Published in final edited form as:
Cancer Causes Control. 2009 May ; 20(4): 491–496. doi:10.1007/s10552-008-9290-y.

Maximizing resources to study an uncommon cancer: E2C2 -- Epidemiology of Endometrial Cancer Consortium

Sara H. Olson[1], **Chu Chen**[2,3], **Immaculata De Vivo**[4,5], **Jennifer A. Doherty**[3], **Virginia Hartmuller**[6], **Pamela L. Horn-Ross**[7], **James V. Lacey Jr.**[8], **Shannon M. Lynch**[6], **Leah Sansbury**[6], **V. Wendy Setiawan**[9], **Leo J. Schouten**[10], and **Xiao Ou Shu**[11]

[1] Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY [2] Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA [3] Department of Epidemiology, School of Public Health and Community Medicine and Department of Otolaryngology: Head and Neck Surgery, School of Medicine, University of Washington, Seattle, WA [4] Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA [5] Department of Epidemiology, Harvard School of Public Health, Boston, MA [6] Epidemiology & Genetics Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD [7] Northern California Cancer Center, Fremont, CA [8]Hormonal and Reproductive Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD [9] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA [10] Department of Epidemiology, GROW – School for Oncology and Developmental Biology, Maastricht University, The Netherlands [11] Vanderbilt Epidemiology Center, Vanderbilt Institute of Medicine and Public Health, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, TN

Keywords

endometrial cancer; epidemiology

Background and formation of E2C2

Endometrial cancer is the most common gynecologic cancer in the U.S., with more than 40,000 new cases per year(1). Since older age and high body mass index (BMI) are major risk factors, the growing number of older women and high prevalence of obesity are likely to lead to increased numbers of women diagnosed in the coming years. As for many other diseases, little is known about the role of genetic variation or interactions between genes and environmental exposures. Although steroid hormones are clearly important in the etiology of endometrial cancer(2-4), the potential role of genetic variation in genes related to hormone biosynthesis and metabolism has not been defined. Research funding in the U.S. per case and per death for this disease has historically been low relative to breast, ovarian, or cervical cancer; however, several studies with substantial numbers of participants have recently been completed or are nearing completion and cohort studies are identifying larger numbers of affected women as their participants age. Individually, these studies face challenges because of relatively low statistical power, particularly for study of interactions, narrow range of exposures for several risk factors, and homogeneous ethnic background. To overcome these challenges, the

Correspondence: Sara H. Olson, 307 East 63 Street, New York NY 10065, olsons@mskcc.org.

Epidemiology of Endometrial Cancer Consortium (E2C2) was formed in 2006 by an international group of investigators. The overall objective of E2C2 is to build on resources from existing studies funded by the National Cancer Institute (NCI) and other sources by combining data across studies in order to advance our understanding of the etiology of this disease. This commentary presents the progress and challenges of the consortium to date and the current projects addressing etiologic questions.

The consortium was born at a September, 2005, meeting on Understudied Rare Cancers sponsored by the NCI's Epidemiology and Genetics Research Program (EGRP). At that meeting, a group of investigators with special interests in endometrial cancer summarized the state of the science and the gaps that could be addressed by combining resources from funded studies. Subsequently, a Planning Committee, including representatives from EGRP, met regularly by teleconference. We compiled lists of epidemiologic studies of endometrial cancer, both case-control and cohort, that had been published on endometrial cancer risk or that had the resources to do so in the future. Our goal was to be as inclusive as possible. We contacted individual investigators from each of these studies to determine their interest in the new consortium. More than 30 investigators responded positively, agreed to join, and sent detailed information about their studies. E2C2 includes studies conducted outside the US, including studies in Europe, Australia, and China.

E2C2 became a formally designated NCI consortium in June 2006. This official NCI recognition provides sponsorship for some meetings and a web portal. The portal serves as a central hub, with email ready lists of all consortium members; E2C2 policies; descriptions and status of ongoing projects; schedules and minutes for conference calls and meetings; descriptions of and questionnaires used in individual studies; and a list of “core” and other variables for the joint dataset. Four meetings open to all interested investigators have been held, coordinated with meetings of the American Association for Cancer Research.

