



HHS Public Access

Author manuscript

Cancer Causes Control. Author manuscript; available in PMC 2022 March 19.

Published in final edited form as:

Cancer Causes Control. 2011 March ; 22(3): 463–468. doi:10.1007/s10552-010-9717-0.

Cohorts and Consortia Conference: a summary report (Banff, Canada, June 17–19, 2009)

Paolo Boffetta^{1,2}, Graham A Colditz³, John D Potter⁴, Laurence Kolonel⁵, Paula J Robson⁶, Reza Malekzadeh⁷, Daniela Seminara⁸, Ellen L Goode⁹, Keun-Young Yoo¹⁰, Paul Demers¹¹, Richard Gallagher¹², Ross Prentice⁴, Yutaka Yasui¹³, Kieran O'Doherty¹⁴, Gloria M Petersen⁹, Cornelia M Ulrich^{4,15}, Ilona Csizmadj¹⁶, Ernest K Amankwah¹⁶, Nigel T Brockton¹⁶, Christine M Friedenreich¹⁶, Karen Kopciuk¹⁶, S Elizabeth McGregor¹⁶, Linda E Kelemen¹⁶

¹The Tisch Cancer Institute, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY, USA, 10029

²International Prevention Research Institute, 95 cours Lafayette, Lyon, France, 69006

³Institute For Public Health, Washington University In St Louis, St. Louis, MO 63110, USA

⁴Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle WA, PO Box 19024, Seattle WA 98109-1024, USA

⁵Cancer Research Center, University of Hawaii, 1236 Lauhala Street, Honolulu, HI 96813, USA

⁶Alberta Health Services - Cancer Care, Department of Population Health Research Suite 1400, Sun Life Place, 10123 99th Street, Edmonton, AB, Canada, T5N 3T2

⁷Digestive Disease Research Center, Tehran University of Medical Sciences, Shariati Hospital 14117 Tehran, Iran

⁸National Cancer Institute, Bethesda, MD, USA; 20892-7335

⁹Department of Health Sciences Research, College of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

¹⁰Seoul National University College of Medicine, Department of Preventive Medicine, 103 Daehangno, Chongno-gu, Seoul 110-799, Korea

¹¹School of Environmental Health, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2

¹²Cancer Control Research Program, BC Cancer Agency, Vancouver, BC, Canada

¹³University of Alberta, School of Public Health, 3-27 University Terrace, Edmonton, AB, Canada, T5R 3B9

¹⁴University of British Columbia, W. Maurice Young Centre for Applied Ethics, 227-6356 Agricultural Road, Vancouver, BC, Canada, V6T 1Z2

Correspondence: Paolo Boffetta, MD, MPH (paolo.boffetta@i-pri.org) or Linda E Kelemen, MSc, ScD (lkelemen@post.harvard.edu).

¹⁵German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany, 69120

¹⁶Alberta Health Services-Cancer Care, Department of Population Health Research, 1331 29th Street NW, Calgary, AB, Canada, T2N 4N2

Abstract

Epidemiologic studies have adapted to the genomics era by forming large international consortia to overcome issues of large data volume and small sample size. Whereas both cohort and well-conducted case-control studies can inform disease risk from genetic susceptibility, cohort studies offer the additional advantages of assessing lifestyle and environmental exposure-disease time sequences often over a life-course. Consortium involvement poses several logistical and ethical issues to investigators, some of which are unique to cohort studies, including the challenge to harmonize prospectively-collected lifestyle and environmental exposures validly across individual studies. An open forum to discuss the opportunities and challenges of large-scale cohorts and their consortia was held in June 2009 in Banff, Canada and is summarized in this report.

Keywords

Biobanks; cohort studies; consortia; cancer; ethics; data harmonization; molecular epidemiology

