HHS Public Access

Author manuscript

Cancer Causes Control. Author manuscript; available in PMC 2016 July 01.

Published in final edited form as:

Cancer Causes Control. 2015 July ; 26(7): 959–972. doi:10.1007/s10552-015-0596-2.

Proceedings of The Second International Molecular Pathological Epidemiology (MPE) Meeting

Shuji Ogino, **Peter T. Campbell**, **Reiko Nishihara**, **Amanda I. Phipps**, **Andrew H. Beck**, **Mark E. Sherman**, **Andrew T. Chan**, **Melissa A. Troester**, **Adam J. Bass**, **Kathryn C. Fitzgerald**, **Rafael A. Irizarry**, **Karl T. Kelsey**, **Hongmei Nan**, **Ulrike Peters**, **Elizabeth M. Poole**, **Zhi Rong Qian**, **Rulla M. Tamimi**, **Eric J. Tchetgen Tchetgen**, **Shelley S. Tworoger**, **Xuehong Zhang**, **Edward L. Giovannucci**, **Piet A. van den Brandt**, **Bernard A. Rosner**, **Molin Wang**, **Nilanjan Chatterjee**, and **Colin B. Begg**

Department of Pathology, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA (SO); Department of Medical Oncology, Dana-Farber Cancer Institute, and Harvard Medical School, Boston, MA, USA (SO, RN, AJB, ZRQ); Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA (SO, RN, KCF, RMT, EJT, SST, ELG, MW); Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA (PTC); Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA (RN, RAI, EJT, BAR, MW); Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA (RN, KCF, ELG); Department of Epidemiology, University of Washington, Seattle, WA, USA (AIP, UP); Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA (AIP, UP); Department of Pathology, Beth Israel Deaconess Medical Center, and Harvard Medical School, Boston, MA, USA (AHB); The Broad Institute, Cambridge, MA, USA (AHB, AJB); Breast and Gynecologic Cancer Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD, USA (MES); Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA (MES, NC); Division of Gastroenterology, Massachusetts General Hospital, Boston, MA, USA (ATC); Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, and Harvard Medical School, Boston, MA, USA (ATC, EMP, RMT, SST, XZ, ELG, BAR, MW); Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA (MAT); Department of Biostatics and Computational Biology, Dana-Farber Cancer Institute, and Harvard Medical School, Boston, MA, USA (RAI); Department of Pathology and Laboratory Medicine, Brown University, Providence, RI, USA (KTK); Department of Epidemiology, Richard M. Fairbanks School of Public Health, Melvin and Bren Simon Cancer Center, Indiana University,

Corresponding author: Shuji Ogino, MD, PhD, MS, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Department of Pathology, Brigham and Women's Hospital, 450 Brookline Ave., Room M422, Boston, MA 02215 USA, Telephone: +1-617-632-1972; Fax: +1-617-582-8558, shuji_ogino@dfci.harvard.edu.

SO, PTC, RN, and AIP contributed equally.

BAR, MW, NC, and CBB contributed equally.

Conflict of Interest: The Second International Molecular Pathological Epidemiology (MPE) Meeting was sponsored in part by Enzymatics, Inc. ATC previously served as a consultant for Bayer Healthcare, Millennium Pharmaceuticals, Pozen Inc, and Pfizer Inc. The work was not funded by Enzymatics, Inc, Bayer Healthcare, Millennium Pharmaceuticals, Pozen Inc, or Pfizer Inc. All of the other authors declare no conflict of interest.

Use of Human Genome Organisation (HUGO) Gene Nomenclature Committee (HGNC)-approved symbols for genes and gene products: We use symbols approved by HGNC and described at www.genenames.org; those include BRAF, CD274, ERBB2, ESR1, FASN, KRAS, MLH1, PDCD1LG2, PGR, PIK3CA, and VHL. Gene names are italicized while names of gene products are non-italicized. Non-official names are described in parenthesis where helpful.

Indianapolis, IN, USA (HN); Department of Epidemiology, Maastricht University, Maastricht, Netherlands (PAvdB); Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA (CBB)

Abstract

Disease classification system increasingly incorporates information on pathogenic mechanisms to predict clinical outcomes and response to therapy and intervention. Technological advancements to interrogate omics (genomics, epigenomics, transcriptomics, proteomics, metabolomics, metagenomics, interactomics, etc.) provide widely-open opportunities in population-based research. Molecular pathological epidemiology (MPE) represents integrative science of molecular pathology and epidemiology. This unified paradigm requires multidisciplinary collaboration between pathology, epidemiology, biostatistics, bioinformatics, and computational biology. Integration of these fields enables better understanding of etiologic heterogeneity, disease continuum, causal inference, and the impact of environment, diet, lifestyle, host factors (including genetics and immunity), and their interactions on disease evolution. Hence, the Second International MPE Meeting was held in Boston in December 2014, with aims to: (1) develop conceptual and practical frameworks; (2) cultivate and expand opportunities; (3) address challenges; and (4) initiate the effort of specifying guidelines for MPE. The meeting mainly consisted of presentations of method developments and recent data in various malignant neoplasms and tumors (breast, prostate, ovarian and colorectal cancers, renal cell carcinoma, lymphoma, and leukemia), followed by open discussion sessions on challenges and future plans. In particular, we recognized need for efforts to further develop statistical methodologies. This meeting provided an unprecedented opportunity for interdisciplinary collaboration, consistent with the purposes of the BD2K (Big Data to Knowledge), GAME-ON (Genetic Associations and Mechanisms in Oncology), and Precision Medicine Initiatives of the U.S.A. National Institute of Health. The MPE Meeting Series can help advance transdisciplinary population science, and optimize training and education systems for 21st century medicine and public health.

Keywords

epidemiologic method; molecular pathologic epidemiology; personalized medicine; systems biology; translational epidemiology; unique disease principle

Introduction

A fundamental premise of epidemiology is that individuals with the same disease name have similar etiologies, and exhibit a similar disease evolution. These important principles (homogeneity and generalizability) constitute a firm basis to generate evidence for practice in clinical medicine and public health. Under this premise, the field of epidemiology has been established as the core scientific discipline of public health. Over recent decades, our improved knowledge of disease pathogenesis has transformed disease classification systems. It is also increasingly evident that pathogenic processes are fundamentally heterogeneous as indicated by the “unique disease principle”.[1, 2] The goal of using molecular disease signatures is to sub-classify disease to improve prediction of disease occurrence and progression for precision medicine and public health. With this evolution, the field of

molecular pathological epidemiology (MPE) has emerged as an integrative interdisciplinary field of molecular pathology and epidemiology.[3, 4] The concept of MPE has been well-recognized in the recent literature.[5–60] Similar ideas and concepts have also been discussed.[61–64] At this point, most molecular pathology tests (hence, MPE studies) focus on neoplastic diseases, although the concepts and approaches of MPE can potentially transform epidemiology in virtually all disease areas including non-neoplastic conditions.[2, 65] MPE can be one of the next steps of genome-wide association studies (GWAS; GWAS-MPE approach), to decipher pathogenic roles of putative disease-causing genetic variants.[4] MPE has also contributed to the development of related paradigms and concepts, including the colorectal continuum model,[66, 67] the etiologic field effect model,[68] and the integrative lifecourse epidemiology-MPE.[69] The advancement and availability of molecular pathology technologies have opened enormous opportunities in population-based research, but also pose considerable challenges, including paucity of interdisciplinary experts and interdisciplinary education programs, and lack of research guidelines specific to MPE.

To advance integrative molecular and population-level health science and to address the unique research challenges specific to the field of MPE, experts in various fields including pathology, epidemiology, genomics, biostatistics, bioinformatics, and computational biology must work together. Along with the BD2K (Big Data to Knowledge), GAME-ON (Genetic Associations and Mechanisms in Oncology) and Precision Medicine Initiatives by the U.S.A. National Institute of Health (NIH), our effort of integrating these seemingly divergent fields can lead to a greater understanding of heterogeneity of the pathogenic process, and the impact of environmental, dietary, lifestyle, and host factors (including genetics and immunity), and their interactions on that process. Brainstorming and discussion fostered at a scientific meeting will lead to the development of new methodologies to address the unique research questions and challenges in this emerging field.

Thus, the First International MPE Meeting (led by Dr. Shuji Ogino) was held as a closed meeting at the Harvard School of Public Health in Boston on April 24, 2013, where 10 investigators gathered. Based on the success and productivity of this small, closed meeting, plans were made for a larger international MPE meeting.