Organizational structure

The structure of E2C2 is illustrated in figure 1. E2C2 membership is and will continue to be open to all interested investigators, including those with NCI funding and others; at present, most members are epidemiologists but investigators from other disciplines are also welcome. Junior investigators are strongly encouraged to join and have taken an active role in ongoing projects.

The Executive Committee is responsible for developing policies for the consortium, such as those concerning publication and authorship, and is the clearing house for new proposals for use of E2C2 resources. Consortium members wishing to use data from E2C2 are asked to prepare a brief proposal to the Executive Committee, using information available on the web portal about individual studies. The Executive Committee determines whether the proposals address an appropriate scientific question, are consistent with the overall goals of the consortium, and are not in conflict with other ongoing projects. The committee meets regularly by conference call. An Advisory Committee, consisting of senior scientists with extensive experience in endometrial cancer epidemiology and a gynecologist with strong research interest, was formed in 2007. Members provide overall direction and consult on ongoing projects.

Several Working Groups have been formed to address specific areas of interest. When studies are planned, Working Groups identify a data coordinating center and a data analyst, plan analyses and discuss study results, and prepare papers for publication. Obtaining funding is the responsibility of the primary investigator or group. Individual investigators are responsible for ensuring that required agreements on data sharing are in place prior to the transfer of data.

E2C2 resources

Table 1 summarizes the numbers of cases and controls in studies with and without DNA in E2C2. Details of the studies are provided in auxiliary tables 1 through 4. From case-control and nested case-control studies there are 14,523 cases and 25,217 controls. In addition there are 5,818 cases from cohort studies that have not identified matched controls. For genotyping studies, there are 8,555 cases and 12,653 controls and an additional 2,114 cases from cohort studies. The numbers will continue to increase as the cohorts mature. Extensive information on exposures such as BMI, weight change, reproductive variables, menopausal status, use of hormone replacement therapy and oral contraceptives, smoking, presence of other conditions (e.g., diabetes and hypertension), physical activity, and education is available from most studies. All the U.S. studies have information on race and ethnicity. Nearly all the studies collected diet data. Although most of the participants in the studies are Caucasian, reflecting disease incidence and the location of the studies, a large case-control study in Shanghai(5) and the Multiethnic Cohort (6) provide the ability to compare genetic and other risk factors in different racial and ethnic groups and environmental settings. The large numbers of cases and controls and extensive collection of data on risk factors provide a rich resource that will allow us to make substantial contributions to knowledge about the etiology of this disease.

Current and future E2C2 projects

The broad areas identified that would benefit from analysis of pooled data are described below along with ongoing projects in these areas.

Genetic factors, including gene-environment and gene-gene interactions

Investigations of genetic susceptibility to endometrial cancer have consisted mainly of studies of small numbers of variants in candidate genes. In addition, most studies have had small or moderate numbers of cases and controls, although newer studies are larger. Because of the strong association of endometrial cancer with sex-steroid hormones, several studies focused on variants in genes related to hormone biosynthesis and metabolism. Current studies are also investigating other genetic pathways, including DNA repair, insulin-like growth factors, and the stimulation and inhibition of phase I and II enzymes. E2C2 provides a resource to confirm or refute associations found in individual studies; in addition, the large number of subjects allows for study of interactions between genetic variants and established risk factors such as BMI and menopausal status, as well as other exposures such as diet, hormone use, or smoking. Study of combinations of genotypes is also likely to be important and can only be studied with precision through large collaborative efforts.

