Published in final edited form as:

Am J Obstet Gynecol. 2013 May ; 208(5): 360–365. doi:10.1016/j.ajog.2012.11.030.

Phenotyping Clinical Disorders: Lessons Learned From Pelvic Organ Prolapse

Jennifer M. Wu, MD, MPH¹, Renée M. Ward, MD², Kristina L. Allen-Brady, PhD³, Todd L. Edwards, PhD^{5,7}, Peggy A. Norton, MD⁴, Katherine E. Hartmann, MD, PhD^{2,5}, Elizabeth R. Hauser, PhD⁶, and Digna R. Velez Edwards, PhD^{5,7}

¹Department of Obstetrics and Gynecology, Division of Urogynecology, Duke University, Durham, NC

²Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, TN

³Division of Genetic Epidemiology, Department of Internal Medicine, University of Utah, Salt Lake City, UT

⁴Department of Obstetrics and Gynecology, University of Utah, Salt Lake City, UT

⁵Vanderbilt Epidemiology Center, Institute of Medicine and Public Health, Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, Tennessee

⁶Center for Human Genetics, Duke University, Durham, NC

⁷Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee

Abstract

Genetic epidemiology, the study of genetic contributions to risk for disease, is an innovative area in medicine. While research in this arena has advanced in other disciplines, few genetic epidemiologic studies have been conducted in obstetrics and gynecology. It is crucial that we study the genetic susceptibility for issues in women's health, as this information will shape the new frontier of "personalized medicine." To date, preterm birth may be one of the best examples of genetic susceptibility in obstetrics and gynecology, but many areas are being evaluated including endometriosis, fibroids, polycystic ovarian syndrome and pelvic floor disorders. An essential component to genetic epidemiologic studies is to characterize, or "phenotype," the disorder in order to identify genetic effects. Given the growing importance of genomics and genetic epidemiology, we discuss the importance of accurate phenotyping of clinical disorders and highlight critical considerations and opportunities in phenotyping, using pelvic organ prolapse as a clinical example.

Genomics, the study of genes or gene products in a person or organism, and genetic epidemiology, the study of the genetic contributions to health and disease in families and populations, are emerging fields in medicine. Although these areas of research are quite advanced in other disciplines,¹ such as cardiology² and oncology,³ genetic epidemiology

© 2012 Mosby, Inc. All rights reserved.

Corresponding Author: Jennifer M. Wu, MD, MPH, Duke University Medical Center Box 3192, 5324 McFarland Dr., Suite 310, Durham, NC 27707, Work: (919) 401-1006, Cell: (919) 452-4577, Fax: (919) 401-1033, jennifer.wu@duke.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: The authors report no conflict of interest.

studies in obstetrics and gynecology are relatively limited. Several disorders are being evaluated including preterm birth,⁴⁻⁷ fibroids,⁸ endometriosis,⁹ polycystic ovarian syndrome,^{10, 11} and pelvic floor disorders, especially pelvic organ prolapse.¹²⁻¹⁷ In this discussion, we will use prolapse as an example to illustrate important considerations in characterizing, or phenotyping, clinical disorders.

In Mendelian or monogenic traits, there is a simple relationship between genotype and phenotype, such that a mutation at a single locus results in an expressed trait/disease (phenotype) that is inherited according to Mendel's laws. Examples include the characteristics of Mendel's pea plants, ABO blood groups, cystic fibrosis, and sickle cell anemia. However, the etiology of the vast majority of clinical disorders is more complex. Instead of simple causal relationships between a single genotype and phenotype, the etiologies of these disorders are multifactorial, resulting from the effects of multiple genes, individual history, behaviors, concomitant medical conditions, and exposures including medications, diet and environmental agents, as well as the interactions among these factors.

