INTRODUCTION TO GENETIC EPIDEMIOLOGY

(EPID0754 + GBIO0015)

Prof. Dr. Dr. K. Van Steen

CHAPTER 1: SETTING THE PACE

1 Course Responsible

Contact details

2 Administrative Issues

Course details and examination methods

3 Exploring the Scene

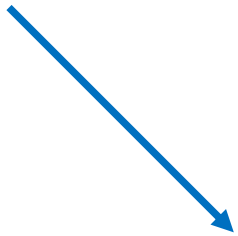
Expectations

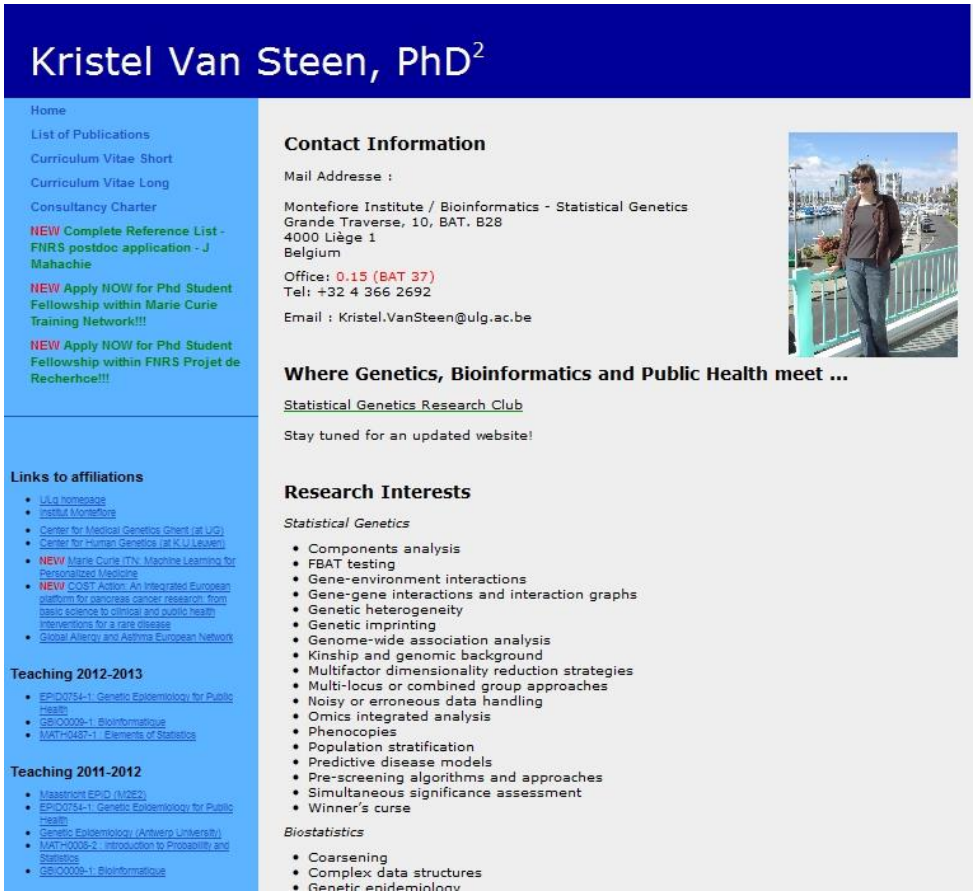
4 Background Information: Medical Genomics

Recent evolutions in medical genomics

5 Workshop Papers

1 Course Responsible





Kristel Van Steen, PhD[2]

Home
List of Publications
Curriculum Vitae Short
Curriculum Vitae Long
Consultancy Charter
NEW Complete Reference List - FNRS postdoc application - J Mahachie
NEW Apply NOW for Phd Student Fellowship within Marie Curie Training Network!!!
NEW Apply NOW for Phd Student Fellowship within FNRS Projet de Recherhce!!!