On December 4–5, 2014, the Second International Molecular Pathological Epidemiology (MPE) Meeting, open to the worldwide research community, was held at the Dana-Farber Cancer Institute (Boston, MA). The conference handbook is available as Supplementary Material. There were a total of approximately 150 attendees from 16 countries. The specific aims of this meeting were to: 1) develop a conceptual and practical framework for MPE; 2) cultivate and expand opportunities; 3) address challenges in MPE; and 4) initiate the effort of designing guidelines for MPE research (“STROBE-MPE”), as first proposed in 2012.[70, 71] Making consensus guidelines that will improve study reporting is one of well-established strategies to build a field and enhance its contributions, as evidenced by the development of the guidelines for clinical trials.[72, 73]

Overview session

The meeting started with the overview session (moderated by Dr. Hongmei Nan). Dr. Ogino presented an introductory lecture on MPE, emphasizing its strengths, opportunities, and challenges. MPE research has shown that smoking and obesity are risk factors for different subtypes [microsatellite instability (MSI) and non-MSI subtypes, respectively] of colorectal cancer,[11, 33, 74–84] which is a single entity in conventional epidemiology. This paradigm shift is important in cancer prevention. For example, colonoscopy may be less effective for reducing risk of MSI-high colorectal cancers than for non-MSI cancers,[85, 86] and hence, smokers who are at risk of MSI-high colorectal cancer may need to have different screening strategies. With the simplest scenario of a binary disease subtypes (e.g., A and B), to show etiologic heterogeneity, three hypotheses must be evaluated: (1) the exposure relates to subtype A; (2) the exposure relates to subtype B; and (3) there is heterogeneity between the "(1)" and "(2)" relationships. Notably, many molecular biomarkers are not simple binary measures, and disease can be classified by multiple markers into multiple subtypes. Thus, there exists a need to develop new statistical methods to address etiologic heterogeneity in many different scenarios. As open opportunities, molecular pathology tests have become routine clinical practice in many parts of the world, and we can potentially utilize accumulating disease molecular data.[87]

Dr. Edward Giovannucci presented a lecture on the utility of the MPE approach to enhance causality in epidemiology. One of the strongest criteria used in observational studies to assess causality is the magnitude of the relative risk. Not uncommonly, a risk factor may be associated only with a minority subtype of a given cancer. In that case, the relative risk for all subtypes combined as a singular disease is diluted and may be imperceptible. A recent example is the association of vegetable intake and ESR1 (estrogen receptor 1, ER)-negative breast cancer. In a large pooling study,[88] high intake of vegetables was associated with an 18% lower risk of ESR1-negative breast cancer. As ESR1-negative breast cancers represent only approximately 20% of all breast cancer, no discernible association was apparent for total breast cancer. A similar example is seen for smoking in relation to colorectal cancer risk. Smoking is associated with a two-fold risk for CIMP-high (or MSI-high) colorectal cancer, whereas a null association is observed for non-CIMP (or non-MSI) subtypes.[74–76] As CIMP-high colorectal cancer accounts for 15% to 20% of the total, the association between smoking and total colorectal cancer has been weak-to-modest. These two examples illustrate that important etiologic factors could easily be missed when cancer heterogeneity is not taken into account.

Colorectal and breast cancers

As the two cancers that have been most widely studied using MPE, a session was devoted to research on colorectal and breast cancers (moderated by Dr. Kana Wu). With respect to colorectal cancer,[89–92] Dr. Peter Campbell presented a lecture summarizing MPE research in the area of obesity. Energy balance and metabolism have been implicated in cancer evolution.[93–98] Recently, MPE studies have investigated whether the association of risk factors (including high body mass index, BMI) with colorectal cancer differs by tumor molecular features. While high BMI is a consistent and convincing risk factor for

colorectal cancer overall (relative risks in the range of 1.3 to 1.5 are common for an obese BMI relative to a normal BMI), MPE studies have overall suggested that the associations are stronger for, or even restricted to, non-MSI colorectal cancer (relative risks up to around 2 are common for an obese BMI relative to a normal BMI).[11, 33, 82–84] Other recent studies have shown that these results might be corroborated with expression status of FASN, [99] which has been associated with MSI-high in colorectal cancer.[100] These data also indicate the existence of molecular confounding in MPE, and a need for new methodologies to disentangle correlated molecular biomarkers.

Dr. Andrew Chan presented a lecture on MPE research into the relationship between aspirin use and colorectal cancer. Consistent experimental and epidemiologic evidence indicates that aspirin reduces colorectal cancer risk.[18, 101, 102] Data also support an association between aspirin use and improved outcomes among colorectal cancer survivors. Nonetheless, current clinical guidelines recommend against the routine use of aspirin to prevent colorectal cancer in individuals at average risk largely due to concerns about its potential gastrointestinal toxicity. His group has conducted several MPE studies [103–109] to shed light on the mechanistic basis of aspirin's anti-cancer effect, and to identify tumor, germline, and plasma biomarkers for potential risk stratification to more effectively target aspirin chemoprevention and treatment. In particular, aspirin use appears to be strongly associated with lower mortality among patients with *PIK3CA*-mutated colorectal cancer. [106, 110] Moreover, recent results from secondary analyses of randomized controlled trials of aspirin for cardiovascular prevention has shown that aspirin's effects may extend to benefits for cancers beyond the colon, persuasively making the case for a broader role for aspirin in cancer prevention. In summary, evidence supports a role for aspirin in the prevention and treatment of colorectal cancer, and there will be novel strategies for molecular risk stratification.

With respect to breast cancer, Dr. Montserrat Garcia-Closas presented research into this molecularly-heterogeneous disease. The breast microenvironment and etiologic factors for breast cancer change throughout the life course, in step with hormonal changes during puberty, pregnancy, lactation, and peri-menopausal periods. Although there are many well-established risk factors for breast cancer overall, increasing evidence indicates that such risk factors differ in the magnitude of their impact according to common somatic attributes. In particular, several reproductive risk factors demonstrate different associations with breast cancer subtypes based on their expression of hormone receptors, such as ESR1 (estrogen receptor 1). Recently, GWAS successfully identified many risk variants for breast cancer, and efforts have started to clarify specific associations of risk variants with major and minor subtypes of the disease. A recent study successfully identified risk variants for the ESR1-negative subtype,[111] which had not been identified as risk alleles in GWAS of overall breast cancer. This attests to utility of the MPE approach to discover hidden etiologic factors which have not been uncovered by traditional genetic epidemiology.

Dr. Rulla Tamimi further expanded on the MPE of breast cancer, describing multiple disease subtypes. These subtypes vary in tumor gene expression and phenotype, and are most commonly grouped into four major subtypes: luminal A-like, luminal B-like, ERBB2 (HER2)-positive, and triple-negative (or basal-like). Reproductive and hormonal factors

(e.g., younger age at menarche, older age at first child birth, having fewer children, older age at menopause, and use of hormonal therapy) have been associated with increased risk of luminal A-like breast cancer and, less consistently, luminal B-like cancer.[112, 113] Family history of breast cancer is the only established risk factor for the ERBB2-positive subtype.[114, 115] Younger age at menarche is positively associated with triple-negative breast cancer, while breastfeeding appears to be inversely associated with this disease subtype.[112, 113] Future studies focusing on novel risk factors with consideration for disease heterogeneity may help identify risk factors for less common and more aggressive types of breast cancer. Such studies of less common subtypes will require pooling of data across studies, as individual studies are underpowered to evaluate etiologic heterogeneity.

Special session: hot topics

A special session was held to highlight two emerging hot topics relevant to MPE (moderated by Dr. Ogino). Dr. Matthew Meyerson, a pioneer in cancer genomics, presented a lecture on cancer tissue microbiome analysis. He developed PathSeq which is a computational subtraction method to detect pathogen sequences from next generation sequencing data,[116] and applied this approach in the discovery of *Fusobacterium species*, in particular *Fusobacterium nucleatum*, in colorectal cancer tissue.[117] The amount of *Fusobacterium nucleatum* has been shown to be higher in colorectal cancer tissue compared to adjacent normal colon, and is associated with specific molecular attributes in colorectal cancer tissue: MSI-high and CIMP-high status.[118, 119] *Fusobacterium nucleatum* has also been shown to promote tumorigenesis in a mouse model of colorectal cancer, potentially by inhibiting anti-tumor adaptive T-cell immune response.[120] Thus, tumor tissue microbiome analyses can reveal potential pathogens which can represent both epidemiologic exposures and tumor molecular signatures, and will provide enormous opportunities in MPE research.

Dr. Adam Bass, co-chair for both gastric cancer and esophageal cancer projects in The Cancer Genome Atlas (TCGA), presented a lecture on updates of the gastric TCGA project.[121] There are four major molecular subtypes of gastric carcinoma: EB virus (EBV)-associated, MSI (hypermutator), genomically stable (commonly diffuse histopathology subtype), and chromosomal instability subtypes. Biogeographical differences in tumors within the stomach, as well as histopathological diversity, have also been identified through TCGA and were described. Features of EBV-associated gastric cancer include frequent *PIK3CA* mutation and amplification and up-regulation of CD274 (PD-L1) and PDCD1LG2 (PD-L2), which are immune checkpoint ligands and can be targets of immunotherapy. These findings support the importance of molecular classification for gastric cancers in clinical and epidemiologic research to identify specific risk factors and therapeutic targets. In summary, TCGA findings are useful in designing large-scale MPE studies on gastric cancers.

MPE pooling projects

Considering the unique disease (or tumor) principle, it is necessary to examine a large number of cases, most likely by designing pooling consortium projects; thus, a session was devoted to existing pooling projects that have facilitated MPE research (moderated by Dr. Liam Murray). Dr. Lindsay Morton described the InterLymph Consortium, a pooling project

on the epidemiology of lymphomas. InterLymph was initiated in 2001 and presently includes 20 studies with 17,500 cases of non-Hodgkin lymphomas (NHLs) and 23,000 controls.[122] Despite the challenge of harmonizing data across the different studies, MPE research through InterLymph has demonstrated epidemiologic similarities and differences across NHL subtypes. For example, autoimmune diseases, hepatitis and alcohol are risk factors for T-cell NHLs, marginal zone lymphoma, Burkitt lymphoma, diffuse large B-cell lymphoma, while genetic variants (identified by GWAS) are the only established risk factors for chronic lymphocytic leukemia/small lymphocytic lymphoma, follicular lymphoma, and mantle cell lymphoma. As InterLymph matures, this resource can generate a wealth of MPE data.