Advances in high-throughput genotyping and sequencing capabilities launched a paradigm shift from studying Mendelian disorders to studying *common, complex diseases*, in which there is a polygenic model of genetic risk, with multiple gene variants conferring moderate risk in a general population ("common disease, common variant" [CDCV] hypothesis).¹⁸ While the genome-wide association study (GWAS) design, a consequence of CDCV, has led to many discoveries for heritable traits, limitations to this study design include the challenge of identifying causal, functional variants after finding a significant association and a limited evaluation of single nucleotide polymorphisms with a prevalence of less than 5%.^{19, 20} Thus, a more recent paradigm is that multiple rare variants result in common diseases ("common disease, rare variant" [CDRV] hypothesis).²¹ Despite these advances in genetic epidemiology, the genetic determinants of most obstetric and gynecologic diseases remain uninvestigated.

Why identify genetic risk factors for disease?

One major reason to identify genetic risk factors for a clinical disease is to advance scientific knowledge about the pathophysiology of the disease in order to facilitate targeted diagnostic, therapeutic and preventative interventions. For example, pelvic organ prolapse is thought to have a multifactorial etiology with a genetic component;^{12, 13, 22, 23} however, the underlying mechanisms that lead to defects are not yet known. By understanding the genetic susceptibility for a disease, mutation screening may be developed and utilized for risk-stratification to individualize recommendations for prevention or treatment.

In studies of common, complex diseases, a critical component is to characterize, or "phenotype," the disorder accurately to facilitate the identification of genetic effects. The process of developing a phenotype definition involves the use of epidemiologic, biological, molecular and/or bioinformatic methods to systematically select characteristics of a disorder that might result from distinct genetic influences. Defining the phenotype is critical because poorly-phenotyped studies will misclassify participants, which decreases the power to identify significant genetic effects. While it is important to carefully define and phenotype "cases" with a disease, trait, or susceptibility, it is also necessary to define "controls" without disease as genetic epidemiologic studies will often compare cases to controls. To illustrate these concepts in further detail and to discuss the nuances of designing genetic epidemiologic studies, we will use pelvic organ prolapse as an example to highlight critical considerations in phenotyping clinical disorders.

Pelvic Organ Prolapse as an Example

Pelvic organ prolapse is a major women's health issue with a prevalence as high as 40% in postmenopausal women.²⁴ Risk factors for prolapse include advancing age, Caucasian race, higher body mass index (BMI), increasing parity, smoking, chronic constipation, chronic cough and family history of prolapse.^{22, 23, 25–28} Although childbirth is often considered a major risk factor, a majority of parous women do not develop prolapse while some nulliparous women do develop the condition. Furthermore, the development of prolapse is often temporally distant from birth injury, thereby strongly suggesting that additional factors beyond vaginal parity, such as genetic susceptibility, that contribute to or protect against prolapse.

General Principles in Genetic Epidemiologic Studies

In developing a study to evaluate the genetic epidemiology of prolapse, a common approach is to recruit cases with prolapse versus controls without prolapse; however, a critical scientific question is how will pelvic organ prolapse be defined. Prior to discussing different approaches to defining prolapse, it is important to discuss key principles that are unique to genetic epidemiologic studies.

Diseases that become more prevalent with advancing age

It is important to consider the age of onset and natural history of the disease of interest as this will directly impact who to recruit for cases and controls. For example, a 45 year old woman may not show any evidence of prolapse, and thus could be considered a control subject. However, it is possible that she may not develop prolapse until 50 or 60 years of age. In many genetic epidemiologic schemes, these younger “unaffected” subjects are considered “unknowns.” Thus, older women without evidence of prolapse are more ideal control subjects. The same rationale applies to studying other phenotypes, which are characterized by older age at onset, such as coronary artery disease and breast cancer.