Links to affiliations
- ULg homepage
- Institut Montefiore
- Center for Medical Genetics Ghent (at UG)
- Center for Human Genetics (at K.U.Leuven)
- NEW Marie Curie ITN: Machine Learning for Personalized Medicine
- NEW COST Action: An integrated European platform for pancreas cancer research: from basic science to clinical and public health interventions for a rare disease
- Global Allergy and Asthma European Network

Teaching 2012-2013
- EPID0754-1: Genetic Epidemiology for Public Health
- GBIO0009-1: Bioinformatique
- MATH0487-1 : Elements of Statistics

Teaching 2011-2012
- Maastricht EPID (M2E2)
- EPID0754-1: Genetic Epidemiology for Public Health
- Genetic Epidemiology (Antwerp University)
- MATH0008-2 : Introduction to Probability and Statistics
- GBIO0009-1: Bioinformatique

Contact Information

Mail Addresse :

Montefiore Institute / Bioinformatics - Statistical Genetics
Grande Traverse, 10, BAT. B28
4000 Liège 1
Belgium

Office: 0.15 (BAT 37)
Tel: +32 4 366 2692

Email : Kristel.VanSteen@ulg.ac.be

Where Genetics, Bioinformatics and Public Health meet ...

Statistical Genetics Research Club

Stay tuned for an updated website!

Research Interests

Statistical Genetics
- Components analysis
- FBAT testing
- Gene-environment interactions
- Gene-gene interactions and interaction graphs
- Genetic heterogeneity
- Genetic imprinting
- Genome-wide association analysis
- Kinship and genomic background
- Multifactor dimensionality reduction strategies
- Multi-locus or combined group approaches
- Noisy or erroneous data handling
- Omics integrated analysis
- Phenocopies
- Population stratification
- Predictive disease models
- Pre-screening algorithms and approaches
- Simultaneous significance assessment
- Winner's curse

Biostatistics
- Coarsening
- Complex data structures
- Genetic epidemiology

Contact details via

www.montefiore.ulg.ac.be/~kvansteen

Questions or remarks via e-mail

kristel.vansteen@ulg.ac.be or the TA kbessonov@student.ulg.ac.be

Use “genetic epidemiology” in subject title when sending a mail to ask questions or to make a face-to-face appointment for a meeting

2 Administrative Issues

Old learning outcomes are revised ~last year's organization

Acquis d'apprentissage (objectifs d'apprentissage) du cours	Learning outcomes of the course
Familiariser les étudiants avec les différentes facettes de l'épidémiologie génétique, avec une emphase à la génétique statistique.	To familiarise students with the different facets of genetic epidemiology, with an emphasis on statistical genetics.
Prérequis et corequis / Modules de cours optionnels recommandés	**Prerequisites and co-requisites/ Recommended optional programme components**
Connaissance de base de concepts de statistiques. Les termes génétiques seront répresentés.	Basic knowledge of statistical concepts. Genetic terms will be re-introduced.
Activités d'apprentissage prévues et méthodes d'enseignement	**Planned learning activities and teaching methods**
Travail personnel par l'intermédiaire des tâches de lecture, homeworks et plus grands projets, à l'aide principalement des outils de Bioconductor dans le logiciel libre R. Les informations seront disponibles sur la page www.montefiore.ulg.ac.be/~kvansteen/	Personal work through tasks set during lectures, homework and larger projects, with the help primarily of Bioconductor tools using the 'R' open source software. Further information is available at www.montefiore.ulg.ac.be/~kvansteen/
Mode d'enseignement (présentiel ; enseignement à distance)	**Mode of delivery (face-to-face ; distance-learning)**
2eme semestre. Le cours sera interactif dans anglais/français. Slides en anglais. Devoirs peut être en français. Les examens ont organisés en français. Toutes les modalités pratiques du cours seront disponibles sur la page www.montefiore.ulg.ac.be/~kvansteen/	2nd semester. The course will be interactive in English/French. All course notes are in English. Homeworks can be handed in in French. Examination will be in French. All practical details about the course will be available at www.montefiore.ulg.ac.be/~kvansteen/
Lectures recommandées ou obligatoires et notes de cours	**Recommended or required readings**
Slides en anglais (www.montefiore.ulg.ac.be/~kvansteen/)	English-language slides via www.montefiore.ulg.ac.be/~kvansteen/

