2003

Genetic Analysis Workshop 13: Analysis of Longitudinal Family Data for Complex Diseases and Related Risk Factors

Laura Almasy
Southwest Foundation for Biomedical Research

Christopher I. Amos
University of Texas Health Science Center at Houston

Joan E. Bailey-Wilson
National Institutes of Health

Rita M. Cantor
University of California - Los Angeles

Cashell E. Jaquish
National Institutes of Health

See next page for additional authors

Authors
Laura Almasy, Christopher I. Amos, Joan E. Bailey-Wilson, Rita M. Cantor, Cashell E. Jaquish, Maria Martinez, Rosalind J. Neuman, Jane M. Olson, Lyle J. Palmer, Stephen S. Rich, M Anne Spence, and Jean W. MacCluer

BMC Genetics

BioMed Central

Introduction

Genetic Analysis Workshop 13: Analysis of Longitudinal Family Data for Complex Diseases and Related Risk Factors

Laura Almasy*[1], Christopher I Amos[2], Joan E Bailey-Wilson[3], Rita M Cantor[4], Cashell E Jaquish[5], Maria Martinez[6], Rosalind J Neuman[7], Jane M Olson[8], Lyle J Palmer[9], Stephen S Rich[10], M Anne Spence[11] and Jean W MacCluer[1]

Address: [1]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas, 78245, USA, [2]Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas, 77025, USA, [3]National Human Genome Research Institute, National Institutes of Health, Baltimore, Maryland, 21224, USA, [4]Department of Human Genetics, University of California Los Angeles School of Medicine, Los Angeles, California, 90095, USA, [5]National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland, 20892, USA, [6]INSERM, Hôpital Saint Louis, Evry Cedex, 91034, France, [7]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, 63110, USA, [8]Department of Biostatistics and Epidemiology, Case Western Reserve University, Cleveland, Ohio, 44109, USA, [9]Genetic Epidemiology Group, Western Australia Institute for Medical Research, Perth, WA 6000, Australia, [10]Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina, 27517, USA and [11]Department of Pediatrics, University of California Irvine Medical Center, Orange, California, 92868, USA

Email: Laura Almasy* - almasy@darwin.sfbr.org; Christopher I Amos - camos@request.mdacc.tmc.edu; Joan E Bailey-Wilson - jebw@nhgri.nih.gov; Rita M Cantor - rcantor@mednet.ucla.edu; Cashell E Jaquish - jaquishc@nhlbi.nih.gov; Maria Martinez - maria@evry.inserm.fr; Rosalind J Neuman - roz@gretta.wustl.edu; Jane M Olson - olson@darwin.epbi.cwru.edu; Lyle J Palmer - lpalmer@cyllene.uwa.edu.au; Stephen S Rich - srich@wfubmc.edu; M Anne Spence - maspence@uci.edu; Jean W MacCluer - jean@darwin.sfbr.org

* Corresponding author

from Genetic Analysis Workshop 13: Analysis of Longitudinal Family Data for Complex Diseases and Related Risk Factors
New Orleans Marriott Hotel, New Orleans, LA, USA, November 11–14, 2002

Published: 31 December 2003

BMC Genetics 2003, **4**(Suppl 1):S1

This article is available from: http://www.biomedcentral.com/1471-2156/4/s1/S1

Preface

This supplement to BMC Genetics contains the proceedings of Genetic Analysis Workshop 13 (GAW13), which was held November 11–14, 2002, at the Marriott hotel in New Orleans, Louisiana. The Genetic Analysis Workshops, which began in 1982 and are currently held biannually, provide a venue for the development and testing of statistical genetic methods. For each GAW, one or more data sets are made available to potential participants and GAW attendees must submit an analysis of one of these data sets, or be the providers of a data set or otherwise involved in workshop organization. The workshop itself is then a head-to-head comparison of different ways of analyzing the same data. New analytical methods are introduced and the performance of established methods is compared.

GAW13 focused on analytical issues relating to localization of genes influencing common, complex diseases and their risk factors, with an emphasis on use of longitudinal data. For the first time in a GAW, longitudinal family data, covering a 40-year time period, was made available to workshop participants. As has been the case in recent GAWs, both simulated and actual human data were available. New this year, however, was a close correspondence between these data sets, with the simulated data mirroring the family structure, trait data, and longitudinal sampling scheme of the real data set.

The Framingham Heart Study data provided for GAW13 included longitudinal observations from two cohorts. The original Framingham cohort (Cohort 1), was first examined in 1948 and has been examined every 2 years thereafter. Cohort 2, composed primarily of offspring of the original cohort and the spouses of these offspring, was examined first in 1971 and has been examined approximately every 4 years. The data provided for GAW13 are a subset of these two cohorts and include 2885 individuals

in 300 pedigrees. Both diagnostic phenotypes, such as hypertension, and quantitative risk factors, such as HDL cholesterol levels, weight, and systolic blood pressure were available. Genotypes were provided for microsatellite markers representing a whole genome scan at an average 10-cM density. Marker maps and estimated marker allele frequencies were also supplied.