Important considerations regarding race and ethnicity

It is critical to carefully consider race and ethnicity in genetic epidemiologic studies, as different racial/ethnic groups may have different frequencies of genetic variants, which is a concept known as population stratification. These systematic differences in allele frequencies exist because of differences in ancestry rather than the association between genes and diseases.^{29, 30} Confounding by race results when race is associated with both differences in genetic variant frequencies as well as differences in disease prevalence, but there is no association between the genetic variants and disease. Thus, population stratification in genetic studies can lead to excess false positive results and failures to detect true associations.³¹ Because of population stratification, genetic studies will often focus on one specific race/ethnicity or utilize advanced analytic techniques to adjust for racial genetic differences.^{32–34}

Consequences of heterogeneity in phenotype definitions

Heterogeneities in complex disease, such as prolapse, arise due to different clinical manifestations, differences in biological pathways leading to disease, and differences in symptom severity. Often phenotypic heterogeneity (inconsistency in the definition of phenotype) across studies makes it difficult to generalize study findings and to replicate genetic associations. This inconsistency in trait definition may be due to lack of consensus or a general disagreement among researchers in the field as to the correct phenotypic definition. Establishing a consistent operational definition and description of the phenotype to be used across studies may help avoid this dilemma. Below, we discuss different methods

of phenotyping prolapse and make a recommendation regarding how to phenotype prolapse in order to move towards developing a consistent phenotype definition for future genetic epidemiologic studies.

Defining the Phenotype of Pelvic Organ Prolapse

There are several different clinical definitions for prolapse, and none are universally accepted. Possible definitions include 1) bothersome symptoms, which focuses on subjective factors, 2) surgical treatment for prolapse, and 3) stage of prolapse, which is an objective measure of severity.

It has been shown that bothersome symptoms are a critical component of evaluating prolapse³⁵ and validated questionnaires^{36, 37} include a question regarding bothersome prolapse symptoms: “Do you usually have a bulge of something falling out that you can see or feel in the vaginal area? If yes, how much does it bother you?”³⁸ While this is key to evaluating treatment outcomes, it is less suited to designing early genetic association studies.

Another option for phenotyping prolapse is to define prolapse based on those who have chosen to undergo surgery. Presumably only women with severe enough bothersome symptoms warrant surgical management.¹² One important consideration is that it is necessary to confirm the procedures performed and the indication for surgery. Obtaining prior operative reports can be somewhat challenging, especially if the surgery was in the distant past. Furthermore, this phenotype definition fails to capture women who may have very advanced prolapse but pursue non-surgical treatment options.

A more objective method of phenotyping is based on examination findings, and for prolapse, the two primary examination techniques are the pelvic organ prolapse quantification (POP-Q)³⁹ or Baden-Walker system.⁴⁰ A common strategy is to compare cases with POP-Q stage III/IV prolapse versus controls with stage 0/I prolapse.^{14, 41} The rationale is that there may be some degree of prolapse that occurs as a “normal” part of aging, as the proportion of postmenopausal women with a mild degree of prolapse is relatively high.²⁴ By excluding stage II, the more extreme phenotypes are captured. Selective sampling of more extreme phenotypes is a common approach in genetic epidemiologic studies, as recruitment of more extreme phenotypes increases the probability of observing a genetic effect. For example, a woman with a high genetic risk might be a young, premenopausal, nulliparous woman with stage IV prolapse and a family history of prolapse. In contrast, a low genetic risk woman might be an 80 year old woman who has had 5 vaginal deliveries yet has stage 0 prolapse. While it may be difficult to identify and recruit these ideal extreme phenotypes, it is important to select cases and controls which are clinically disparate. Thus far, a majority of case control association studies have defined cases and controls based on physical examination findings.^{14, 42, 43, 44} Currently, we recommend using physical examination findings to phenotype prolapse, given that it is an objective measure that is likely to be more reproducible. Also, for initial genetic epidemiologic studies, the definition of cases as stage III to IV and controls as stage 0 to I may increase the power to identify a genetic effect. Although we have focused on the definition of cases for this phenotype, it is also critically important to define control subjects carefully, in order to decrease the risk of misclassification bias. It is necessary to ensure that controls have been evaluated fully for the disease/phenotype of interest and have had equal opportunity for detection of the condition. For example, if we defined prolapse solely based on symptoms, it is possible that an asymptomatic woman who actually has prolapse on examination could be selected as a control. It is important to exclude those who have had prior surgery for prolapse; thus, a careful medical and surgical history and physical examination are needed.