Public health (PH) and non-public health (non-PH) combined

- Difference in number of hours / work load (20-20 / 15-15)
- Difference in type of homeworks:
  - PH:
    - papers around topics or themes are presented and discussed in class
    - two presentation rounds (presentation and report are marked)
    - exam is oral: organized around “list of thematic questions”
  - non-PH:
    - practical analysis of real-life data using PLINK
    - presentation of results (twice; presentations are marked)
    - exam: final report, with possibility to defend “orally” (“list of thematic questions”)

Course website

Teaching 2013-2014 - UNDER CONSTRUCTION

EPID0754-1 : Introduction à l'épidémiologie génétique
GBIO0015-1 : A tour in genetic epidemiology

SOME PRACTICAL ARRANGEMENTS:

- When does the first class take place?
  - Wednesday February 26, from 1-4pm
  - If you have a laptop, please bring it to this and subsequent classes
  - No other material is necessary, unless announced
- When and where do the next classes take place?
  - Click here for an updated course schedule. All classes take place from 1-5pm, unless stated otherwise.
    Updated ()
- What about homeworks?
  - Click here for some guidelines
- What if I have additional questions?
  - Consult Kristel Van Steen (kristel.vansteen@ulg.ac.be) before or after class or by e-mail. In order for the mail not to get lost, please mention "genetic epidemiology" in the subject title. Teaching assistent for this course: Kirill Bessonov (kbessonov@student.ulg.ac.be).
- Reference books:
  - Click here

Course website

INFORMATION BY CLASS:

26 February 2014 - CLASS 1 : 1pm-4pm

- Course notes:
  - Chapter 1: Logistics
  - Chapter 2: Introduction to Genetics
  - Downloading instructions R and Bioconductor: Follow this link
  - Information on using R:
    - Part 1: R manuals
    - Part 2: Basic Commands !!!
    - Part 3: An introduction to R
    - Part 4: Advanced
    - Part 5: A few illustrative examples in R
- Workshop papers: Some motivations for genetic epidemiology
  - Paper 1: Prospects of genetic epidemiology in the 21st century
  - Paper 2: Recent advances in the genetic epidemiology and molecular genetics of substance use disorders
- Background Information:
  - Basic epidemiology

Course organization

	HW1	HW2	Particip. in discussions	Oral Exam (*)	Total
Max	**20 (presentation + slides/ report)**	**30 (presentation + slides/report)**	**10**	**40 (no final report)**	**100**

	HW1	HW2	Final Report (optional oral defense *)
Max	**20 (presentation)**	**20 (presentation)**	**60 (final report)**