Dr. Ulrike Peters presented a lecture on the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO). GECCO was first funded by the NIH as a GWAS consortium to characterize genetic susceptibilities to colorectal cancer and modifying effects of gene-by-environment interaction. The newest GECCO U01 project, “Molecular Pathological Epidemiology of Colorectal Cancer”, was recently successfully renewed in 2014. The renewal U01 grant has MPE projects on various exposures and standard tumor molecular biomarkers, including mutations in *KRAS*, *BRAF*, and *PIK3CA*, and MSI and CIMP statuses. The aims of this ongoing collaborative effort also include analyses of over 4000 colorectal cancers by targeted sequencing of 200 driver genes. As most of the driver genes are only mutated in a small fraction of the cases (most are expected to be mutated in <5% of cases), it is critical to group mutated genes in meaningful ways, such as mutated pathways to ensure sufficient statistical power for subtype analysis. Furthermore, deep targeted sequencing will allow identifying mutations in small fractions of the tumor, to investigate tumor heterogeneity. These novel data will require novel statistical analytical approaches.

The lectures were followed by a panel discussion (by Drs. Lindsay Morton, Ulrike Peters, Stephanie Smith-Warner, and Piet van den Brandt). The panelists emphasized the importance of harmonization of tumor molecular data in MPE pooling projects. Dr. Piet van den Brandt presented results on pooling of MPE data from two prospective cohort studies: the Netherlands Cohort Study and the Melbourne Collaborative Cohort Study. These collaborative projects have shown consistent data with improved robustness of findings on anthropometric factors and colorectal cancer risk according to tumor status of MSI and *BRAF* mutation.[11]

Statistical opportunities and challenges in MPE

Methodologic issues and statistical challenges in MPE were highlighted in a session (moderated by Dr. Donna Spiegelman). Dr. Colin Begg presented statistical methods to address etiologic heterogeneity. His talk focused particularly on the discovery of etiologically distinct sub-types, a task that is conceptually and computationally challenging because of the complexity of the somatic portraits of individual tumors and of the corresponding profiles of risk factors. To permit an organized framework for investigating etiological heterogeneity he proposed a scalar statistical measure that captures quantitatively the degree of heterogeneity exhibited by any candidate set of tumor sub-types. This facilitates the exploration of countless sub-typing options to identify the set or sets that

possess the greatest evidence of heterogeneity.[64] He also explained the unique insights that can be obtained by investigating and correlating the mutational profiles of double primary malignancies,[123] and illustrated the concepts using studies of breast cancer, melanoma and kidney cancer.

Dr. Molin Wang described statistical methods to study etiologic heterogeneity using categorical, ordinal and multi-marker classifiers. Methods were reviewed for scenarios where disease subtypes are categorical in cohort studies, matched and unmatched case-control studies, and case-case study designs. New methods were discussed for ordinal disease subtypes. For analyzing studies with disease subtypes defined by multiple categorical and/or ordinal markers, Dr. Wang presented a meta-regression method which uses existing statistical software for the mixed model analysis.[124] That method can be used to test for heterogeneity across multiple disease subtypes classified by multiple markers, to assess whether the exposure-subtype associations are different across subtypes defined by one marker while controlling for other markers, and to evaluate whether the difference in exposure-subtype association across subtypes defined by one marker depends on any other marker.

Dr. Nilanjan Chatterjee presented on the development of methodologies to address etiologic heterogeneity,[125, 126] including one novel method: Association Analysis based on Subset (ASSET). ASSET was described as a means for detecting disease-exposure associations in the presence of heterogeneous subtypes. The method explores all possible subsets of disease subtypes and asses significance of the association for the best subset after accounting for multiple-testing adjustment using an efficient procedure. The method has robust power compared to alternative approaches for detecting overall association in the presence of etiologic heterogeneity and produces readily interpretable results through identification of the disease types that drive a given association. Examples were shown from a recent application of the method to cross-cancer analysis of GWAS data.

Dr. Bernard Rosner described statistical methods to address etiologic heterogeneity and disentangle molecular confounding by multiple correlated biomarkers. The correlation of markers makes it difficult to ascribe interaction effects of a risk factor with a specific marker without considering other markers. In addition, there are often tumor markers with missing data for some subjects. To resolve this, Dr. Rosner discussed a two-stage regression approach.[127] At a first stage, markers are cross-classified and the Beta coefficient for a given risk factor is obtained for the outcome defined by each combination of markers. At a second stage, a regression is performed of the regression coefficients from the first stage on a vector of tumor characteristics. This yields a test of heterogeneity of effects of a risk factor by a specific marker while controlling for the effects of other markers and an associated estimate of effect characterized by an “adjusted hazard ratio.” Dr. Rosner also compared different methods for handling missing marker data including complete case, missing indicator, inverse probability weighting, and multiple imputation methods. Additional simulation studies are needed to compare these methods for handling missing marker data.

Dr. Eric Tchetgen Tchetgen presented a lecture on unique methodological considerations in the face of outcome heterogeneity. He first discussed the so-called obesity paradox. In an

example of this paradox, obesity is associated with better clinical outcomes among patients with a disease for which obesity has been shown to be a risk factor. This paradox can be explained by considering heterogeneity of disease. He then focused on the situations in which a risk factor is important for some subtypes but not for others. He discussed the possible underestimation of association if the standard nominal polytomous logistic regression models are used when the disease is not rare. To avoid underestimation, he suggested using the subtype-specific logistic regression model for each subtype, with a constraint that the sum of the subtype-specific probability is one.

High dimensional data and cellular heterogeneity

Reflecting the increasingly complex and high-dimensional nature of MPE data, we had a session on bioinformatics and computational pathology (moderated by Dr. Reiko Nishihara). Dr. Rafael Irizarry presented a lecture on cellular heterogeneity, batch effects and confounding. Technical artifacts and measurement errors can become serious confounders which occur due to the difference in tissue collection design, laboratory conditions, etc. among samples and batches of collection and processing.[128] These factors can easily lead to spurious findings which have nothing to do with biological implications. Variability in cellular composition can be another source of serious confounding, for example, in epigenetic research on specimens with mixture of different cell types.[129] An example of data was presented from a study to identify DNA methylation profiles using peripheral blood cells, which consist of many different cell types in varying proportions among individuals and among different time points even within one individual. Appropriate statistical and experimental solutions should be considered to adjust for batch effects and to estimate the relative proportion of cell types in tissue or blood specimens.

Dr. Karl Kelsey described the use of epigenetic markers to differentiate and quantify specific cell types in a cell mixture. The epigenome, including DNA methylation, is responsible for the overall control of gene expression and can be altered by the environment. DNA methylation is associated both with lineage specific differentiation in development and somatic differentiation of stem cells throughout life. Epigenetic plasticity might contribute to numerous environmentally-associated adverse health outcomes. The epidemiologic literature now includes numerous reports of highly significant associations of both environmental exposures and disease states with changes in the DNA methylation profile of blood cells. In assessing the nature of these associations, Dr. Kelsey noted consistent observations indicating that the vast majority of the variation can be explained by accounting for shifts in leukocyte subtypes.[130] Dr. Kelsey showed how it can be exploited in epidemiologic studies (amenable to use with archived blood samples), providing a window into the assessment of an individual immune profile.

Dr. Andrew Beck presented on computational pathology to address disease heterogeneity. Inter-tumoral heterogeneity is a major challenge in MPE, and there is a need for the development of new computational and statistical approaches for significance testing for Omics analyses in large, heterogeneous patient populations. Dr. Beck described two computational methods he recently developed to address this challenge. The first approach, Significance Analysis of Prognostic Signatures (SAPS),[131] is a method for identifying

prognostic gene sets that: 1) stratify patients into groups that show significant survival differences; 2) perform significantly better than random gene-sets at predicting patient survival; and 3) are enriched for prognostic genes. He also described his recently developed Earth Movers Distance Omics (EMDomics) method, which uses the Earth Movers Distance to identify genes that are differentially expressed between heterogeneous classes of samples. Using both simulations and real biological data, Dr. Beck has shown that EMDomics outperforms competing approaches for the identification of genes differentially expressed between heterogeneous groups. The SAPS and EMDomics R packages should be useful for prognostic studies and two-class significance testing (respectively) in MPE studies.