Introduction

In past decades, epidemiologic study designs have embraced high-throughput technologic advances that generate increasingly large volumes of molecular and genetic data. Information from these investigations has revolutionized our ability to understand cancer pathways and disease etiology. In turn, these technologies have fostered “big science” epidemiology [1]. For example, large consortia of case-control studies arose to overcome many of the limitations associated with large data volume and small sample size, and their reports of associations with common, low-penetrant alleles from genome-wide association studies have provided clues to the genetic etiology of various cancers [2–3]. A similar movement is occurring with large-scale cohort studies and their consortia [4–5], which offer the advantages of assessing lifestyle and environmental exposure-disease time sequences, as recently demonstrated by the NCI Cohort Consortium [6]. There is continued emphasis by the epidemiology community to elucidate the contribution to cancer risk of the environment and of gene-environment interactions. Many of the cohort studies that currently participate, or wish to participate, in consortia face the challenge of harmonizing prospectively-collected lifestyle and environmental exposures validly across individual studies, sustained cohort funding, distributed expertise for multiple phenotype assessment, and inadequate funding for consortium efforts. An open forum to discuss the opportunities and challenges of large-scale cohorts and their consortia was organized by the Alberta Health Services-Alberta Cancer Board, and co-sponsored by the Canadian Partnership Against Cancer and the International Agency for Research on Cancer, and held in June 2009 in Banff, Canada. A summary of the highlights from speaker presentations and discussions are presented in this report.

Established and New Cohort Studies

The conference began with presentations by speakers representing large successful cohort studies including the American Cancer Society Cohorts, European Perspective into Cancer and Nutrition, the Multi-Ethnic Cohort, the Nurses' Health Study and the Shanghai Women's Health Study. Speakers agreed that existing cohorts have the capacity to examine many emerging scientific questions. Indeed, many are incorporating molecular technologies that complement existing resources, including adding prospectively collected biologic samples for '-omics' applications, and evaluating DNA, RNA and proteins from formalin-fixed tissue blocks collected from cancer patients. Many agreed that priorities for future research should include the assessment of childhood and adolescent exposures, repeated exposure measurement, and the addition to cancer-focused cohorts of non-cancer endpoints to further address the public-health issues of risk and benefits of exposures [7].

Other speakers highlighted the existence of several newly established cohort studies, including those in low- and medium-resource countries, which can provide new information on cancer etiology and that are actively harmonizing their protocols for future integration into existing biobanks and consortia (Table 1). These investigations are sampling populations with a wide variation in cancer incidence and mortality within a country [8–12], as well as populations undergoing an epidemiologic transition or populations at very high incidence of specific cancers [9–12]. Some of these studies [8] are also including improved measurement of certain environmental exposures, such as dietary assessment and enhanced lifetime residential and work histories that utilize exposure matrices and linkages with industry databases to examine effects of environmental and occupational agents. Finally, novel applications of the cohort design were presented including a model of precision medicine [13] initiated at the Moffitt Cancer Center in Tampa, Florida, whereby information collected from a cohort of patients seen at the clinic is used to generate evidence to select or develop therapies for their disease.

There were several models of successful cohort studies that have been designed to evaluate defined populations that cannot be studied with sufficient power in the general population cohorts. These specialized cohorts and their networks include unique populations, rare events, or specialized study designs (Table 2). Readers interested in obtaining additional information on these studies can contact the authors.

Challenges Faced by Epidemiologists leading Cohort Studies

Conference attendees agreed that every cohort study encounters major challenges including those associated with participant recruitment and retention, which can threaten internal validity, the follow-up time to accrue cases, inability to study very rare diseases, and the acquisition of detailed clinical and pathologic data from multiple hospitals/health care systems that may be laborious, expensive, and difficult to standardize. In addition to these factors, two main challenges encountered by investigators of cohort studies were vocalized throughout the conference: issues related to dietary and physical activity assessment and consortium participation.

Diet and Physical Activity Assessment

The measurement of environmental exposures, particularly diet and physical activity, while not identified as a separate session topic, emerged as an issue that clearly challenged the expertise of the panel and attendees. Reoccurring themes within this discussion included optimizing the quality of dietary data through improved assessment methods and analytic techniques, the role of biomarkers in the calibration of dietary and physical activity exposure estimates, and the role of technology in overcoming the limitations of current methods used in measuring diet and physical activity. Proposed strategies that have the potential to be more widely applied to better ascertain diet and activity exposures included combining diet and activity information from various data collection instruments and the prospective integration of calibration and validation sub-studies into the design of large epidemiological studies to address the fundamental measurement challenges that are at the core of ascertaining these exposures in large observational studies. The conference highlighted the need for future discussions on these themes. As a result, improving current methods of dietary assessment will be the focus of a pre meeting workshop of the 2011 North American Congress of Epidemiology conference (<http://www.epicongress2011.org/>).