(*) “themes” are posted on the course website

Course outline

CHAPTER 2: INTRODUCTION TO GENETICS

1 Basics of molecular genetics

2 Overview of human genetics

CHAPTER 3: DIFFERENT FACES OF GENETIC EPIDEMIOLOGY

1 Basic epidemiology

2 Genetic epidemiology

3 Phenotypic aggregation within families

4 Genetic epidemiology and public health

CHAPTER 4: GENOME-WIDE ASSOCIATION STUDIES

1 Introduction

2 Preliminary analyses

3 Tests of association: single SNP

4 Tests of association: multiple SNPs

5 Dealing with population stratification

6 Multiple testing

7 Assessing the function of genetic variants

8 Validation and replication

CHAPTER 5: INCORPORATING THE ENVIRONMENT

1 Beyond main effects: GxG interactions

2 Multifactor Dimensionality Reduction techniques

3 Challenges: GxE interactions, omics integration

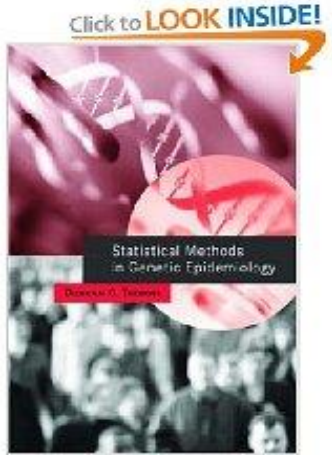
CHAPTER 6: INCORPORATING ADDITIONAL LEVELS OF CELLULAR COMPLEXITY

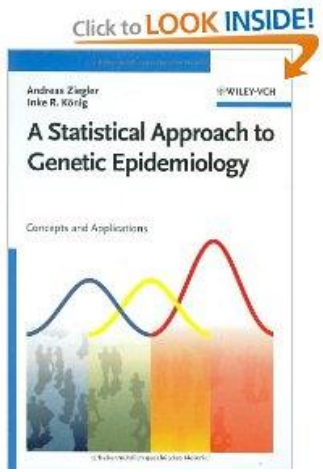
1 Biological interpretation

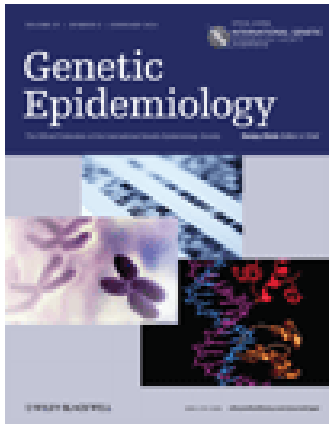
2 Integrated analyses

Course material / References

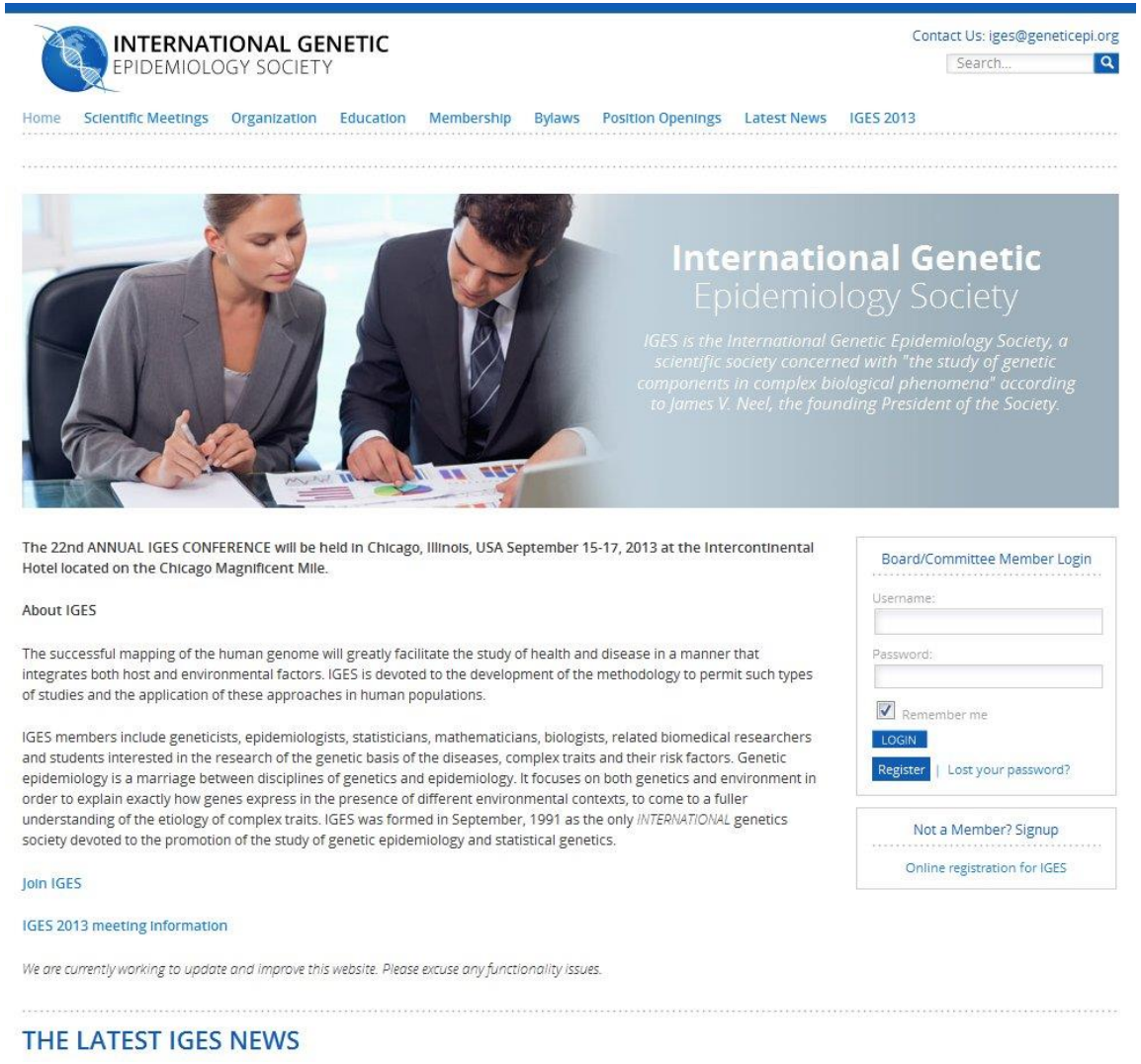
- Check out course website for slides and assignments
- These slides are comprehensive enough for the subset of material that will be covered in class
- For those who are interested, key references are included on the course website and below