Pathology and pathogenesis

Understanding of disease pathology and pathogenesis is central to MPE,[61] and was the focus of another session (moderated by Dr. Paul Lochhead). Dr. Mark Sherman and Dr. Melissa Troester presented related lectures on pathogenic insights into breast cancer development, with a focus on research into terminal duct lobular units (TDLUs). The morphological and molecular characteristics of “normal appearing” tissues that harbor potentially carcinogenic alterations, collectively referred to herein as “molecular histology,[132] undergo changes over the life course. Environmental exposures interact with normal tissues to produce physiological functions, benign alterations, and in some instances, carcinogenesis. A similar concept has been recently consolidated as “etiologic field effect”.[68] Studying the transition from normal to the earliest identifiable preneoplastic changes represents an important opportunity to better understand the underlying biology of risk. TDLUs are the structures of the breast that produce milk and also give rise to breast cancer precursors. TDLUs undergo a range of morphological changes related to puberty, pregnancy, lactation, aging, and menopause,[133, 134] and can also accompany altered molecular histology and a change in mammographic density. Studies have found that, among women with benign breast biopsy, reduced levels of TDLU involution in the surrounding normal breast are associated with breast cancer risk.[135, 136] Thus, the topic of “molecular histology” and MPE are linked, and cancer precursor lesions can be used as an intermediary phenotype to examine carcinogenic processes in MPE studies.[137] Studies have shown that the dominant expression subtypes of normal breast tissue are highly correlated with both mammographic density and histology, suggesting that such expression patterns could follow patterns of risk, and moreover, that some of the pathways driving histological change could be identified in gene expression studies.[138, 139] Further research is needed to understand key associations between normal appearing breast epithelium, with or without specific molecular alterations, and breast cancer, including: 1) possible short-term increased risk following a live birth; 2) the long-term protection of pregnancy for late onset cancers (mainly hormone receptor positive); 3) the possible relationship of giving birth and increased risk of basal-like breast cancers and 4) the suggested protective effect of breastfeeding, especially for basal-like cancer.[140] In tandem, mechanistic studies conducted in animal models and based on insights from gene expression profiling may deepen insight into the molecular processes that drive TDLU development, involution and carcinogenesis, and may suggest interventions that can modify risk.

MPE of various cancers

Day two of the Meeting began with a session devoted to various cancers not covered in discussions from the first day (moderated by Dr. Amanda Phipps). Dr. van den Brandt presented research into the MPE of renal cell carcinoma with respect to heterogeneity in both etiology and prognosis. Risk factors associated with clear cell renal cell carcinoma appear to differ by *VHL* mutation status. MPE research on renal cell carcinoma has also given clues to the so-called obesity paradox.[141] The obesity paradox refers to the fact that, while obesity is a risk factor for a certain disease (such as renal cell carcinoma), obesity is associated with better clinical outcome among some individuals with the disease. Renal cell carcinomas in obese individuals are more likely tumors with low-level FASN expression, which appear to be an indolent subtype associated with better outcome.

Dr. Lorelei Mucci described her MPE research into prostate cancer. Prostate cancer follows a disease model where premalignant and malignant tumors can be detected by screening or symptomatology, and benign and malignant tumors form histopathological and clinical spectra. Therefore, unlike with many other cancers, prostate cancer cases are often sub-classified into lethal and non-lethal subtypes based on clinical behavior.[142] Dr. Mucci presented data from the Health Professionals Follow-up Study and Physicians' Health Study on obesity and lethal prostate cancer.[143] Whole genome mRNA profiling has been used to conduct gene expression profiling of prostate cancers in these cohorts, and genes related to chromatin remodeling pathway and RNA processing and metabolism pathway have been shown to be differentially expressed in lethal vs. non-lethal subtypes. Molecular signatures for aggressive versus indolent cancers may be useful in clinical management of prostate cancer.

Dr. Shelley Tworoger presented a lecture on the MPE of ovarian cancer. As with other cancers under discussion, ovarian cancer is a group of heterogeneous diseases, and risk factor associations appear to differ by tumor subtype.[144, 145] Because there is an issue of limited statistical power in each subtype in any single study, consortia are needed to fully explore subtype-specific associations with exposures and clinical outcomes. In ovarian cancer, meaningful metrics to categorize disease can be obtained from pathology reports (e.g., histopathologic features) and other data sources (e.g., death information). Considering multiple metrics of disease heterogeneity (e.g., histopathology, molecular pathology, and time to death) can provide a deeper insight into etiology and help target prevention recommendations to the most aggressive forms of ovarian cancer.

Open discussion sessions

To maximize the utility of the International MPE Meeting, and to augment the invited lecture sessions, emphasis was placed on open discussion sessions where experts in diverse disciplines shared their perspectives and brainstormed ideas for new directions, priorities, and collaborations. We systematically discussed a number of important opportunities and issues in MPE, as described in the following sections.

Interdisciplinary education (moderated by Drs. Troester and Mucci)

Paucity of interdisciplinary education and training has been a challenge in MPE. Because molecular pathology has become pervasive in all areas in medicine and public health, doctoral competencies for epidemiologists in terms of adequate knowledge and skills on both pathology and epidemiology should be defined. In particular, knowledge in multiple areas including biology, physiology, pathology, clinical medicine, statistics, and bioinformatics is important. As technologies continue to advance, skills to update own knowledge and competencies are essential. Meeting attendees suggested possible workshops for interested students and junior investigators, and the development of online resources. Attendees also emphasized the importance of transdisciplinary collaborations in facilitating the interdisciplinary training so important to MPE.

Study design and statistical methods (moderated by Drs. Wang and Aya Kuchiba)

Various topics related to the statistics of MPE research were discussed, especially themes that were not adequately discussed by the lectures (e.g., including missing data and measurement error in disease subtyping biomarkers). Challenges to the analysis of high-dimensional molecular marker data and rare subtypes were highlighted as research priorities. Other challenges discussed included various sources of bias caused by pre-analytic, analytic and post-analytic processes in pathology procedures and data; concerns surrounding multiple hypothesis testing due to the evaluation of multiple exposures and outcomes (disease subtypes); timing and latency issues with respect to exposures and molecular events in cells and tissues. In summary, meeting attendees expressed the importance of addressing challenges in study design and statistics, in order to fully exploit enormous opportunities which await investigators in MPE.

Routine collection of molecular pathology data (moderated by Drs. Beck and Sherman)

Moderators highlighted routine pathology procedures from tissue handling to molecular data generation. Tissue life-cycle includes sampling, fixation, processing, assay, scoring/interpretation, database, and data analyses. The availability of tissues for assay is limited by clinical requirements. Compared to disease tissue (e.g., cancer tissue) which is collected for pathology diagnosis and patient management, normal (or pre-disease) tissue is not easy to obtain from individuals. To resolve this issue, ultimately, *in vivo* pathology technologies will become available for *in vivo* monitoring of pathologic processes, and transform all areas of population health sciences. Intra-tumor heterogeneity was also discussed as an unresolved issue in epidemiology. It is an active area of investigation, and evidence is accumulating for its role in resistance of cancer to various types of drugs including targeted therapeutic agents. In summary, there are widely open opportunities in utilization of tissue resources in epidemiology and population health sciences.

Pooling MPE data (moderated by Drs. Morton, Peters, van den Brandt, and Smith-Warner)

Continuing conversations initiated in earlier sessions, several investigators with experience in pooling projects led a discussion on the challenges (and advantages) of such collaborative efforts. MPE inherently deals with disease subtypes which have smaller sample sizes than an overall disease sample. Therefore, pooling projects are essential in MPE to increase sample sizes, and examine generalizability and effect modification of exposure – disease subtype associations. Challenges include standardization and reproducibility of biomarker measurements, and funding which is necessary to conduct centralized laboratory tests. Ongoing MPE pooling projects such as InterLymph and GECCO will continue to provide guidance in other consortium initiatives.

Research guideline development (STROBE-MPE) (moderated by Dr. Ogino)

Dr. Ogino presented introductory slides on background of the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) initiative,[146, 147] and suggested the need for MPE guidelines. There is an extension of STROBE in molecular epidemiology (STROBE-ME) Guidelines;[148, 149] however the STROBE-ME guideline does not address many of the challenges in MPE which were discussed in the Meeting. Thus, the “STROBE-MPE” guideline project has been proposed,[70, 71] and the STROBE group has agreed with the MPE group to have Dr. Matthias Egger as a liaison for the STROBE-MPE project. It was discussed whether a recommendation (rather than a guideline) by an expert panel might serve for the research community. At this time, as a relatively young field, MPE researcher needs to accumulate more data, and investigators need to develop new methodologies to address etiologic heterogeneity. Once made, a recommendation or guideline must be updated to keep up technological advancements.

Possible collaboration projects (moderated by Drs. Giovannucci and Wang)

The purpose of this session was to conduct brainstorming of experts in diverse disciplines. Since molecular pathology tests have become routine clinical practice, molecular biomarker data are accumulating in hospitals. Routine collection of such data may be very useful. Collaboration can be possible for molecular pathological epidemiologists with researchers in areas such as causal inference, comparative effectiveness research and health communication. Collaboration among statisticians can be effective for solving statistical issues. Possibly, a consortium can form to ensure the external validity and reproducibility of new statistical methods.

Relationships with other societies (moderated by Drs. Murray, Campbell, and van den Brandt)

To continue the International MPE Meeting Series, attendees discussed the worth of synergizing with activities of other well-established societies, including those in epidemiology, pathology, laboratory medicine, statistics, cancer research, oncology and other clinical disciplines. Since routine molecular analyses on disease are most commonly

conducted on cancer and neoplastic diseases, much MPE research currently focuses on neoplastic diseases. Though many societies were discussed, in particular, international societies have broad influence on global health, and we can consider those as potential candidate societies to which we can relate; those include American Association for Cancer Research (AACR), Society for Epidemiologic Research (SER), American College of Epidemiology (ACE), International Epidemiological Association (IEA), United States and Canadian Academy of Pathology (USCAP), American Society for Investigative Pathology (ASIP), and Association for Molecular Pathology (AMP).

Planning the Third International MPE Meeting (moderated by Drs. Phipps, Ogino, and Elizabeth Poole)

Based on comments offered by Meeting attendees, it was suggested that the International MPE Meeting Series continue to be an interactive forum of combining succinct lectures and discussion sessions. Possibilities of settings dates close to another scientific society or consortium meeting were discussed, as were possible meeting themes, location, and issues of timing, with the goal of attracting diverse groups of researchers including pathologists, epidemiologists, biostatisticians and bioinformaticians. It was suggested that incorporating hands-on workshops may increase the value of the meeting. Attendees tentatively agreed that the Third International Meeting be held in Boston in Spring 2016. In the future, the International MPE Meeting may be held in other cities.