Consortium Participation

Issues related to participation in consortium projects dominated the second half of the conference. Important concerns were raised, such as the need to better harmonize questionnaire-based data across various existing studies to avoid substantial heterogeneity in exposure definitions. In recognition of this challenge, the Public Population Project in Genomics (P3G) [14] developed the DataSHaPER (DataSchema and Harmonization Platform for Epidemiological Research) tool to facilitate harmonization of variables across studies (both retrospective and prospective). The P3G is an open-source central portal that provides an available archive of all member cohort questionnaires, which lends itself to careful review and selection of exposure variables of similar quality across studies for investigators wishing to address specific hypotheses. The portal is also a resource for governance models, DNA processing catalogues, and information technology tools that can assist investigators in the daily management of newly-developed and ongoing studies.

The validity of combining biomarker data across cohort studies was another potential concern, and could be improved with prospectively-implemented pilot and recalibration studies to evaluate heterogeneity between studies from the effect of different collection and storage protocols. To combine data for a specific analyte from previously-assayed specimens, researchers may consider setting aside a subset of samples from each study to allow for future calibration with other studies. Other concerns regarding consortium participation included the potential to deplete valuable banked biospecimens rapidly and the pressures to pool data before separate analyses can be completed and published. Although it was agreed that consortia should balance the complexity of the projects they initiate with well-defined objectives and specific deadlines, it was acknowledged that, in practice this does not always occur. To achieve this balance, attendees agreed that consortium research and resources should be prioritized to address questions that are uniquely addressable through pooled or meta-analyses (e.g., rare diseases or exposures or validating small-sample study results). Recognition of the contribution of early career scientists (e.g., through

inclusion among the lead authors of important papers) is a critical issue to ensure the success of their careers, and several consortia have developed authorship policies that promote a leading role for early career scientists in research efforts, grant preparation and publications.

Finally, financial considerations to consortium participation were also noted. Although administrative supplements or funding mechanisms exist for some core functions (e.g., NIH R24 or U24 - Research Infrastructure and Capacity Building or Resource-Related Research Project applications), inadequate funds exist for common database management, member institution costs for individual consortium projects (e.g., compilation of analytic datasets with requested variables, biospecimen maintenance and retrieval costs), and, in some instances, the large commitment of investigator time.

Balancing Data Sharing and Privacy of Genomic and Epidemiologic Data

The importance of harmonizing biospecimen collection and storage methods across studies to facilitate consortium efforts was countered with ethical issues of upholding participant privacy and institutional review board regulations at time of participant consent.

From a logistical and validity perspective, speakers noted that standard operating procedures exist for large biobanking activities [15] but many protocols are based solely on feasibility. To better harmonize these initiatives using an evidenced-based approach, the Forum for International Biobanking Organizations (FIBO) [16] developed criteria for reporting on the quality of samples and related information in biobanks. Speakers also identified a need for an international protocol on biospecimen sharing for public-health research that would have sets of regulations that are permissive enough to facilitate sample exchange. The pan-European Biobanking and Biomolecular Resources Infrastructure (BBMRI) was presented as a good example of a recent initiative to combine different types of biobanks (e.g., cohort, clinical samples) to overcome many obstacles for sharing of biobank collections in international collaborative studies [15]. These issues also support the need for a cadre of professionals with training and expertise in biospecimen management.

Speakers also noted that there are ethical and legal challenges associated with making data available publicly from large-scale human genetics research [17–18]. For example, research participants face heightened risks of privacy infringement owing to long term storage of data and samples and increased access to those data. Sharing of samples and data across national boundaries means that security measures implemented to protect participants in the source jurisdiction may not be enforceable in other jurisdictions. As DNA is by its very nature a unique identifier of individuals, cohorts involved in genomic studies face the challenge that even aggregation of donors' data is no longer sufficient to protect against the identification of an individual's participation in a given study [19–21]. Traditional informed consent protocols cannot be applied in many cases as re-consenting participants for different studies may not be possible or financially prohibitive, and no consensus has yet emerged as to ethically acceptable alternatives. There is often a lack of clarity on what the ethical obligations of researchers are (if any) to act on incidental findings of clinically relevant information. In addition to scientific expertise, these issues involve trade-offs of societal values against each other; for example, increased levels of donor

anonymity are associated with decreased utility for research, and vice versa: increased utility of donors' samples for certain types of research (in particular, studies focusing on the relationship between genotype and environmental factors) is associated with increased risk of privacy infringement [22–23]. For these reasons, it was agreed that resolution of the ethical conundrums must involve some form of public engagement to maintain trust over the lifespan of the studies, which in turn will strengthen the internal validity of the study through increased subject recruitment and retention [24]. Balancing privacy and data-sharing is thus possible, but necessitates public engagement. Although commercial involvement is important in order to translate the public data to improved treatment and management of patients, attendees recognized that a balance is needed with independent scientific scrutiny to avoid self-serving commercial interests.