IGES (http://www.geneticepi.org/)



3 Exploring the Scene

Round-Table Discussion

Q1: What is your background? What is your thesis about? What do you want to achieve in your professional life?

Q2: Have you analyzed data before? How? Which tools have you used? What was the most difficult part? Data manipulation? Interpretation? Implementation?

Q3: What is epidemiology?

Q4: What do you think genetic epidemiology includes? Personalized medicine? Genetic testing?

Q5: What are your expectations of this course? What would you really like to do / achieve in this course?

4 Background Information: Medical Genomics

Genomics Primer from Mayo Clinic Proceedings

Mayo staff have authored several articles that will be useful for anyone developing a thorough understanding of medical genomics. These articles have been published as a series in the Mayo Clinic Proceedings journal. The following sections have been copied from their website:

http://mayoresearch.mayo.edu/mayo/research/grc/proceedings.cfm

Part I: History of Genetics and Sequencing of the Human Genome

Cindy Pham Lorentz, MS; Eric D. Wieben, PhD; Ayalew Tefferi, MD; David A. H. Whiteman, MD; and Gordon W. DeWald, PhD

The first part of this overview gives an account of the history of genetics that spans from humankind's first attempts at understanding and influencing heredity, to the early scientific work in the field of genetics, and then to the advancements in modern genetics. The second part summarizes the Human Genome Project (HGP) from inception to the publishing of the "first draft" of the human genome sequence.

Part II: Background Principles and Methods in Molecular Genetics

Ayalew Tefferi, MD; Eric D. Wieben, PhD; Gordon W. DeWald, PhD; David A. H. Whiteman, MD; Matthew E. Bernard, MD; and Thomas C. Spelsberg, PhD

In this second part of an educational series in medical genomics, selected principles and methods in molecular biology are recapped, with the intent to prepare the reader for forthcoming articles with a more direct focus on aspects of the subject matter

Part III: Microarray Experiments and Data Analysis

Ayalew Tefferi, MD; Mark E. Bolander, MD; Stephen M. Ansell, MD, PhD; Eric D. Wieben, PhD; and Thomas C. Spelsberg, PhD

Genomics has been defined as the comprehensive study of whole sets of genes, gen products, and their interactions as opposed to the study of single genes or proteins. Microarray technology is one of many novel tools that are allowing global and high-throughput analysis of genes and gene products. In addition to an introduction on underlying principles, the current review focuses on the use of microarrays in gene expression analysis. … The current review should serve as an introduction to the subject for clinician investigators, physicians and medical scientists in training, practicing clinicians, and other students of medicine.

Part IV: Expression Proteomics

Animesh Pardanani, MD, PhD; Eric D. Wieben, MD; Thomas C. Spelsberg, PhD; and Ayalew Tefferi, MD

Proteomics, simply defined is the study of proteomes. The three broad areas are expression proteomics, which catalogues the relative abundance of proteins; cell-mapping or cellular proteomics, which delineates functional protein-protein interactions and organelle-specific protein distribution; and structural proteomics, which characterizes the 3-dimensional structure of proteins. This articles reviews the area of expression proteomics.