Meeting abstracts, posters, and awards

Rapid dissemination of new research findings and knowledge is an important purpose of our meeting. Hence, a fraction of participants submitted abstracts, and made posters for distribution at the meeting. We also projected poster slides during each break time. For presenters who applied for an award, poster referees discussed findings with them and scored merits of each poster presentation. Awards were made to recognize the efforts of junior scientists in conducting high quality interdisciplinary MPE research, and to facilitate further transdisciplinary integration. Referees selected the following awardees for their outstanding abstract and poster presentations: Molecular Pathological Epidemiology Rising Investigator Award (junior faculty) to Drs. Aditi Hazra and Xuehong Zhang; Molecular Pathological Epidemiology Trainee Award (post-doctoral fellow or other trainee) to Drs. Mingyang Song and Atsuhiro Masuda. We plan to continue to provide these opportunities at The International MPE Meeting Series.

Meeting evaluation by participants

Participants filled out an evaluation form to assess quality of the Second International MPE Meeting with respect to the overall program, topics, speakers, discussion sessions, and logistics. On a scale of 1 (poor) to 5 (excellent), average scores were 4.5 for the overall program, 4.7 for the speakers and discussions, and 4.3 for logistics, indicating overall success of the meeting. We also provided the participants with a write-in space for comments. Reflecting attendee satisfaction with the quality of the Meeting, we herein quote some examples: “This is a great meeting”; “Really enjoyed the built-in discussions”;

"Timing well coordinated"; "It was great to have a session on statistical analyses"; "Many areas well explored." In terms of room for improvement in future meetings, attendees commented as follows: "Bring more pathologists and touch more on bioinformatics"; "Would be good to provide tea as well and provided food for all"; "There is no capacity for big posters." In summary, evaluations and feedback by the participants indicated their general satisfaction and enthusiasm, and it is the basis for our plan to organize and have the Third International MPE Meeting in Spring 2016.

Conclusions

The Second International MPE Meeting brought together experts in divergent fields who are working on big data of exposures, molecular pathology of disease, and disease evolution in population-based settings. Brainstorming and discussions fostered by the meeting are and will be helpful for the meeting participants in pursuit of this relatively young field. In the foreseeable future, molecular pathology tests will become prevalent in many different diseases, and make disease molecular data widely available. Because heterogeneity of pathogenic processes is an undisputable phenomenon, medical and health research must consider this fundamental nature of disease, along with the disease continuum theory.[150] Thus, the MPE paradigm should become ubiquitous in all fields of population health sciences. We also recognize efforts to further develop statistical methodologies are needed. Given advancements of various omics technologies to analyze diseases, new opportunities are widely open to study virtually any disease. We look forward to continuing brainstorming and discussions at the Third International MPE Meeting in spring of 2016.

Acknowledgments

We thank all of the members of the Program Committee, the speakers, the discussants, and the other participants of The Second International Molecular Pathological Epidemiology (MPE) Meeting on December 4 to 5 in 2014 in Boston, MA, USA. We thank the Dana-Farber Cancer Institute (Edward J. Benz, Jr., President and CEO) for providing the meeting venue; and the Department of Pathology, the Brigham and Women's Hospital (Jeffrey A. Golden, Chairperson), and Enzymatics, Inc., for providing meals and refreshments, respectively. We also thank the Department of Pathology, the Brigham and Women's Hospital (Jeffrey A. Golden, Chairperson), the Department of Epidemiology, the Harvard T.H. Chan School of Public Health (Michelle A. Williams, Chairperson), the Dana-Farber Harvard Cancer Center (Giovanni Parmigiani, Meir J. Stampfer, Lorelei A. Mucci, Deborah Schrag, and Charles S. Fuchs; Program Leaders), and the Channing Division of Network Medicine, the Department of Medicine, the Brigham and Women's Hospital (Edwin K. Silverman, Division Chief) for providing morale supports and helping in announcements. This work was supported in part by grants from the U.S.A. National Institute of Health (NIH) [R01 CA151993 (to SO), K07 CA190673 (to RN), R01 CA137178 (to ATC), K24 DK098311 (to ATC), and K07 CA172298 (to AIP)], and the Friends of the Dana-Farber Cancer Institute (to SO). ATC is Damon Runyon Clinical Investigator. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH. The funders (including Enzymatics, Inc.) did not have any role in planning the meeting, the decision to submit the manuscript for publication, or the writing of the manuscript.

Abbreviations

BD2K	big data to knowledge
BMI	body mass index
CIMP	CpG island methylator phenotype
GAME-ON	genetic associations and mechanisms in oncology

GECCO	the Genetics and Epidemiology of Colorectal Cancer Consortium
GWAS	genome-wide association study
MPE	molecular pathological epidemiology
MSI	microsatellite instability
NIH	National Institute of Health
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology

References

1. Ogino S, Fuchs CS, Giovannucci E. How many molecular subtypes? Implications of the unique tumor principle in personalized medicine. Expert Rev Mol Diagn. 2012; 12:621–628. [PubMed: 22845482]
2. Ogino S, Lochhead P, Chan AT, et al. Molecular pathological epidemiology of epigenetics: Emerging integrative science to analyze environment, host, and disease. Mod Pathol. 2013; 26:465–484. [PubMed: 23307060]
3. Ogino S, Stampfer M. Lifestyle factors and microsatellite instability in colorectal cancer: The evolving field of molecular pathological epidemiology. J Natl Cancer Inst. 2010; 102:365–367. [PubMed: 20208016]
4. Ogino S, Chan AT, Fuchs CS, Giovannucci E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut. 2011; 60:397–411. [PubMed: 21036793]
5. Jacobs R, Voorneveld P, Kodach L, Hardwick J. Cholesterol metabolism and colorectal cancers. Current opinion in pharmacology. 2012; 12:690–695. [PubMed: 22884562]
6. Curtin K, Slattery ML, Samowitz WS. CpG island methylation in colorectal cancer: past, present and future. Pathology Research International. 2011; 2011:902674. [PubMed: 21559209]
7. Hughes LA, Simons CC, van den Brandt PA, et al. Body size, physical activity and risk of colorectal cancer with or without the CpG island methylator phenotype (CIMP). PLoS One. 2011; 6:e18571. [PubMed: 21483668]
8. Hughes LA, Khalid-de Bakker CA, Smits KM, et al. The CpG island methylator phenotype in colorectal cancer: Progress and problems. Biochimica et biophysica acta. 2012; 1825:77–85. [PubMed: 22056543]
9. Iwagami S, Baba Y, Watanabe M, et al. Pyrosequencing Assay to Measure LINE-1 Methylation Level in Esophageal Squamous Cell Carcinoma. Ann Surg Oncol. 2012; 19:2726–2732. [PubMed: 22187122]
10. Limburg PJ, Limsui D, Vierkant RA, et al. Postmenopausal Hormone Therapy and Colorectal Cancer Risk in Relation to Somatic KRAS Mutation Status among Older Women. Cancer Epidemiol Biomarkers Prev. 2012; 21:681–684. [PubMed: 22337533]
11. Hughes LA, Williamson EJ, van Engeland M, et al. Body size and risk for colorectal cancers showing BRAF mutation or microsatellite instability: a pooled analysis. Int J Epidemiol. 2012; 41:1060–1072. [PubMed: 22531127]
12. Ku CS, Cooper DN, Wu M, et al. Gene discovery in familial cancer syndromes by exome sequencing: prospects for the elucidation of familial colorectal cancer type X. Mod Pathol. 2012; 25:1055–1068. [PubMed: 22522846]
13. Rex DK, Ahnen DJ, Baron JA, et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol. 2012; 107:1315–1329. [PubMed: 22710576]
14. Koshiol J, Lin SW. Can Tissue-Based Immune Markers be Used for Studying the Natural History of Cancer? Ann Epidemiol. 2012; 22:520–530. [PubMed: 22481034]