Role of Risk Factors and Confounders

After defining the cases and controls, it is necessary to identify risk factors and potential confounders. The presence of risk factors, in either cases or controls, will be adjusted for in the statistical analysis in order to ultimately estimate the contribution from genetic factors. Family history may be used to further refine the phenotype, if desired.

Risk Factors for Prolapse

Given that the etiology of prolapse is likely multifactorial with a genetic component as well as environmental factors, which include health behaviors, medical conditions, and exposures, it is important to collect data on other risk factors and to think about potential exclusion criteria. The presence of risk factors, in either cases or controls, will be adjusted for in the statistical analysis in order to ultimately estimate the contribution from genetic factors. Data collection should include detailed information regarding obstetrics history, including parity; mode of deliveries (vaginal, cesarean and assisted vaginal deliveries), age at first delivery, infant birth weight, and perineal trauma; diseases associated with Valsalva maneuvers, such as emphysema/COPD or constipation; prior surgical history, including hysterectomy; prior anti-incontinence and prolapse procedures; medications, especially hormone replacement therapy and steroid use; smoking history; and body mass index.²² Another consideration is to collect additional markers for connective tissue issues such as joint hypermobility, varicose veins, rectal prolapse, and hernias.⁴⁵ However, those with connective tissue disease, such as Ehlers-Danlos or Marfans syndrome, should be excluded as these subjects may have an inherently higher risk for prolapse.

Family History as a Risk Factor

Previous studies examining the risk of pelvic organ prolapse in nulliparous and parous sister pairs have observed a concordance rate of prolapse suggesting a familial predisposition.²³ Studies involving extended pedigrees using a population database have shown similar results with significantly increased relative risks in first and third degree relatives ($RR=4.15$, $p<0.001$, $RR=1.24$, $p=0.05$, respectively) and increased average pairwise relatedness for all individuals with prolapse ($p<0.001$).⁴⁶ Hence, having close and distant relatives with pelvic organ prolapse are known risk factors for the disease. Study of extended pedigrees with a significant excess of prolapse has multiple benefits including (1) affected relatives are more likely to share the same genetic risk factors than unrelated cases with prolapse, (2) a significant excess number of prolapse cases in a family results in a high probability that the cases are attributable to genetic effect(s), and (3) knowledge of pedigree structure increases power of the study as it can lead to a predictable correlation structure of the genetic risk. Selective sampling of prolapse cases based on family history of prolapse increases the likelihood that genetic effects will be detected in a study.

Coexisting Pelvic Floor Disorders

Prolapse often coexists with other pelvic floor disorders, such as stress or urgency urinary incontinence. Currently, we do not know if stress incontinence and prolapse share the same causal pathway or if they are distinct genetic conditions. The incorporation of coexisting pelvic floor disorders becomes an important issue when defining cases and controls because if stress incontinence and prolapse share the same causal pathway, there is the potential to introduce misclassification bias if controls have stress incontinence even if they do not have prolapse. Thus, one recommendation is to collect additional data to define the phenotypes of stress and urge incontinence when conducting a genetic study of prolapse until we have more data regarding the etiologies of these conditions.

Resources for Genetic Epidemiology Studies

After carefully phenotyping the disease of interest and identifying risk factors, the next step is to determine what clinical and genetic data resources are available. The two broad categories are those that (1) exist already or (2) need to be collected prospectively. There are a number of existing resources to investigate (Table 1). A valuable source is the National Center for Biotechnology Information's (NCBI) database of Genotypes and Phenotypes (dbGaP) which is a public repository that archives and distributes the results of studies that have investigated numerous phenotypes with a wide variety of genomic technologies (<http://www.ncbi.nlm.nih.gov/gap>).⁴⁷ An investigator can search dbGaP to browse what genetic studies have been conducted, what variables were included in these studies, and what type of analysis was performed. These datasets from dbGaP can be downloaded at no expense to an investigator and are particularly useful if additional phenotype data are available from an existing genome-wide association study (GWAS) dataset on dbGaP but was not analyzed as part of the primary analyses. For example, if a GWAS was performed to primarily evaluate diabetes but other variables of interest were assessed, such as urinary incontinence, an investigator interested in incontinence can request access to the GWAS data and then evaluate the genetic epidemiology of incontinence.