Part V: Bioinformatics

Peter L Elkin, MD

Bioinformatics is the discipline that develops and applies informatics to the field of molecular biology. Although a comprehensive review of the entire field of bioinformatics is beyond the scope of this article, I review the basic tenets of the field and provide a topical sampling of the popular technologies available to clinicians and researchers. These technologies include tools and methods for sequence analysis (nucleotide and protein sequences), rendering of secondary and tertiary structures for these molecules, and protein fold prediction that can lead to rational drug design. I then discuss signaling pathways, new standards for data representation of genes and proteins, and finally the promise of merging these molecular data with the clinical world (the new science of phenomics).

Part VI: Genomics and Molecular Genetics in Clinical Practice

Stephen M. Ansell, MD, PhD; Michael J. Ackerman, MD, PhD; John L. Black, MD; Lewis R. Roberts, MD, PhD; and Ayalew Tefferi, MD

An important milestone in medical science is the recent completion of a "working draft" of the human genome sequence. The identification of all human genes and their regulatory regions provides the framework to expedite our understanding of the molecular basis of disease. This advance has also formed the foundation for a broad range of genomic tools that can be applied to medical science. These developments in global gene and gene product analysis as well as targeted molecular genetic testing are destined to change the practice of modern medicine.

…

Despite these exciting advances, many practicing clinicians perceive that the role of molecular genetics, especially that of genomics, is confined primarily to the research arena with little current clinical applicability. The aim of the article is to highlight advances in DNA/RNA-based methods of susceptibility screening, disease diagnosis and prognostication, and prediction of treatment outcome in regard to both drug toxicity and response as they apply to various areas of clinical medicine.

Part VII: The Evolving Concept of the Gene

Eric D. Wieben, PhD

The draft sequence of the human genome was reported 2 years ago, and the task of filling gaps and polishing the sequence is nearing completion. However, despite this remarkable achievement, there is still no definitie assessment of the number of genes contained in the human genome. In part, this uncertainty reflects our growing understanding of the complexity and diversity of gene structure. Examples of complex gene structure are considered in the context of the discussion about the evolution of our understanding of gene structure and function.

Part VIII: Essentials of Medical Genetics for the Practicing Physician

Regina E. Ensenauer, MD; Shanda S. Reinke; Michael J. Ackerman, MD, PhD; David J. Tester; David A. H. Whiteman, MD; and Ayalew Tefferi, MD

After the mapping and sequencing of the human genome, medical professionals from essentially all specialties turned their attention to investigating the role genes play in health and disease. Until recently, medical genetics was considered a specialty of minor practical relevance. This view has changed with the development of new diagnostic and therapeutic possibilities. It is now realized that genetic disease represents an important part of medical practice. Achievements in cancer genetics, in the field of prenatal diagnostics (including carrier testing for common recessive disorders), and in newborn screening for treatable metabolic disorders reinforce the rapidly expanding role of genetics in medicine.

...

Diagnosing a genetic disorder not only allows for disease-specific management options but also has implications for the affected individual's entire family. A working understanding of the underlying concepts of genetic disease with regard to chromosome, single gene, mitochondrial, and multifactorial disorders is necessary for today's practicing physician. Routine clinical practice in virtually all medical specialties will soon require integration of these fundamental concepts for use in accurate diagnosis and ensuring appropriate referrals for patients with genetic disease and their families.

Part IX: Scientific and Clinical Applications of DNA Microarrays -- Multiple Myeloma as a Disease Model

John Shaughnessy, Jr., PhD

Multiple myeloma (MM) is a poorly understood and uniformly fatal malignancy of antibody-secreting plasma cells. …

This review discusses progress made in the development of molecular-based diagnostics and prognostics for MM through the dissection of the transcriptome of plasma cells from healthy individuals and patients with MM and other plasma cell dyscrasias.