15. Fini, L.; Grizzi, F.; Laghi, L. Adaptive and Innate Immunity in Colorectal Cancer Progression. In: Ettarh, R., editor. Colorectal Cancer Biology – From Genes to Tumor. InTech; 2012. p. 323-340.
16. Gay LJ, Mitrou PN, Keen J, et al. Dietary, lifestyle and clinico-pathological factors associated with APC mutations and promoter methylation in colorectal cancers from the EPIC-Norfolk Study. J Pathol. 2012; 228:405–415. [PubMed: 22864938]
17. Galon J, Franck P, Marincola FM, et al. Cancer classification using the Immunoscore: a worldwide task force. Journal of translational medicine. 2012; 10:205. [PubMed: 23034130]
18. Chia WK, Ali R, Toh HC. Aspirin as adjuvant therapy for colorectal cancer-reinterpreting paradigms. Nature reviews. Clinical oncology. 2012; 9:561–570. [PubMed: 22910681]
19. Dogan S, Shen R, Ang DC, et al. Molecular Epidemiology of EGFR and KRAS Mutations in 3026 Lung Adenocarcinomas: Higher Susceptibility of Women to Smoking-related KRAS-mutant Cancers. Clin Cancer Res. 2012; 18:6169–6177. [PubMed: 23014527]
20. Spitz MR, Caporaso NE, Sellers TA. Integrative cancer epidemiology–the next generation. Cancer discovery. 2012; 2:1087–1090. [PubMed: 23230187]
21. Shanmuganathan R, Nazeema Banu B, Amirthalingam L, Muthukumar H, Kaliaperumal R, Shanmugam K. Conventional and Nanotechniques for DNA Methylation Profiling. J Mol Diagn. 2013; 15:17–26. [PubMed: 23127612]
22. Rosty C, Young JP, Walsh MD, et al. Colorectal carcinomas with KRAS mutation are associated with distinctive morphological and molecular features. Mod Pathol. 2013; 26:825–834. [PubMed: 23348904]
23. Weijenberg MP, Hughes LA, Bours MJ, Simons CC, van Engeland M, van den Brandt PA. The mTOR Pathway and the Role of Energy Balance Throughout Life in Colorectal Cancer Etiology and Prognosis: Unravelling Mechanisms Through a Multidimensional Molecular Epidemiologic Approach. Current nutrition reports. 2013; 2:19–26. [PubMed: 23396869]
24. Buchanan DD, Win AK, Walsh MD, et al. Family History of Colorectal Cancer in BRAF p.V600E mutated Colorectal Cancer Cases. Cancer Epidemiol Biomarkers Prev. 2013; 22:917–926. [PubMed: 23462926]
25. Burnett-Hartman AN, Newcomb PA, Potter JD, et al. Genomic aberrations occuring in subsets of serrated colorectal lesions but not conventional adenomas. Cancer Res. 2013; 73:2863–2872. [PubMed: 23539450]
26. Alvarez MC, Santos JC, Maniezzo N, et al. MGMT and MLH1 methylation in Helicobacter pylori-infected children and adults. World journal of gastroenterology: WJG. 2013; 19:3043–3051. [PubMed: 23716983]
27. Hagland HR, Berg M, Jolma IW, Carlsen A, Soreide K. Molecular Pathways and Cellular Metabolism in Colorectal Cancer. Digestive surgery. 2013; 30:12–25. [PubMed: 23595116]
28. Zaidi N, Lupien L, Kuemmerle NB, Kinlaw WB, Swinnen JV, Smans K. Lipogenesis and lipolysis: The pathways exploited by the cancer cells to acquire fatty acids. Progress in lipid research. 2013; 52:585–589. [PubMed: 24001676]
29. Abbenhardt C, Poole EM, Kulmacz RJ, et al. Phospholipase A2G1B polymorphisms and risk of colorectal neoplasia. International journal of molecular epidemiology and genetics. 2013; 4:140–149. [PubMed: 24046806]
30. Hughes LA, Melotte V, de Schrijver J, et al. The CpG island methylator phenotype: what's in a name? Cancer Res. 2013; 73:5858–5868. [PubMed: 23801749]
31. Bae JM, Kim JH, Cho NY, Kim TY, Kang GH. Prognostic implication of the CpG island methylator phenotype in colorectal cancers depends on tumour location. Br J Cancer. 2013; 109:1004–1012. [PubMed: 23900220]
32. Amirian ES, Petrosino JF, Ajami NJ, Liu Y, Mims MP, Scheurer ME. Potential role of gastrointestinal microbiota composition in prostate cancer risk. Infectious agents and cancer. 2013; 8:42. [PubMed: 24180596]
33. Hoffmeister M, Blaker H, Kloor M, et al. Body mass index and microsatellite instability in colorectal cancer: a population-based study. Cancer Epidemiol Biomarkers Prev. 2013; 22:2303–2311. [PubMed: 24127414]
34. Araujo RF Jr, Lira GA, Guedes HG, et al. Lifestyle and family history influence cancer prognosis in Brazilian individuals. Pathology, research and practice. 2013; 209:753–757.

35. Esterhuyse MM, Kaufmann SH. Diagnostic biomarkers are hidden in the infected host's epigenome. Expert Rev Mol Diagn. 2013; 13:625–637. [PubMed: 23895131]
36. Zhu Y, Yang SR, Wang PP, et al. Influence of pre-diagnostic cigarette smoking on colorectal cancer survival: overall and by tumour molecular phenotype. Br J Cancer. 2014; 110:1359–1366. [PubMed: 24448365]
37. Hagland HR, Soreide K. Cellular metabolism in colorectal carcinogenesis: Influence of lifestyle, gut microbiome and metabolic pathways. Cancer letters. 2015; 356:273–280. [PubMed: 24614287]
38. Shaheen NJ. Editorial: what is behind the remarkable increase in esophageal adenocarcinoma? Am J Gastroenterol. 2014; 109:345–347. [PubMed: 24594951]
39. Brandstedt J, Wangefjord S, Nodin B, Eberhard J, Jirstrom K, Manjer J. Associations of hormone replacement therapy and oral contraceptives with risk of colorectal cancer defined by clinicopathological factors, beta-catenin alterations, expression of cyclin D1, p53, and microsatellite-instability. BMC cancer. 2014; 14:371. [PubMed: 24885829]
40. Coppede F. The role of epigenetics in colorectal cancer. Expert review of gastroenterology & hepatology. 2014:1–14.
41. Bishehsari F, Mahdavinia M, Vacca M, Malekzadeh R, Mariani-Costantini R. Epidemiological transition of colorectal cancer in developing countries: Environmental factors, molecular pathways, and opportunities for prevention. World journal of gastroenterology: WJG. 2014; 20:6055–6072. [PubMed: 24876728]
42. Cross AJ, Moore SC, Boca S, et al. A prospective study of serum metabolites and colorectal cancer risk. Cancer. 2014; 120:3049–3057. [PubMed: 24894841]
43. Simons CC, van den Brandt PA, Stehouwer C, van Engeland M, Weijenberg MP. Body size, physical activity, early life energy restriction, and associations with methylated insulin-like growth factor binding protein genes in colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2014; 23:1852–1862. [PubMed: 24972776]
44. Haque TR, Bradshaw PT, Crockett SD. Risk Factors for Serrated Polyps of the Colorectum. Digestive diseases and sciences. 2014; 59:2874–2889. [PubMed: 25030942]
45. Ryan BM, Wolff RK, Valeri N, et al. An analysis of genetic factors related to risk of inflammatory bowel disease and colon cancer. Cancer epidemiology. 2014; 38:583–590. [PubMed: 25132422]
46. Li P, Wu H, Zhang H, et al. Aspirin use after diagnosis but not prediagnosis improves established colorectal cancer survival: a meta-analysis. Gut. 2015 in press (published online).
47. Huser V, Sincan M, Cimino JJ. Developing genomic knowledge bases and databases to support clinical management: current perspectives. Pharmacogenomics and personalized medicine. 2014; 7:275–283. [PubMed: 25276091]
48. Wennersten C, Andersson G, Boman K, Nodin B, Gaber A, Jirstrom K. Incident urothelial cancer in the Malmo Diet and Cancer Study: cohort characteristics and further validation of ezrin as a prognostic biomarker. Diagnostic pathology. 2014; 9:189. [PubMed: 25278252]
49. Mikeska T, Craig JM. DNA methylation biomarkers: cancer and beyond. Genes. 2014; 5:821–864. [PubMed: 25229548]
50. Campbell PT, Deka A, Briggs P, et al. Establishment of the Cancer Prevention Study II Nutrition Cohort Colorectal Tissue Repository. Cancer Epidemiol Biomarkers Prev. 2014; 23:2694–2702. [PubMed: 25472679]
51. Wild CP, Bucher JR, de Jong BW, et al. Translational cancer research: balancing prevention and treatment to combat cancer globally. J Natl Cancer Inst. 2015; 107:353. [PubMed: 25515230]
52. Caiazza F, Ryan EJ, Doherty G, Winter DC, Sheahan K. Estrogen receptors and their implications in colorecal carcinogenesis. Front Oncol. 2015; 5 Article 19.
53. Ng JM, Yu J. Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer. Int J Mol Sci. 2015; 16:2472–2496. [PubMed: 25622259]
54. Tillmans LS, Vierkant RA, Wang AH, et al. Associations between Environmental Exposures and Incident Colorectal Cancer by ESR2 Protein Expression Level in a Population-Based Cohort of Older Women. Cancer Epidemiol Biomarkers Prev. 2015; 24:713–719. [PubMed: 25650184]
55. Witvliet MI. World health survey: a useful yet underutilized global health data source. Austin J Public Health Epidemiol. 2014; 1:id1012.

