specific prevalence and incidence) had not been previously estimated for many urologic conditions, and disparities in these conditions were poorly understood. The BACH Survey provided prevalence rates by racial/ethnic for urine leakage, LUTS, painful bladder syndrome, nocturia and diabetes. The BACH Survey contributed prevalence estimates and identified risk factors for female sexual dysfunction, erectile dysfunction and symptomatic androgen deficiency.

Before the BACH Survey, urologic symptoms were not considered important clinical or public health problems. The BACH Survey helped identify an epidemic of urologic conditions and estimated that 52 million adults in the USA will have symptoms of LUTS, urine leakage, painful bladder syndrome or prostatitis in 2025. The BACH Survey demonstrated that urologic symptoms were significantly associated with other major medical conditions (type 2 diabetes, cardiac disease, hypertension and depression), and a dose-response relationship between the severity and severity of urologic symptoms and their associated co-morbidities.

---

**Table 2** Retention and attrition of participants in the BACH study cohorts

<table>
<thead>
<tr>
<th>Response categories</th>
<th>BACH I</th>
<th>BACH II</th>
<th>Pre-diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondents</td>
<td>5502</td>
<td>4145</td>
<td>3151</td>
</tr>
<tr>
<td>Non-respondents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineligible</td>
<td>348</td>
<td>324</td>
<td></td>
</tr>
<tr>
<td>Refusal</td>
<td>350</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>Unable to contact</td>
<td>657</td>
<td>535</td>
<td></td>
</tr>
<tr>
<td>Total eligible</td>
<td>5152</td>
<td>3856</td>
<td></td>
</tr>
<tr>
<td>Retention as % eligible</td>
<td>80.5%</td>
<td>81.7%</td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 2** Research model for the BACH study

---
duration of urologic symptoms and chronic illnesses was identified. Urologic symptoms were also shown to have a negative impact on quality of life, with an effect on quality of life similar to that of having diabetes, high blood pressure or cancer.

Given the newfound importance of urologic symptoms and conditions, the BACH Survey’s estimates on the risk factors for these conditions, the overlap between these conditions and the unmet medical care needs for these conditions are important contributions to the field of urology.

Table 3

Measures available from the BACH cohort study (2002–12)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Sociodemographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential address (geo-coded), mobility</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Income, education, work status, occupation, marital status</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Sociological questionnaire, including acculturation, alienation, neighbourhood order/disorder, perceived discrimination, health literacy</td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>II. Health care access/utilization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health care access/utilization</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Health insurance status/type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of care, satisfaction with care</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Inclination to seek care</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III. Lifestyles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical activity, diet</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Abuse history</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Tobacco and alcohol use</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Sleep</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>IV. Psychosocial factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depressive symptoms, interpersonal stress, major life events</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Depressive symptoms, interpersonal stress</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Major life events</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>V. Health status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of life (self-rated health, current and projected life satisfaction)</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Chronic disease/events, family medical history, pain, fatigue, menopausal status</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Inventoried prescription/non-prescription medications and supplements</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>VI. Physical/anthropomorphic measures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height, weight, body fat percentage, hip/waist circumference, blood pressure, pulse</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>VII. Biochemical parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol, HDL, LDL, triglycerides</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Testosterone, estradiol, SHBG, FSH, LH (men only)</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Cortisol, C-reactive protein</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Fasting blood glucose, HbA1c, insulin</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Serum aliquots stored at −80 F</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Stored DNA, ancestry informative markers</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
</tbody>
</table>

HDL, high-density lipoprotein; LDL, low-density lipoprotein; SHBG, sex hormone-binding globulin; FSH, follicle-stimulating hormone; LH, luteinizing hormone.
What are the main strengths and weaknesses?