Variants in *CYP19A1* have been of particular interest; this gene codes for aromatase, which converts androgen precursors to estrone and estradiol. *CYP19A1* is a large gene with >70 known variants in four regions of linkage disequilibrium(7). Associations with endometrial cancer have been found for some polymorphisms (8-11). Recently, strong associations were found between several SNPs and serum levels of estrone and estradiol in healthy postmenopausal women(12). In a recently completed project within E2C2 (13), ten centers genotyped the two SNPs with the strongest relation to these hormones (rs749292 and rs727479). Pooled results from about 5,000 cases and 8,200 controls supported the hypothesis that these variants were related to risk; interaction with BMI was also noted. The large sample size allows for detection of small odds ratios (ORs) and interactions. Future work is planned on other genes important in hormone biosynthesis and metabolism.

Another joint project has been to develop a proposal for a whole genome association study, with the objective of identifying unknown and unexplored genetic factors that are associated with this disease. Participants in this effort include investigators on seven cohort studies and

five case-control studies. The combined sample size (more than 6,000 cases and 6,000 controls in three stages) will provide sufficient power to identify variants with ORs in the range of 1.18 to 1.29, accounting for multiple comparisons, consistent with expected effect sizes for common variants in complex diseases. The large sample size and planned collection of information on several exposures will allow for determination of interactions with exposures.

Diet

Several studies have investigated various aspects of diet, such as consumption of meats, fats, and fruits and vegetables, as risk factors for endometrial cancer. Areas of emerging interest in the nutritional epidemiology of endometrial cancer are the role of non-nutritive components of plant foods (e.g., phytochemicals) (14,15) and other approaches to dietary analysis such as investigation of glycemic index/load. Building on findings from a case-control study in Shanghai (16), showing that tea consumption was associated with reduced risk of endometrial cancer among subgroups of women with specific *CYP19A1* genotypes, a group of E2C2 investigators is studying whether these findings can be replicated in other populations. Investigators representing 24 studies that have collected data on tea consumption, including several that have the capacity to genotype variants in *CYP19A1*, have agreed to join a pooling project to address the hypothesis of gene-environment interaction.

Risk factors for rare tumor types

Survival among women diagnosed with the common endometrioid histologic type of tumors is high; however, women diagnosed with rare histologic types, primarily serous and clear cell tumors, have aggressive disease and poor survival. Almost nothing is known about risk factors for these tumors, although it appears that they are less strongly related to BMI and estrogens (17,18). Since these tumors comprise only about 10% of endometrial cancers, combining studies is the only way to attain large enough sample sizes to study their risk factors. A pooled analysis to examine risk factors for these rare and aggressive tumors is planned using the NIH Small Research Grant (R03) mechanism. More than 20 studies will take part in this joint venture, providing data from about 1,100 women with these rare histologic tumor types, allowing us to address this underinvestigated and important issue in endometrial cancer etiology.

Other etiologic questions

In addition to these active ongoing projects, a number of other areas have been identified that can be addressed effectively by combining data from studies. This includes questions about other pathways, aside from hormonal, that link obesity to risk of endometrial cancer; factors such as insulin-like growth factors and insulin resistance are also likely to be important. The roles of body fat distribution and lifetime changes in weight are poorly understood. The joint association of obesity, body fat distribution, and hormone levels with risk of endometrial cancer is unknown and requires large sample sizes to assess. Although physical activity may be associated with reduced risk of endometrial cancer (19), little if any work has been done on integrating measures of body size, diet, and physical activity to understand how energy balance affects risk of endometrial cancer. Cigarette smoking has been found to reduce risk of endometrial cancer, although the mechanisms are not well understood(20). The large number of cases and controls in studies in E2C2 provides an opportunity to obtain precise estimates of the association of smoking with disease in strata defined by menopausal status, body mass index, use of hormone replacement therapy, and other factors, that can help our understanding of hormonal and other mechanisms behind this association. Data from E2C2 can also be used to determine the effects of intensity and duration of smoking on risk.