56. Potter, S. Kindle edition ed. Emeroe Publishing; 2014. Body mass index 112 Success Secrets – 112 Most Asked Questions On Body mass index – What You Need To Know.
57. Cisyk AL, Penner-Goeke S, Lichtensztejn Z, et al. Characterizing the prevalence of chromosome instability in interval colorectal cancer. Neoplasia. 2015; 17:306–316. [PubMed: 25810015]
58. Weisenberger DJ, Levine AJ, Long TI, et al. Association of the Colorectal CpG Island Methylator Phenotype with Molecular Features, Risk Factors and Family History. Cancer Epidemiol Biomarkers Prev. 2015; 24:512–519. [PubMed: 25587051]
59. Gao C. Molecular pathological epidemiology: an interdisciplinary field for study of hepatocellular carcinoma. Austin J Gastroenterol. 2015; 2:1040.
60. Szylberg L, Janiczek M, Popiel A, Marszalek A. Serrated polyps and their alternative pathway to the colorectal cancer: a systematic review. Gastroenterol Res Pract. 20152015 ID 573814.
61. Sherman ME, Howatt W, Blows FM, Pharoah P, Hewitt SM, Garcia-Closas M. Molecular pathology in epidemiologic studies: a primer on key considerations. Cancer Epidemiol Biomarkers Prev. 2010; 19:966–972. [PubMed: 20332257]
62. Gaudet MM, Sherman ME, Thun MJ. Learning from disease heterogeneity. Lancet Oncol. 2012; 13:862–863. [PubMed: 22863524]
63. Begg CB, Zabor EC. Detecting and Exploiting Etiologic Heterogeneity in Epidemiologic Studies. Am J Epidemiol. 2012; 176:512–518. [PubMed: 22922440]
64. Begg CB, Zabor EC, Bernstein JL, Bernstein L, Press MF, Seshan VE. A conceptual and methodological framework for investigating etiologic heterogeneity. Statistics in medicine. 2013; 32:5039–5052. [PubMed: 23857589]
65. Field AE, Camargo CA, Ogino S. The merits of subtyping obestity: one size does not fit all. JAMA. 2013; 310:2147–2148. [PubMed: 24189835]
66. Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. Gut. 2012; 61:847–854. [PubMed: 22427238]
67. Yamauchi M, Lochhead P, Morikawa T, et al. Colorectal cancer: a tale of two sides or a continuum? Gut. 2012; 61:794–797. [PubMed: 22490520]
68. Lochhead P, Chan AT, Nishihara R, et al. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. Mod Pathol. 2015; 28:14–29. [PubMed: 24925058]
69. Nishi A, Kawachi I, Koenen KC, Wu K, Nishihara R, Ogino S. Lifecourse Epidemiology and Molecular Pathological Epidemiology. Am J Prev Med. 2015; 48:116–119. [PubMed: 25528613]
70. Ogino S, Giovannucci E. Commentary: Lifestyle factors and colorectal cancer microsatellite instability – molecular pathological epidemiology science, based on unique tumour principle. In J Epidemiol. 2012; 41:1072–1074.
71. Ogino S, King EE, Beck AH, Sherman ME, Milner DA, Giovannucci E. Interdisciplinary education to integrate pathology and epidemiology: towards molecular and population-level health science. Am J Epidemiol. 2012; 176:659–667. [PubMed: 22935517]
72. DerSimonian R, Charette LJ, McPeek B, Mosteller F. Reporting on methods in clinical trials. N Engl J Med. 1982; 306:1332–1337. [PubMed: 7070458]
73. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med. 2010; 7:e1000251. [PubMed: 20352064]
74. Samowitz WS, Albertsen H, Sweeney C, et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. J Natl Cancer Inst. 2006; 98:1731–1738. [PubMed: 17148775]
75. Limsui D, Vierkant RA, Tillmans LS, et al. Cigarette Smoking and Colorectal Cancer Risk by Molecularly Defined Subtypes. J Natl Cancer Inst. 2010; 102:1012–1022. [PubMed: 20587792]
76. Nishihara R, Morikawa T, Kuchiba A, et al. A prospective study of duration of smoking cessation and colorectal cancer risk by epigenetics-related tumor classification. Am J Epidemiol. 2013; 178:84–100. [PubMed: 23788674]
77. Curtin K, Samowitz WS, Wolff RK, Herrick J, Caan BJ, Slattery ML. Somatic alterations, metabolizing genes and smoking in rectal cancer. Int J Cancer. 2009; 125:158–164. [PubMed: 19358278]

78. Poynter JN, Haile RW, Siegmund KD, et al. Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. Cancer Epidemiol Biomarkers Prev. 2009; 18:2745–2750. [PubMed: 19755657]
79. Lindor NM, Yang P, Evans I, et al. Alpha-1-antitrypsin deficiency and smoking as risk factors for mismatch repair deficient colorectal cancer: A study from the colon cancer family registry. Molecular genetics and metabolism. 2010; 99:157–159. [PubMed: 19853488]
80. Chia VM, Newcomb PA, Bigler J, Morimoto LM, Thibodeau SN, Potter JD. Risk of microsatellite-unstable colorectal cancer is associated jointly with smoking and nonsteroidal anti-inflammatory drug use. Cancer Res. 2006; 66:6877–6883. [PubMed: 16818666]
81. Barrow TM, Michels KB. Epigenetic epidemiology of cancer. Biochem Biophys Res Commun. 2014; 455:70–83. [PubMed: 25124661]
82. Slattery ML, Curtin K, Anderson K, et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. J Natl Cancer Inst. 2000; 92:1831–1836. [PubMed: 11078760]
83. Campbell PT, Jacobs ET, Ulrich CM, et al. Case-control study of overweight, obesity, and colorectal cancer risk, overall and by tumor microsatellite instability status. J Natl Cancer Inst. 2010; 102:391–400. [PubMed: 20208017]
84. Satia JA, Keku T, Galanko JA, et al. Diet, lifestyle, and genomic instability in the north Carolina colon cancer study. Cancer Epidemiol Biomarkers Prev. 2005; 14:429–436. [PubMed: 15734969]
85. Arain MA, Sawhney M, Sheikh S, et al. CIMP Status of Interval Colon Cancers: Another Piece to the Puzzle. Am J Gastroenterol. 2010; 105:1189–1195. [PubMed: 20010923]
86. Nishihara R, Wu K, Lochhead P, et al. Long-term Colorectal Cancer Incidence and Mortality after Lower Endoscopy. N Engl J Med. 2013; 369:1095–1105. [PubMed: 24047059]
87. Ogino S, Lochhead P, Giovannucci E, Meyerhardt JA, Fuchs CS, Chan AT. Discovery of colorectal cancer PIK3CA mutation as potential predictive biomarker: power and promise of molecular pathological epidemiology. Oncogene. 2014; 33:2949–2955. [PubMed: 23792451]
88. Jung S, Spiegelman D, Baglietto L, et al. Fruit and vegetable intake and risk of breast cancer by hormone receptor status. J Natl Cancer Inst. 2013; 105:219–236. [PubMed: 23349252]
89. Lao VV, Grady WM. Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol. 2011; 8:686–700. [PubMed: 22009203]
90. Colussi D, Brandi G, Bazzoli F, Ricciardiello L. Molecular Pathways Involved in Colorectal Cancer: Implications for Disease Behavior and Prevention. Int J Mol Sci. 2013; 14:16365–16385. [PubMed: 23965959]
91. Bardhan K, Liu K. Epigenetics and colorectal cancer pathogenesis. Cancers. 2013; 5:676–713. [PubMed: 24216997]
92. Zoratto F, Rossi L, Verrico M, et al. Focus on genetic and epigenetic events of colorectal cancer pathogenesis: implications for molecular diagnosis. Tumour Biol. 2014; 35:6195–6206. [PubMed: 25051912]
93. Aleman JO, Eusebi LH, Ricciardiello L, Patidar K, Sanyal AJ, Holt PR. Mechanisms of obesity-induced gastrointestinal neoplasia. Gastroenterology. 2014; 146:357–373. [PubMed: 24315827]
94. Lin JH, Giovannucci E. Environmental exposure and tumor heterogeneity in colorectal cancer risk and outcomes. Curr Colorectal Cancer Rep. 2014; 10:94–104.
95. Song M, Garrett WS, Chan AT. Nutrients, Foods, and Colorectal Cancer Prevention. Gastroenterology. 2015 in press (published online).
96. Jeon JY, Meyerhardt JA. Energy in and energy out: what matters for survivors of colorectal cancer? J Clin Oncol. 2012; 30:7–10. [PubMed: 22124098]
97. Campbell PT, Patel AV, Newton CC, Jacobs EJ, Gapstur SM. Associations of Recreational Physical Activity and Leisure Time Spent Sitting With Colorectal Cancer Survival. J Clin Oncol. 2013; 31:876–885. [PubMed: 23341510]
98. Bathe OF, Farshidfar F. From Genotype to Functional Phenotype: Unraveling the Metabolomic Features of Colorectal Cancer. Genes. 2014; 5:536–560. [PubMed: 25055199]
99. Kuchiba A, Morikawa T, Yamauchi M, et al. Body mass index and risk of colorectal cancer according to fatty acid synthase expression in the nurses' health study. J Natl Cancer Inst. 2012; 104:415–420. [PubMed: 22312135]