Another useful resource is DNA banks or biorepositories that are linked to electronic medical records (EMR). One example is the BioVU DNA Repository at Vanderbilt University, which is designed to support large-scale genomic research.^{48, 49} De-identified DNA samples are linked to electronic medical records (EMRs) of patients from Vanderbilt University Medical Center clinics, and detailed human subjects protection is applied to BioVU.⁵⁰ Sample accrual started in 2007 and is ongoing with an average of 600 new samples per week. As of June 2012, there are over 128,480 adult DNA samples.

BioVU is part of the eMERGE Network (electronic MEDical Records and GENomics) which is an NHGRI-funded consortium of DNA biobanks linked to electronic medical records (EMR) with the goal of performing large-scale genomic analyses.⁵¹ To-date eMERGE (<http://www.gwas.org>) has performed genome-wide association studies of peripheral arterial disease,⁵² red blood cell traits,⁵³ HDL cholesterol,⁵⁴ and electrocardiographic PR interval.⁵⁵ Another potential resource is the UK Biobank (<http://www.ukbiobank.ac.uk/>).⁵⁶

Another option is to establish a biorepository by collecting samples prospectively. This approach ensures the collection of detailed, uniform clinical information as well as phenotype data; however, it is critically important to have the necessary infrastructure to establish a biorepository, and the interdisciplinary collaborations needed to conduct genetic epidemiologic studies. When designing a new genetic epidemiologic study, a useful resource is the PhenX toolkit (www.phenx.org/).⁵⁷ PhenX provides 295 well-established, high-quality measures within 21 different domains that can be used by the scientific community to facilitate cross-study comparisons and analyses.⁵⁷ Although specific measures for a disease of interest may not be listed (i.e. there is no information regarding pelvic organ prolapse or urinary incontinence), suggestions on how to collect data on possible covariates (i.e. what is the ideal question to use to assess tobacco use?) will help to ensure that these data could be useful for studies of different traits.

Lastly, another unique aspect of genetic epidemiologic studies is that after conducting an initial study, it is important to validate your findings in an independent cohort, often called a validation cohort. This is in contrast to randomized controlled clinical trials in which the results do not need to be validated initially to be considered a high quality study. Thus, collaboration is a critical component of conducting genetic epidemiologic studies, as collaborators may be able to provide a validation, or replication, cohort. The strong

collaborative nature of this field is evidenced by the numerous consortia that have developed to study different diseases and traits.

Conclusion

In this article, we have outlined important considerations in genetic epidemiologic studies and key issues to consider when phenotyping a clinical disorder. Phenotyping is a critical step in advancing our understanding of the genetic etiology of disorders within obstetrics and gynecology. There are clearly unique considerations in how to operationally define, or phenotype, cases and controls in genetic epidemiologic studies in order to optimize the ability to detect genetic risk factors. Hopefully, in the future there will be dedicated and funded efforts to help uncover the genetic susceptibilities to the multitude of women's health issues.

Acknowledgments

Dr. Wu is supported by K23HD068404, Eunice Kennedy Shriver National Institute of Child Health & Human Development

Dr. Allen-Brady is supported by R01HD061821 (Co-PIs: Lisa Cannon-Albright and Peggy Norton), Eunice Kennedy Shriver National Institute of Child Health & Human Development

Dr. Edwards is supported by KL2RR024977-04, the Vanderbilt Clinical and Translational Research Scholars Program

Dr. Elizabeth Hauser is supported by P30AG028716-06 (PI: Dr. Harvey Cohen), National Institute of Aging

Dr. Velez Edwards is supported by K12HD4383, Building Interdisciplinary Research Careers in Women's Health career development program