Conclusions and Future Priorities

Cohort studies remain a powerful tool in the medical and public-health research agenda. Since research costs and feasibility are such that few randomized controlled prevention trials can be conducted, cohort studies may provide the only viable approach for testing multiple-exposure and multiple-cancer hypotheses concurrently. The last decade has seen several exciting and unprecedented developments in cohort studies, including their integration into consortia. Based on the issues discussed at the conference and highlighted in this report, several priority areas for cohorts and consortia were identified to maximize their benefits to society.

For cohort studies, there is a need for:

- Creation and support of internationally uniform review mechanisms by major funding agencies for cancer cohorts, including the development of guidelines for assessing the performance of existing and new cohort studies, to ensure successful cohorts are financially sustainable. Guidelines such as those proposed by Colditz and Winn [25] can be used as a model;
- Support for new cohort studies that emphasize racial/ethnic diversity and unique phenotypes and exposures, and of protocols that utilize standardized, validated measurements and that collect high-quality biospecimens. Integration of new cohort studies into international collaborations such as the P3G or other established entities at the outset (e.g., consortia, biobanks, linkages with other health-outcome registries) increases the value of a cohort by allowing studies of health and disease trajectories over time, and by optimizing use of pre- and post-diagnostic biospecimens;
- Incorporation of prospectively-implemented quality control and calibration sub-studies to evaluate measurement error and heterogeneity across different studies from the effect of different data-collection and biospecimen-storage protocols;
- Continued multi-level co-operation (investigators and public and private sectors) to ensure the shared governance, longevity and financial support of cohort studies particularly for the maintenance of linked resources (e.g., tumor or health registry databases) within the cohort;

- Continued publication by cohorts of detailed data-access policies to facilitate transparency in data sharing;

For consortia, there is a need for:

- Adaptation of existing funding mechanisms to support different aspects of consortium activities including costs associated with data harmonization, meta-data database creation and dissemination of tools and protocols particularly for individual consortium projects; and
- Up-front recognition of the roles, responsibilities and contributions of early career investigators in consortia so as to make these undertakings compatible with the needs for career advancement.

Although some of these priorities require pressing attention, the conference presentations and discussions revealed that many of the mechanisms and infrastructure (e.g., P3G, FIBO, BBMRI) are already in place to navigate successfully most of these priorities. Existing mechanisms thus serve as strong model systems for future comparison and integration that can produce continued discovery to aid medical and public-health action.

Acknowledgements:

The scientific and organizing committees would like to thank the conference speakers for their contributions: Paolo Boffetta, Graham Colditz, John Potter, Elio Riboli, Susan Hankinson, Laurence Kolonel, Michael Thun, Paula Robson, Rahman Jamal, Reza Malekzadeh, Zhengming Chen, Wei Zheng, Prabhat Jha, Daniela Seminara, Ellen Goode, Keun-Young Yoo, Paul Demers, Thomas Sellers, Isabel Fortier, Pierre Hainaut, Richard Gallagher, Jorn Olsen, Les Robison, Ross Prentice, Yutaka Yasui, Thomas Lumley, Kieran O'Doherty, Bartha Maria Knoppers, John McLaughlin, Louise Parker, John McPherson, Eric Paulos and David Duggan.

Financial support: We gratefully acknowledge funding from the Canadian Partnership Against Cancer, the Canadian Institutes of Health Research, the Canadian Breast Cancer Research Alliance, the Canadian Cancer Society, Alberta Cancer Foundation/Alberta Cancer Research Institute, Alberta Heritage Foundation for Medical Research and the University of Calgary.