Challenges

E2C2 has faced several challenges, both substantive and logistic, in setting up a common database to house the core variables of interest in any study of endometrial cancer. Although individual projects will require specific variables and will seek funding for assembling them, it is important to have the core database established in advance to provide resources for these studies and to avoid duplication of work in the future. Substantive challenges include obtaining agreement on what variables should be included in the database and how they will be defined. Definition of menopausal status poses particular challenges, particularly in women taking hormone replacement therapy at the reference date. The definition of unopposed and opposed estrogen replacement therapy in women taking estrogen and progesterone separately also differs among studies, as does definition of live birth. Combining variables from case-control and cohort studies, with different timing of data collection, is also challenging. Logistically, as for other consortia, substantial investments are required by each investigator to prepare data for pooled analyses, including recoding for consistency when necessary. In addition, resources are needed to check and store data at the coordinating center. It is a particular problem for endometrial cancer that this disease receives little attention in the popular press and there are not at this time strong advocacy groups as there are for many other cancers. We have been fortunate to receive a gift through a donation to Memorial Sloan-Kettering Cancer Center (MSKCC) that is dedicated to establishing the common database; however the lack of advocacy groups will continue to present funding challenges for E2C2 projects.

Some studies face obstacles from Institutional Review Boards or cancer registries in terms of disclosing data to consortia. While maintaining confidentiality and appropriate use of the data are of utmost concern, as consortial research is increasingly needed to address important scientific questions, IRB and cancer registry policies need to grow to facilitate such activities rather than provide unduly time-consuming roadblocks that discourage collaborations.

Concurrent with our activities to form a consortium, a Collaborative Group on Endometrial Cancer was being formed, similar in scope and focus to the Collaborative Group on Hormonal Factors in Breast Cancer, based in Oxford, UK. The main interest of this group is in the association of hormonal, reproductive, and other lifestyle factors. Many individual investigators have joined both consortia. At E2C2, we have made our database variables similar to those for the Oxford Collaborative Group in order to reduce the work of investigators sending data to both groups. To facilitate cooperation, one member of the E2C2 Executive Committee serves as a liaison with the Oxford group, and other members of E2C2 attended their recent meeting on endometrial cancer. Overall, we expect that the two consortia will continue to work together and that the etiologic questions they address will be complementary.

Conclusion

Progress on planned and ongoing projects within E2C2 reflects the interest and enthusiasm of the investigators and the recognition that combining results from individual studies provides a powerful method for answering etiologic questions that cannot be addressed otherwise. E2C2 has provided an infrastructure for investigators to work together to address questions of etiology of this disease and eventually to identify preventive measures. In the future, collaborations established through this resource can be used to address methodological issues such as optimal methods for data collection, management and pooling, and questions related to survivorship. In addition, establishment of a consortium allows for collaboration among investigators with similar interests and exchange of ideas, hypotheses, and methodologies that can benefit future research in this disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Robert Howard Family Foundation for its gift to MSKCC that enabled the establishment of the E2C2 database. We thank Dana Christo, Heather Feigelson, Marc Goodman, Timothy Rebbeck, Daniela Seminara, Elisabete Weiderpass, and Anne Zeleniuch-Jacquotte for their participation in the development of E2C2.