Part X: Gene Therapy

Stephen J. Russell, MD, PhD; and Kah-Whye Peng, PhD

Gene therapy is defined as any therapeutic procedure in which genes are intentionally introduced into human somatic cells. Both preclinical and clinical gene therapy research have been progressing rapidly during the past 15 years; gene therapy is now a highly promising new modality for the treatment of numerous human disorders. Since the first clinical test of gene therapy in 1989, more than 600 gene therapy protocols have been approved, and more than 3000 patients have received gene therapy. However, at the time of writing this article, no gene therapy products have been approved for clinical use.

…

This article explains the potential clinical scope of gene therapy and the underlying pharmacological principles, describes some of the major gene transfer systems (or vectors) that are used to deliver genes to their target sites, and discusses the various strategies for controlling expression of therapeutic transgenes. Safety issues regarding clinical use of gene therapy are explored, and the most important technical challenges facing this field of research are highlighted. This review should serve as an introduction to the subject of gene therapy for clinician investigators, physicians and medical scientists in training, practicing clinicians, and other students of medicine.

4 Background Information (continued)

- On the course website, "**background information**" is for your information only. It is up to you to what extent you consult this information
- This is NOT EXAM MATERIAL

- In contrast, "**workshop papers**" do provide complementary information to the course slides and may help in better understanding certain concepts or may be used as reference material for your presentations
- The idea of workshop papers is to see concepts applied to relevant (other) contexts and to grow awareness about pros and cons of certain strategies
- Therefore, this material is useful when preparing the EXAM (cfr. "guiding questions")

5 Workshop papers

Workshop paper 1

European Journal of Epidemiology **18**: 607–616, 2003.

REVIEW

Prospects of genetic epidemiology in the 21st century

Marieke C.J. Dekker & Cornelia M. van Duijn
Department of Epidemiology and Biostatistics, Erasmus MC, Rotterdam, The Netherlands

Accepted in revised form 14 April 2003

Abstract. Genetic epidemiology is a young but rapidly developing discipline. Although its early years were largely dedicated to family-based research in monogenic disorders, now genetic–epidemiologic research increasingly focuses on complex, multifactorial disorders. Along with the development of the human-genome map and advances in molecular technology grows the importance of genetic–epidemiologic applications. Large-scale population-based studies, requiring close integration of genetic and epidemiologic research, determine future research in the field. In this paper, we review the basic principles underlying genetic–epidemiologic research, such as molecular genetics and familial aggregation of disease, as well as the typical study approaches of genome screening and candidate-gene studies.

Key words: Familial aggregation, Genetics, Genetic epidemiology, Polymorphisms, Study design

Workshop paper 2

Published in final edited form as:
Nat Neurosci. ; 15(2): 181–189. doi:10.1038/nn.3018.

Recent advances in the genetic epidemiology and molecular genetics of substance use disorders

Kenneth S Kendler[1,2,3], **Xiangning Chen**[1,2,3], **Danielle Dick**[1,2], **Hermine Maes**[1,2,3], **Nathan Gillespie**[1,2], **Michael C Neale**[1,2,3], and **Brien Riley**[1,2,3]

[1]Virginia Institute of Psychiatric and Behavioral Genetics, Medical College of Virginia and Virginia Commonwealth University, Richmond, Virginia, USA.

[2]Department of Psychiatry, Medical College of Virginia and Virginia Commonwealth University, Richmond, Virginia, USA.

[3]Department of Human and Molecular Genetics, Medical College of Virginia and Virginia Commonwealth University, Richmond, Virginia, USA.

Abstract

This article reviews current advances in the genetics of substance use disorders (SUDs). Both genetic and environmental sources of risk are required to develop a complete picture of SUD etiology. Genetic sources of risk for SUDs are not highly substance specific in their effects. Genetic and environmental risks for SUDs typically do not only add together but also interact with each other over development. Risk gene identification for SUDs has been difficult, with one recent success in identifying nicotinic receptor variants that affect risk for nicotine dependence. The impact of genetic variants on SUD risk will individually be small. Although genetic epidemiologic methods are giving us an increasingly accurate map of broad causal pathways to SUDs, gene discovery will be needed to identify the specific biological systems. Identifying these risk genes and understanding their action will require large clinical samples, and interaction between these studies and work in model organisms.

NIH-PA Author Manuscript