References

- [1]. Hoover RN (2007) The evolution of epidemiologic research: from cottage industry to "big" science. *Epidemiology* 18: 13–17 [PubMed: 17179754]
- [2]. Song H, Ramus SJ, Tyrer J, et al. (2009) A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nature genetics* 41:996–1000 [PubMed: 19648919]
- [3]. Milne RL, Benitez J, Nevanlinna H, et al. (2009) Risk of estrogen receptor-positive and negative breast cancer and single-nucleotide polymorphism 2q35-rs13387042. *J Natl Cancer Inst* 101: 1012–1018 [PubMed: 19567422]
- [4]. National Cancer Institute. Cohort Consortium (2010) Available from: <http://epi.grants.cancer.gov/Consortia/cohort.html>. Cited 2010
- [5]. The Asian Cohort Consortium (2010) Available from: <http://www.asiacohort.org/Pages/Default.aspx>. Cited 2010.
- [6]. Helzlsouer KJ (2010) Overview of the Cohort Consortium Vitamin D Pooling Project of Rarer Cancers. *Am J Epidemiol* 172: 4–9 [PubMed: 20562193]
- [7]. Colditz GA Ensuring long-term sustainability of existing cohorts remains the highest priority to inform cancer prevention and control. *Cancer Causes Control* 21: 649–656 [PubMed: 20063074]
- [8]. Borugian MJ, Robson P, Fortier I, et al. (2010) The Canadian Partnership for Tomorrow Project: building a pan-Canadian research platform for disease prevention. *CMAJ* 182: 1197–1201 [PubMed: 20421354]

- [9]. UKM Medical Centre. Malaysia Natational Cohort (2010) Available from: <http://intra.hukm.ukm.my/>. Cited 2010
- [10]. Tehran University of Medical Sciences. Golestan Cohort Study of Esophageal Cancer (2010) Available from: <http://ddrc.tums.ac.ir/>. Cited 2010
- [11]. Oxford University. The Kadoorie Biobank Study in China (2010) Available from: <http://www.ctsu.ox.ac.uk/kadooriebiobank>. Cited 2010
- [12]. Jha P, Gajalakshmi V, Gupta PC, et al. (2006) Prospective study of one million deaths in India: rationale, design, and validation results. *PLoS Med* 3: e18 [PubMed: 16354108]
- [13]. Christensen CM, Grossman JH, Hwang J (2008) *The Innovator's Prescription: A Disruptive Solution for Health Care*. McGraw-Hill, New York, NY
- [14]. The P3G Observatory (2010) <http://www.p3gobservatory.org>. Cited 2010
- [15]. Biobanking and Biomolecular Resources Research Infrastructure (2010) Available at: <http://www.bbmri.org>. Cited 2010
- [16]. International Society for Biological and Environmental Repositories (2010) Available at: <http://www.isber.org/FIBO.html>. Cited 2010
- [17]. UNESCO. International Declaration on Human Genetic Data (2010) Available at: http://portal.unesco.org/en/ev.php-URL_ID=17720&URL_DO=DO_TOPIC&URL_SECTION=201.html. Cited 2010
- [18]. Birney E, Hudson TJ, Green ED, et al. (2009) Prepublication data sharing. *Nature* 461: 168–170 [PubMed: 19741685]
- [19]. Bjorn G (2008) Barriers set up to protect genome databases. *Nature medicine* 14:996
- [20]. Homer N, Szelinger S, Redman M, et al. (2008) Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS genetics* 4: e1000167 [PubMed: 18769715]
- [21]. Lowrance WW, Collins FS (2007) Ethics. Identifiability in genomic research. *Science* 317: 600–602 [PubMed: 17673640]
- [22]. Asslauer M, Zatloukal K (2007) Biobanks: transnational, European and global networks. *Briefings in functional genomics & proteomics* 6: 193–201 [PubMed: 17916592]
- [23]. Oosterhuis JW, Coebergh JW, van Veen EB (2003) Tumour banks: well-guarded treasures in the interest of patients. *Nat Rev Cancer* 3: 73–77 [PubMed: 12509769]
- [24]. O'Doherty KC, Burgess MM (2009) Engaging the public on biobanks: outcomes of the BC biobank deliberation. *Public Health Genomics* 12: 203–215 [PubMed: 19367089]
- [25]. Colditz GA, Winn DM (2008) Criteria for the evaluation of large cohort studies: an application to the nurses' health study. *J Natl Cancer Inst* 100: 918–925 [PubMed: 18577745]
- [26]. Pourshams A, Khademi H, Malekshah AF, et al. (2010) Cohort Profile: The Golestan Cohort Study--a prospective study of oesophageal cancer in northern Iran. *Int J Epidemiol* 39: 52–59 [PubMed: 19332502]
- [27]. Jha P, Jacob B, Gajalakshmi V, et al. (2008) A nationally representative case-control study of smoking and death in India. *N Engl J Med* 358: 1137–1147 [PubMed: 18272886]
- [28]. Brown RC, Dwyer T, Kasten C, et al. (2007) Cohort profile: the International Childhood Cancer Cohort Consortium (I4C). *International journal of epidemiology* 36: 724–730 [PubMed: 17255350]
- [29]. Olsen J, Melbye M, Olsen SF, et al. (2001) The Danish National Birth Cohort--its background, structure and aim. *Scand J Public Health* 29: 300–307 [PubMed: 11775787]
- [30]. Ronningen KS, Paltiel L, Meltzer HM, et al. (2006) The biobank of the Norwegian Mother and Child Cohort Study: a resource for the next 100 years. *Eur J Epidemiol* 21: 619–625 [PubMed: 17031521]
- [31]. Mertens AC, Liu Q, Neglia JP, et al. (2008) Cause-specific late mortality among 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. *J Natl Cancer Inst* 100: 1368–1379 [PubMed: 18812549]
- [32]. Robison LL, Armstrong GT, Boice JD, et al. (2009) The Childhood Cancer Survivor Study: a National Cancer Institute-supported resource for outcome and intervention research. *J Clin Oncol* 27: 2308–2318 [PubMed: 19364948]