References

1. Cookson WO, Moffatt MF. Genetics of asthma and allergic disease. Hum Mol Genet 2000 Oct;9(16): 2359–64. [PubMed: 11005790]
2. Key TJ, Pike MC. The dose-effect relationship between 'unopposed' oestrogens and endometrial mitotic rate: its central role in explaining and predicting endometrial cancer risk. Br J Cancer 1988 Feb;57(2):205–12. [PubMed: 3358913]
3. Potischman N, Hoover RN, Brinton LA, Siiteri P, Dorgan JF, Swanson CA, et al. Case-control study of endogenous steroid hormones and endometrial cancer. Journal of the National Cancer Institute 1996 Aug 21;88(16):1127–35. [PubMed: 8757192]
4. Lukanova A, Lundin E, Micheli A, Arslan A, Ferrari P, Rinaldi S, et al. Circulating levels of sex steroid hormones and risk of endometrial cancer in postmenopausal women. Int J Cancer 2004 Jan 20;108(3): 425–32. [PubMed: 14648710]
5. Zheng W, CW, Yang G, Fan J, Rothman N, Li HL, Wen W, Ji BT, LI Q, Liu DK, Shu XO, Gao YT. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. American Journal of Epidemiology 2005;162(11):1123–31. [PubMed: 16236996]
6. Kolonel LN, Altshuler D, Henderson BE. The Multiethnic Cohort Study: exploring genes, lifestyle and cancer risk. Nature reviews 2004 Jul;4(7):519–27.
7. Haiman CA, Stram DO, Pike MC, Kolonel LN, Burtt NP, Altshuler D, et al. A comprehensive haplotype analysis of CYP19 and breast cancer risk: the Multiethnic Cohort. Hum Mol Genet 2003 Oct 15;12 (20):2679–92. [PubMed: 12944421]
8. Berstein LM, Imyanitov EN, Kovalevskij AJ, Maximov SJ, Vasilyev DA, Buslov KG, et al. CYP17 and CYP19 genetic polymorphisms in endometrial cancer: association with intratumoral aromatase activity. Cancer Lett 2004 Apr 30;207(2):191–6. [PubMed: 15072828]
9. Paynter RA, Hankinson SE, Colditz GA, Kraft P, Hunter DJ, De Vivo I. CYP19 (aromatase) haplotypes and endometrial cancer risk. Int J Cancer 2005 Aug 20;116(2):267–74. [PubMed: 15800924]
10. Tao MH, Cai Q, Zhang ZF, Xu WH, Kataoka N, Wen W, et al. Polymorphisms in the CYP19A1 (Aromatase) gene and endometrial cancer risk in Chinese women. Cancer Epidemiol Biomarkers Prev 2007 May;16(5):943–9. [PubMed: 17507620]
11. Olson SH, Orlow I, Bayuga S, Sima C, Bandera EV, Pulick K, et al. Variants in hormone biosynthesis genes and risk of endometrial cancer. Cancer Causes Control. 2008 Apr 25; Epub ahead of print.
12. Haiman CA, Dossus L, Setiawan VW, Stram DO, Dunning AM, Thomas G, et al. Genetic variation at the CYP19A1 locus predicts circulating estrogen levels but not breast cancer risk in postmenopausal women. Cancer Res 2007 Mar 1;67(5):1893–7. [PubMed: 17325027]
13. Setiawan VW, Doherty JA, Shu X, Akbari MR, Chen C, De Vivo I, et al. Two estrogen-related variants in CYP19A1 and endometrial cancer risk: A pooled analysis in E2C2. Cancer Epidemiol Biomarkers Prev. 2009 in press.
14. Goodman MT, Wilkens LR, Hankin JH, Lyu LC, Wu AH, Kolonel LN. Association of soy and fiber consumption with the risk of endometrial cancer. Am J Epidemiol 1997 Aug 15;146(4):294–306. [PubMed: 9270408]
15. Horn-Ross PL, John EM, Canchola AJ, Stewart SL, Lee MM. Phytoestrogen intake and endometrial cancer risk. Journal of the National Cancer Institute 2003 Aug 6;95(15):1158–64. [PubMed: 12902445]

16. Xu WH, Dai Q, Xiang YB, Long JR, Ruan ZX, Cheng JR, et al. Interaction of soy food and tea consumption with CYP19A1 genetic polymorphisms in the development of endometrial cancer. Am J Epidemiol 2007 Sep;7(12):1420–30. [PubMed: 17827443]

17. Sherman ME, Sturgeon S, Brinton LA, Potischman N, Kurman RJ, Berman ML, et al. Risk factors and hormone levels in patients with serous and endometrioid uterine carcinomas. Mod Pathol 1997 Oct;10(10):963–8. [PubMed: 9346174]

18. Bjorge T, Engeland A, Tretli S, Weiderpass E. Body size in relation to cancer of the uterine corpus in 1 million Norwegian women. Int J Cancer 2007 Jan 15;120(2):378–83. [PubMed: 17066451]