100. Ogino S, kawasaki T, Ogawa A, Kirkner GJ, Loda M, Fuchs CS. Fatty acid synthase overexpression in colorectal cancer is associated with microsatellite instability, independent of CpG island methylator phenotype. Human pathology. 2007; 38:842–849. [PubMed: 17350669]
101. Herbert K, Kerr R, Kerr DJ, Church DN. Are NSAIDs coming back to colorectal cancer therapy or not? Curr Colorectal Cancer Rep. 2014; 10:363–371.
102. Tougeron D, Sha D, Manthravadi S, Sinicrope FA. Aspirin and colorectal cancer: Back to the Future. Clin Cancer Res. 2014; 20:1087–1094. [PubMed: 24327271]
103. Chan AT, Ogino S, Fuchs CS. Aspirin and the Risk of Colorectal Cancer in Relation to the Expression of COX-2. N Engl J Med. 2007; 356:2131–2142. [PubMed: 17522398]
104. Chan AT, Ogino S, Fuchs CS. Aspirin use and survival after diagnosis of colorectal cancer. JAMA. 2009; 302:649–658. [PubMed: 19671906]
105. Chan AT, Ogino S, Giovannucci EL, Fuchs CS. Inflammatory markers are associated with risk of colorectal cancer and chemopreventive response to anti-inflammatory drugs. Gastroenterology. 2011; 140:799–808. quiz e711. [PubMed: 21115010]
106. Liao X, Lochhead P, Nishihara R, et al. Aspirin use, tumor PIK3CA mutation status, and colorectal cancer survival. N Engl J Med. 2012; 367:1596–1606. [PubMed: 23094721]
107. Nishihara R, Lochhead P, Kuchiba A, et al. Aspirin use and risk of colorectal cancer according to BRAF mutation status. JAMA. 2013; 309:2563–2571. [PubMed: 23800934]
108. Nan H, Morikawa T, Suuriniemi M, et al. Aspirin use, 8q24 single nucleotide polymorphism rs6983267, and colorectal cancer according to CTNNB1 alterations. J Natl Cancer Inst. 2013; 105:1852–1861. [PubMed: 24317174]
109. Fink SP, Yamauchi M, Nishihara R, et al. Aspirin and the Risk of Colorectal Cancer in Relation to the Expression of 15-Hydroxyprostaglandin Dehydrogenase (HPGD). Sci Transl Med. 2014; 6:233re232.
110. Domingo E, Church DN, Sieber O, et al. Evaluation of PIK3CA mutation as a predictor of benefit from NSAID therapy in colorectal cancer. J Clin Oncol. 2013; 31:4297–4305. [PubMed: 24062397]
111. Garcia-Closas M, Couch FJ, Lindstrom S, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet. 2013; 45:392–398. 398e391–392. [PubMed: 23535733]
112. Tamimi RM, Colditz GA, Hazra A, et al. Traditional breast cancer risk factors in relation to molecular subtypes of breast cancer. Breast cancer research and treatment. 2012; 131:159–167. [PubMed: 21830014]
113. Millikan RC, Newman B, Tse CK, et al. Epidemiology of basal-like breast cancer. Breast cancer research and treatment. 2008; 109:123–139. [PubMed: 17578664]
114. Phipps AI, Buist DS, Malone KE, et al. Family history of breast cancer in first-degree relatives and triple-negative breast cancer risk. Breast cancer research and treatment. 2011; 126:671–678. [PubMed: 20814817]
115. Yang XR, Sherman ME, Rimm DL, et al. Differences in risk factors for breast cancer molecular subtypes in a population-based study. Cancer Epidemiol Biomarkers Prev. 2007; 16:439–443. [PubMed: 17372238]
116. Kostic AD, Ojesina AI, Pedamallu CS, et al. PathSeq: software to identify or discover microbes by deep sequencing of human tissue. Nature biotechnology. 2011; 29:393–396.
117. Kostic AD, Gevers D, Pedamallu CS, et al. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. Genome research. 2012; 22:292–298. [PubMed: 22009990]
118. Tahara T, Yamamoto E, Suzuki H, et al. Fusobacterium in colonic flora and molecular features of colorectal carcinoma. Cancer Res. 2014
119. Mima K, Sukawa Y, Nishihara R, et al. Fusobacterium nucleatum and T-cells in colorectal carcinoma. JAMA Oncol. 2015 (in press).
120. Kostic AD, Chun E, Robertson L, et al. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell host & microbe. 2013; 14:207–215. [PubMed: 23954159]

121. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014; 513:202–209. [PubMed: 25079317]
122. Morton LM, Sampson JN, Cerhan JR, et al. Rationale and Design of the International Lymphoma Epidemiology Consortium (InterLymph) Non-Hodgkin Lymphoma Subtypes Project. Journal of the National Cancer Institute. Monographs. 2014; 2014:1–14. [PubMed: 25174022]
123. Begg CB. A strategy for distinguishing optimal cancer subtypes. Int J Cancer. 2011; 129:931–937. [PubMed: 20949563]
124. Wang M, Kuchiba A, Ogino S. A meta-regression method for studying etiologic heterogeneity across disease subtypes classified by multiple biomarkers. Am J Epidemiol. 2015 in press.
125. Chatterjee N, Sinha S, Diver WR, Feigelson HS. Analysis of cohort studies with multivariate and partially observed disease classification data. Biometrika. 2010; 97:683–698. [PubMed: 22822252]
126. Chatterjee N. A Two-Stage Regression Model for Epidemiological Studies With Multivariate Disease Classification Data. Journal of the American Statistical Association. 2004; 99:127–138.
127. Rosner B, Glynn RJ, Tamimi RM, et al. Breast cancer risk prediction with heterogeneous risk profiles according to breast cancer tumor markers. Am J Epidemiol. 2013; 178:296–308. [PubMed: 23645624]
128. Leek JT, Scharpf RB, Bravo HC, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010; 11:733–739. [PubMed: 20838408]
129. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome biology. 2014; 15:R31. [PubMed: 24495553]
130. Houseman EA, Kim S, Kelsey KT, Wiencke JK. DNA methylation in whole blood: uses and challenges. Curr Envir Health Rpt. 2015 in press.
131. Beck AH, Knoblauch NW, Hefti MM, et al. Significance analysis of prognostic signatures. PLoS Comput Biol. 2013; 9:e1002875. [PubMed: 23365551]
132. Sherman ME, Figueroa JD, Henry JE, Clare SE, Rufenbarger C, Storniolo AM. The Susan G. Komen for the Cure Tissue Bank at the IU Simon Cancer Center: a unique resource for defining the “molecular histology” of the breast. Cancer Prev Res (Phila). 2012; 5:528–535. [PubMed: 22345117]
133. Figueroa JD, Pfeiffer RM, Patel DA, et al. Terminal duct lobular unit involution of the normal breast: implications for breast cancer etiology. J Natl Cancer Inst. 2014; 106
134. Faupel-Badger JM, Arcaro KF, Balkam JJ, et al. Postpartum remodeling, lactation, and breast cancer risk: summary of a National Cancer Institute-sponsored workshop. J Natl Cancer Inst. 2013; 105:166–174. [PubMed: 23264680]
135. Ghosh K, Vachon CM, Pankratz VS, et al. Independent association of lobular involution and mammographic breast density with breast cancer risk. J Natl Cancer Inst. 2010; 102:1716–1723. [PubMed: 21037116]
136. Ghosh K, Hartmann LC, Reynolds C, et al. Association between mammographic density and age-related lobular involution of the breast. J Clin Oncol. 2010; 28:2207–2212. [PubMed: 20351335]
137. Lochhead P, Chan AT, Giovannucci E, et al. Progress and opportunities in molecular pathological epidemiology of colorectal premalignant lesions. Am J Gastroenterol. 2014; 109:1205–1214. [PubMed: 24935274]
138. Roman-Perez E, Casbas-Hernandez P, Pirone JR, et al. Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. Breast Cancer Res. 2012; 14:R51. [PubMed: 22429463]
139. Sun X, Sandhu R, Figueroa JD, Gierach GL, Sherman ME, Troester MA. Benign breast tissue composition in breast cancer patients: association with risk factors, clinical variables, and gene expression. Cancer Epidemiol Biomarkers Prev. 2014; 23:2810–2818. [PubMed: 25249325]
140. Palmer JR, Viscidi E, Troester MA, et al. Parity, lactation, and breast cancer subtypes in African American women: results from the AMBER Consortium. J Natl Cancer Inst. 2014; 106:dju237. [PubMed: 25224496]
141. Hakimi AA, Furberg H, Zabor EC, et al. An epidemiologic and genomic investigation into the obesity paradox in renal cell carcinoma. J Natl Cancer Inst. 2013; 105:1862–1870. [PubMed: 24285872]

142. Markt SC, Valdimarsdottir UA, Shui IM, et al. Circadian clock genes and risk of fatal prostate cancer. Cancer Causes Control. 2015; 26:25–33. [PubMed: 25388799]

143. Pettersson A, Lis RT, Meisner A, et al. Modification of the association between obesity and lethal prostate cancer by TMPRSS2:ERG. J Natl Cancer Inst. 2013; 105:1881–1890. [PubMed: 24292212]

144. Gates MA, Rosner BA, Hecht JL, Tworoger SS. Risk factors for epithelial ovarian cancer by histologic subtype. Am J Epidemiol. 2010; 171:45–53. [PubMed: 19910378]

145. Poole EM, Merritt MA, Jordan SJ, et al. Hormonal and reproductive risk factors for epithelial ovarian cancer by tumor aggressiveness. Cancer Epidemiol Biomarkers Prev. 2013; 22:429–437. [PubMed: 23307531]

146. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007; 4:e296. [PubMed: 17941714]

147. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007; 4:e297. [PubMed: 17941715]

148. Terry MB, Knight JA. STROBE-ME – Illuminating methodological issues for the reporting of molecular epidemiology data. Preventive medicine. 2011; 53:388–389. [PubMed: 22024220]

149. Gallo V, Egger M, McCormack V, et al. STrengthening the Reporting of OBservational studies in Epidemiology – Molecular Epidemiology (STROBE-ME): An Extension of the STROBE Statement. PLoS Med. 2011; 8:e1001117. [PubMed: 22039356]

150. Ogino S, Nishihara R, VanderWeele TJ, et al. Molecular pathological epidemiology is essential in studying neoplastic and non-neoplastic diseases in the era of precision medicine. Epidemiology. 2015 (in press).