REFERENCES

1. GIANT: Genetic Investigation of ANthropometric Traits. Accessed June 6, 2012 at http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium.
2. Kathiresan S, Voight BF, Purcell S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* 2009; 41:334–341. [PubMed: 19198609]
3. Tomlinson IP, Dunlop M, Campbell H, et al. COGENT (COlorectal cancer GENEtics): an international consortium to study the role of polymorphic variation on the risk of colorectal cancer. *Br J Cancer.* 2010; 102:447–454. [PubMed: 19920828]
4. Romero R, Friel LA, Velez Edwards DR, et al. A genetic association study of maternal and fetal candidate genes that predispose to preterm prelabor rupture of membranes (PROM). *Am J Obstet Gynecol.* 2010; 203:361 e1–361 e30. [PubMed: 20673868]
5. Romero R, Velez Edwards DR, Kusanovic JP, et al. Identification of fetal and maternal single nucleotide polymorphisms in candidate genes that predispose to spontaneous preterm labor with intact membranes. *Am J Obstet Gynecol.* 2010; 202:431 e1–431 e34. [PubMed: 20452482]
6. Uzun A, Sharma S, Padbury J. A bioinformatics approach to preterm birth. *Am J Reprod Immunol.* 2012; 67:273–277. [PubMed: 22385126]
7. Dolan SM, Hollegaard MV, Meriäldi M, et al. Synopsis of preterm birth genetic association studies: the preterm birth genetics knowledge base (PTBGene). *Public Health Genomics.* 2010; 13:514–523. [PubMed: 20484876]
8. Cha PC, Takahashi A, Hosono N, et al. A genome-wide association study identifies three loci associated with susceptibility to uterine fibroids. *Nat Genet.* 2011; 43:447–450. [PubMed: 21460842]
9. Painter JN, Anderson CA, Nyholt DR, et al. Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. *Nat Genet.* 2011; 43:51–54. [PubMed: 21151130]