19. World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington, DC: AICR; 2007.

20. Zhou B, Yang L, Sun Q, Cong R, Gu H, Tang N, et al. Cigarette smoking and the risk of endometrial cancer: a meta-analysis. Am J Med 2008 Jun;121(6):501–8. [PubMed: 18501231]

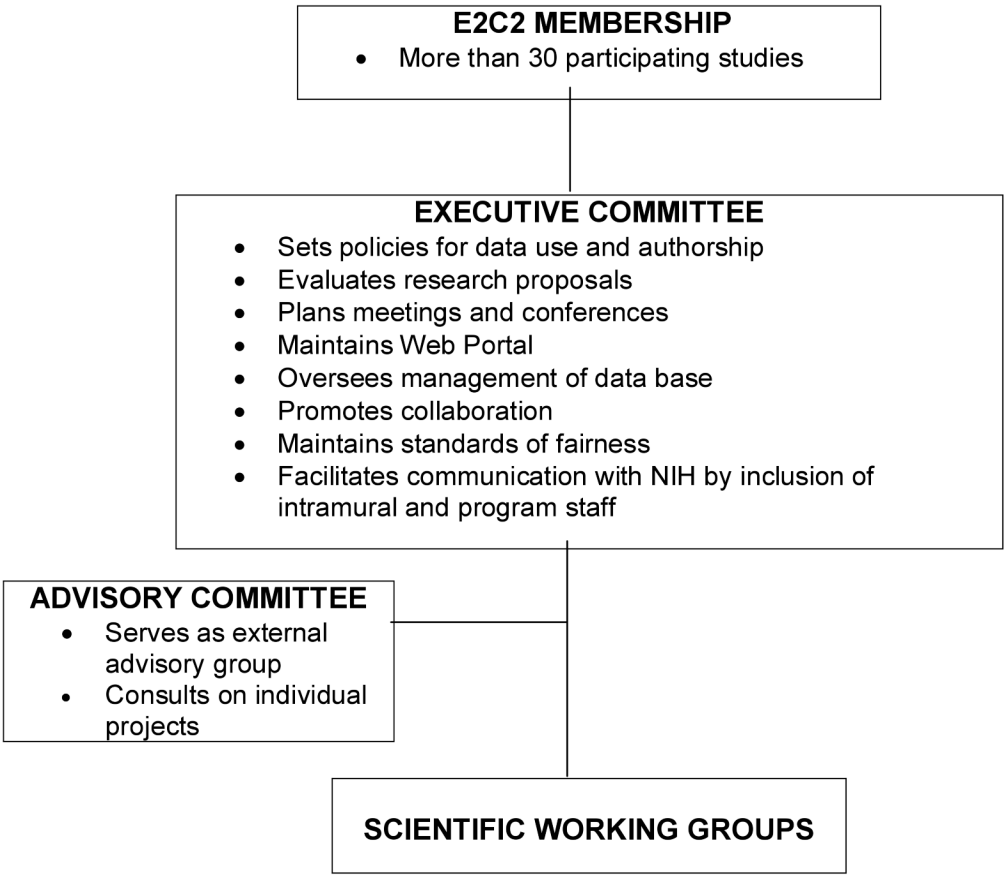

Figure 1.
E2C2 organizational structure

Table 1
Cases and controls in studies in E2C2 (see auxiliary tables for details of individual studies)

	Cases		Controls	
	Total	**With DNA**	**Total**	**With DNA**
Studies that collected DNA				
Case-control studies	7,576	6,675	10,119	8,187
Nested case-control studies	2,514	1,880	5,364	4,466
Cohort studies	2,589	2,114	453,142*	
Studies that did not collect DNA				
Case-control studies	4,433		9,734	
Cohort studies	3,229		321,283*	

* unaffected individuals in cohort studies

Cases and controls from case-control and nested case-control studies (14,523 cases, 25,217 controls)

Cases from cohort studies that have not identified matched controls (5,818 cases)

Cases and controls with DNA from case-control and nested case-control studies (8,555 cases, 12,653 controls)

Cases with DNA from cohort studies that have not identified matched controls (2,114 cases)