- [33]. The Women’s Health Initiative Study Group. (1998) Design of the Women’s Health Initiative clinical trial and observational study. *Controlled clinical trials* 19: 61–109 [PubMed: 9492970]
- [34]. Chlebowski RT, Kuller LH, Prentice RL, et al. (2009) Breast cancer after use of estrogen plus progestin in postmenopausal women. *N Engl J Med* 360: 573–587 [PubMed: 19196674]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Description of Selected Newly Established Cohort Studies

Cohort	Population	Reference
Canadian Partnership for Tomorrow Project	A federation of cohorts in five provinces/regions in Canada enrolling 300,000 adults aged 35 to 69 years by 2012 with long-term follow-up. Efforts were made to maximize harmonization with other existing large international biobanks in order to increase potential for future pooling of data and samples.	http://www.partnershipfortomorrow.ca [8]
Malaysia National Cohort	Population-based cohort sampling 100,000 participants by 2012 from urban areas, rural farming communities and the three main ethnic groups in Malaysia.	http://intra.hukm.ukm.my/cohort .
Golestan Cohort Study	First large-scale prospective study of cancer in Middle Eastern countries undergoing economic and social transitions, and focusing on upper gastrointestinal cancers in Northeastern Iran. Enrolment of over 50,000 healthy adults has been completed.	http://ddrc.tums.ac.ir/modules/news/index [26]
Kadoorie Biobank Prospective Study	Prospective study of over 515,000 people aged 35 to 74 years recruited between 2004 and 2008 from 10 diverse regions in China. The first re-survey of ~20,000 participants was completed in 2008 with 85% response. Utilizes linkage with death and disease registries, and future linkage with health insurance claim systems.	http://www.ctsu.ox.ac.uk/kadooriebiobank
Prospective Study of One Million Individuals in India	A large-scale prospective study of chronic diseases recruiting over one million adults aged over 30 years from 5 to 10 regions in India during 2010 and 2011.	[27]

Table 2.

Opportunities in Selected Existing Specialized Cohorts

Cohort	Population	Reference
Pregnancy cohorts	Various population-based cohorts that examine the periods of pre- and peri-conception and puberty on developmental outcomes.	http://www.birthcohorts.net [28] http://www.dnbc.dk [29] http://www.fhi.no/eway/default.aspx?pid=238&trg=MainArea_5811&MainArea_5811=5903;0:15.3046;1:0:0::0:0 [30] http://www.nationalchildremsstudy.aov/Pages/default.aspx
Childhood Cancer Survivor Study	14,372 subjects with childhood cancer diagnosed between 1970 and 1986 plus 20,720 subjects diagnosed between 1987 and 1999	http://ccss.sjude.org [31–32]
Women's Health Initiative	Multifaceted randomized controlled trial cohort and a companion prospective cohort of over 161,000 post-menopausal women.	http://www.whi.org [33–34]