10. Kim JJ, Choi YM, Cho YM, et al. Polycystic ovary syndrome is not associated with polymorphisms of the TCF7L2, CDKAL1, HHEX, KCNJ11, FTO and SLC30A8 genes. *Clin Endocrinol (Oxf)*. 2012
11. Welt CK, Styrkarsdottir U, Ehrmann DA, et al. Variants in DENND1A Are Associated with Polycystic Ovary Syndrome in Women of European Ancestry. *J Clin Endocrinol Metab*. 2012; 97:E1342–E1347. [PubMed: 22547425]
12. Allen-Brady K, Cannon-Albright L, Farnham JM, et al. Identification of six loci associated with pelvic organ prolapse using genome-wide association analysis. *Obstet Gynecol*. 2011; 118:1345–1353. [PubMed: 22105264]
13. Allen-Brady K, Norton PA, Farnham JM, Teerlink C, Cannon-Albright LA. Significant linkage evidence for a predisposition gene for pelvic floor disorders on chromosome 9q21. *Am J Hum Genet*. 2009; 84:678–682. [PubMed: 19393595]
14. Wu JM, Visco AG, Grass EA, et al. Comprehensive analysis of LAMC1 genetic variants in advanced pelvic organ prolapse. *Am J Obstet Gynecol*. 2012; 206:447 e1–447 e6. [PubMed: 22342894]
15. Wu JM, Visco AG, Grass EA, et al. Matrix metalloproteinase-9 genetic polymorphisms and the risk of advanced pelvic organ prolapse. *Obstet Gynecol*. 2012 Jul. *in press*.
16. Nikolova G, Lee H, Berkovitz S, et al. Sequence variant in the laminin gamma1 (LAMC1) gene associated with familial pelvic organ prolapse. *Hum Genet*. 2007; 120:847–856. [PubMed: 17021862]
17. Chen HY, Lin WY, Chen YH, Chen WC, Tsai FJ, Tsai CH. Matrix metalloproteinase-9 polymorphism and risk of pelvic organ prolapse in Taiwanese women. *Eur J Obstet Gynecol Reprod Biol*. 2010; 149:222–224. [PubMed: 20144500]
18. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996; 273:1516–1517. [PubMed: 8801636]
19. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461:747–753. [PubMed: 19812666]
20. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med*. 2010; 363:166–176. [PubMed: 20647212]
21. Tennessen JA, Bigham AW, O'Connor TD, et al. Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science*. 2012
22. Jelovsek JE, Maher C, Barber MD. Pelvic organ prolapse. *Lancet*. 2007; 369:1027–1038. [PubMed: 17382829]
23. Buchsbaum GM, Duecy EE, Kerr LA, Huang LS, Perevich M, Guzick DS. Pelvic organ prolapse in nulliparous women and their parous sisters. *Obstet Gynecol*. 2006; 108:1388–1393. [PubMed: 17138771]
24. Hendrix SL, Clark A, Nygaard I, Aragaki A, Barnabei V, McTiernan A. Pelvic organ prolapse in the Women's Health Initiative: gravity and gravidity. *Am J Obstet Gynecol*. 2002; 186:1160–1166. [PubMed: 12066091]
25. Rortveit G, Brown JS, Thom DH, Van Den Eeden SK, Creasman JM, Subak LL. Symptomatic pelvic organ prolapse: prevalence and risk factors in a population-based, racially diverse cohort. *Obstet Gynecol*. 2007; 109:1396–1403. [PubMed: 17540813]
26. Chiaffarino F, Chatenoud L, Dindelli M, et al. Reproductive factors, family history, occupation and risk of urogenital prolapse. *Eur J Obstet Gynecol Reprod Biol*. 1999; 82:63–67. [PubMed: 10192487]
27. Lukacz ES, Lawrence JM, Contreras R, Nager CW, Lubner KM. Parity, mode of delivery, and pelvic floor disorders. *Obstet Gynecol*. 2006; 107:1253–1260. [PubMed: 16738149]
28. Mant J, Painter R, Vessey M. Epidemiology of genital prolapse: observations from the Oxford Family Planning Association Study. *Br J Obstet Gynaecol*. 1997; 104:579–585. [PubMed: 9166201]
29. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet*. 2003; 361:598–604. [PubMed: 12598158]
30. Freedman ML, Reich D, Penney KL, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet*. 2004; 36:388–393. [PubMed: 15052270]

31. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet.* 2004; 36:512–517. [PubMed: 15052271]
32. Bryc K, Auton A, Nelson MR, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci U S A.* 2010; 107:786–791. [PubMed: 20080753]
33. Bryc K, Velez C, Karafet T, et al. Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc Natl Acad Sci U S A.* 2010; 107(Suppl 2):8954–8961. [PubMed: 20445096]
34. Edwards TL, Gao X. Methods for detecting and correcting for population stratification. *Curr Protoc Hum Genet.* 2012; Chapter 1(Unit 1):1–14. 22.
35. Nygaard I, Barber MD, Burgio KL, et al. Prevalence of symptomatic pelvic floor disorders in US women. *JAMA.* 2008; 300:1311–1316. [PubMed: 18799443]
36. Barber MD. Questionnaires for women with pelvic floor disorders. *Int Urogynecol J Pelvic Floor Dysfunct.* 2007; 18:461–465. [PubMed: 17120170]
37. Barber MD, Walters MD, Bump RC. Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). *Am J Obstet Gynecol.* 2005; 193:103–113. [PubMed: 16021067]
38. Barber MD, Neubauer NL, Klein-Olarte V. Can we screen for pelvic organ prolapse without a physical examination in epidemiologic studies? *Am J Obstet Gynecol.* 2006; 195:942–948. [PubMed: 16681989]
39. Bump RC, Mattiasson A, Bo K, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *Am J Obstet Gynecol.* 1996; 175:10–17. [PubMed: 8694033]
40. ACOG. Practice Bulletin: Pelvic organ prolapse. *Obstet Gynecol.* 2007; 110:717–729. [PubMed: 17766624]
41. Chen C, Hill LD, Schubert CM, Strauss JF 3rd, Matthews CA. Is laminin gamma-1 a candidate gene for advanced pelvic organ prolapse? *Am J Obstet Gynecol.* 2010; 202:505 e1–505 e5. [PubMed: 20223449]
42. Wu JM, Visco AG, Grass EA, et al. Matrix metalloproteinase-9 genetic polymorphisms and the risk for advanced pelvic organ prolapse. *Obstet Gynecol.* 2012; 120:587–593. [PubMed: 22914468]
43. Feiner B, Fares F, Azam N, Auslender R, David M, Abramov Y. Does COL1A1 SP1- binding site polymorphism predispose women to pelvic organ prolapse? *Int Urogynecol J Pelvic Floor Dysfunct.* 2009; 20:1061–1065. [PubMed: 19655227]
44. Rodrigues AM, Girao MJ, da Silva ID, Sartori MG, Martins KD, Castro RD. COL1A1 Sp1- binding site polymorphism as a risk factor for genital prolapse. *Int Urogynecol J Pelvic Floor Dysfunct.* 2008
45. Lammers K, Lince SL, Spath MA, et al. Pelvic organ prolapse and collagen-associated disorders. *Int Urogynecol J.* 2012; 23:313–319. [PubMed: 21811768]
46. Norton PA, Allen-Brady K, Cannon-Albright L. The familiarity of pelvic organ prolapse in the Utah Population Database. *International Urogyn J Pelvic Floor Dysfunct.* 2012 Jun. *in press.*
47. Mailman MD, Feolo M, Jin Y, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet.* 2007; 39:1181–1186. [PubMed: 17898773]
48. Ritchie MD, Denny JC, Crawford DC, et al. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. *Am J Hum Genet.* 2010; 86:560–572. [PubMed: 20362271]
49. Roden DM, Pulley JM, Basford MA, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther.* 2008; 84:362–369. [PubMed: 18500243]
50. Pulley J, Clayton E, Bernard GR, Roden DM, Masys DR. Principles of human subjects protections applied in an opt-out, de-identified biobank. *Clin Transl Sci.* 2010; 3:42–48. [PubMed: 20443953]
51. McCarty CA, Chisholm RL, Chute CG, et al. The eMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med Genomics.* 2011; 4:13. [PubMed: 21269473]

52. Kullo IJ, Fan J, Pathak J, Savova GK, Ali Z, Chute CG. Leveraging informatics for genetic studies: use of the electronic medical record to enable a genome-wide association study of peripheral arterial disease. *J Am Med Inform Assoc.* 2010; 17:568–574. [PubMed: 20819866]
53. Kullo IJ, Ding K, Jouni H, Smith CY, Chute CG. A genome-wide association study of red blood cell traits using the electronic medical record. *PLoS One.* 2010; 5
54. Turner S, Armstrong LL, Bradford Y, et al. Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet.* 2011; Chapter 1(Unit1 19)
55. Denny JC, Ritchie MD, Crawford DC, et al. Identification of genomic predictors of atrioventricular conduction: using electronic medical records as a tool for genome science. *Circulation.* 2010; 122:2016–2021. [PubMed: 21041692]
56. Collins R. What makes UK Biobank special? *Lancet.* 2012; 379:1173–1174. [PubMed: 22463865]
57. PhenX: Consensus measures for Phenotypes and eXposures. Accessed June 6, 2012 at <https://www.phenx.org/>.

Table 1

Resources for Genetic Epidemiologic Studies

Categories	Sources
Existing Resources	<ul style="list-style-type: none"> • dbGAP⁴⁷ (http://www.ncbi.nlm.nih.gov/gap) • BioVU⁴⁸ • eMERGE⁵¹ (http://www.gwas.org) • UK Biobank⁵⁶ (http://www.ukbiobank.ac.uk/) • Institutional resources / biorepositories
Collecting Data and Samples	<ul style="list-style-type: none"> • PhenX⁵⁷ (https://www.phenx.org/)