Strengths

The strengths of the BACH Survey stem from its community-based random sample design. The study, by design, includes both sexes, a wide age range (30–79 years) and includes a large number of minority participants, representative of Black and Hispanic populations. Key strengths of the BACH Survey include: (i) the wide range of measurements covering six theoretical domains (Figure 2); (ii) these measurements allow for both individual-level and neighbourhood-level (multi-level) analyses; (iii) its longitudinal design that allows for within- and between-person comparisons over a 10-year period; (iv) its focus on pre-diagnostic disparities (e.g. urologic symptoms, pre-diabetes) rather than disparities based on variably diagnosed conditions; and (v) the multi-disciplinary approach measures, the prevalence of disease through both self-report and physiologic (objective) confirmation. In summary, the BACH participants are well-phenotyped in a number of key areas (variety of measures, over time, undiagnosed and diagnosed conditions) that could lead to productive collaborations in many areas where data pooling is needed.

Representativeness and generalizability

Although geographically limited to the city of Boston, MA, the BACH Survey sample has been compared with other large regional (Behavioral Risk Factor Surveillance System, Centers for Disease Control and Prevention) and national (the National Health and Nutrition Examination Survey, the National Health Interview Survey) on a number of different socio-demographic and health-related variables. The results suggest that the BACH Survey is highly representative of the city of Boston and that BACH Survey estimates of key health conditions are comparable with national trends. One key difference is that the BACH Survey does not include a number of other minority groups (e.g. Asian Americans).

Weaknesses

First, several key variables in the BACH Survey (e.g. history of certain medical conditions) rely on self-report data. Relying on self-report data is common among observational studies, and research has shown that self-report of major medical conditions are well correlated with medical record review.62–65 In addition, every attempt was made to directly measure key variables (e.g. height, weight, blood pressure, cholesterol, fasting glucose).

A second limitation of the BACH Survey was the exclusion of Asians from the study. Although a sizeable minority population in the USA, Asians comprised only 7.5% of the Boston population in 2000.66 The feasibility of recruiting and interviewing Asians for inclusion in the BACH Survey (e.g. interviewer language requirements) was weighed against the potential public health impact.

Finally, the initial survey response rate was 57.3%. This response rate, although low, is comparable with response rates among other random sample cohort studies and was not entirely unexpected given the lengthy in-home interview (2 h), the blood draw and the sensitive nature of many of the questions. Nonetheless, the study has maintained high retention rates, thus presumably mitigating concerns regarding internal validity. The BACH Survey staff has fostered a close relationship with the study participants and with the inner-city Boston community. This close contact and continued communication through newsletters, holiday cards and birthday cards helps to ensure a trust between the study participants and the study research team and staff that leads to high retention rates and good response rates to sensitive questions.

Can I get hold of the data? Where can I find out more?

BACH I data are available through the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Central Repositories (www.niddk-repository.org). Further inquiries can be made through the New England Research Institutes Inc. website (www.neriscience.com/epidemiology).

Funding

National Institute of Diabetes and Digestive and Kidney Disorders (DK056842 and DK080786).

Conflict of interest: None declared.

KEY MESSAGES

- The Boston Area Community Health (BACH) Survey is a community-based random sample epidemiologic cohort of 5502 men and women, aged 30–79 years from the population of Boston, MA.

- The diverse cohort (1767 Black, 1876 Hispanic and 1859 White participants) has allowed for estimates on the magnitude of health disparities in urologic conditions (e.g. urine leakage, painful bladder syndrome, nocturia, prostatitis).
- The BACH Survey has identified an epidemic of urologic conditions, helped to demonstrate their public health importance, identified risk factors for these conditions and demonstrated significant unmet treatment needs.
- Moving beyond urologic conditions, the BACH Survey has quantified social disparities in type 2 diabetes, obesity, exposure to medications, bone mineral content and density, nutrition and physical function.

References


31. Link CL, McKinlay JB. Only half the problem is being addressed: underinsurance is as big a problem as uninsured. *Int J Health Serv* 2010;40:507–23.