A second data set was simulated, based as closely as possible on the Framingham Heart Study data made available to GAW13. While the exact details of the pedigrees were modified, the overall pedigree size and structure in the simulated data was based on that of the Framingham Heart Study data. Similarly, phenotypes were simulated to have the same distribution as those in the real data and the observed correlations between some phenotypes were incorporated into the trait model. Genotype data were also closely matched to the actual data, particularly in regard to marker distribution and information content. While participants had the option of using the complete simulated data set, missing data patterns were also generated based on observations from the Framingham Heart Study data. Those who analyzed the simulated data had the choice of obtaining the simulation model prior to their analyses and were asked to indicate whether their analyses were done with or without this knowledge.

In spring of 2002, the availability of the GAW13 data was announced by e-mail to the over 1700 individuals on the GAW mailing list. A total of 97 groups requested GAW13 data. The Framingham data were requested by 75 groups and the simulated data by 90 groups, with 67 groups requesting both data sets. In the summer of 2002, 117 contributions were received describing analyses of the two data sets. Of these, 81 utilized the Framingham Heart Study data and 36 analyzed the simulated data. A book, or CD, containing these contributions plus papers describing the data sets was distributed to workshop participants.

A total of 241 individuals from 13 countries attended GAW13. Attendees included investigators from five continents – Asia, Australia, Europe, and North and South America. As with GAW12, participants were organized into 11 presentation groups and a co-author with past GAW experience was asked to lead discussion in each presentation group. Groups ranged in size from 6 to 14 papers with common themes. Some groups collected papers exploring similar methodological issues, such as methods for longitudinal analysis or derivation of new phenotypes from the data, whereas others were related through a common focus on particular phenotypes, such as analyses of blood pressure and hypertension phenotypes or analyses of tobacco and alcohol use. Given the similarities in the Framingham Heart Study and simulated data sets, presentation groups were assigned without regard to the data set analyzed. Each group met individually during the workshop and in many groups members communicated beforehand to begin comparing and contrasting the approaches taken and the results obtained by group members. At GAW13, many groups used part of their group meeting time for individual presentations by each group member, giving investigators an opportunity to present their work. Although mainly attended by group members, group meetings were open to all GAW13 participants. From this process, each group developed an oral presentation, summarizing and synthesizing the work of the individual papers, which was delivered to the overall workshop during general sessions. Individual contributions were also presented in the form of 52 posters displayed during four poster sessions.

The manuscripts included here are a subset of those presented at GAW13. All of these papers have been reviewed for scientific merit. The proceedings begin with two papers describing the two data sets, followed by 103 individual GAW13 contributions organized by presentation group and alphabetically by first author within each group. Because the data set analyzed was not considered in assignment of group membership, papers reporting analyses of simulated data and Framingham Heart Study data are intermingled. New to GAW this year, the proceedings occupy two volumes. In addition to the individual contributions in this volume, each presentation group has a summary in a forthcoming supplement to the journal *Genetic Epidemiology* in which the present manuscripts are compared and contrasted and larger themes and conclusions are explored. Results of GAW13 analyses provided novel insights into the etiology of cardiovascular disease and some of its risk factors. In addition, many new methods were developed, explored, and applied for the analysis of genetic data from longitudinal studies, an area of research that has been underdeveloped to date.

Acknowledgments

Many people contributed to the success of Genetic Analysis Workshop 13 by selecting workshop topics, providing real and simulated data, preparing and distributing data and participant contributions, communicating with participants, organizing the meeting, chairing sessions, reviewing manuscripts, and editing the GAW13 proceedings.

The Genetic Analysis Workshops would not exist without the generosity of the investigators who provide data for analysis by workshop participants. We are grateful to the investigators from the Framingham Heart Study who provided data to GAW13: Adrienne Cupples, Qiong Yang, Serkalem Demissie, Donna Copenhafer, Richard Myers, Ralph D'Agostino, Philip Wolf, and Daniel Levy, for the Framingham Heart Study Investigators. The Framingham Heart Study is supported by the NHLBI Framingham Heart Study Contract N01-HC-25195.

The GAW13 simulated data set was generated in a collaborative effort among Warwick Daw, Duncan Thomas, John Morrison, Xiaojun Zhou, Carol Bosken, and Dakai Zu, with input from Laura Almasy, Chris Amos,

Adrienne Cupples, Lynn Goldin, Jean MacCluer, and John Rice. We are grateful to all of these people for their efforts in the difficult task of designing and creating a simulated data set that mimicked many of the features of the Framingham data, thus offering participants the opportunity to address issues such as power and false positives. The creation of the simulated data set was supported by start-up funds to Warwick Daw, and by grants CA52862 and GM58897 to John Morrison and Duncan Thomas.

At GAW13, contributions were organized into groups, each focused on a single topic. Group leaders had the difficult task of generating discussion among strangers via e-mail, and organizing presentations that summarized all of the contributions in their group. Their efforts deserve special recognition. We are grateful to the following individuals who led the group presentations: Max Baur, Jim Gauderman, Laura Almasy, Gail Jarvik, Beth Hauser, Deborah Meyers, Catherine Falk, Ellen Wijsman, Heike Bickeböller, Lynn Goldin, and Nancy Saccone. Ample time was scheduled for discussion, with four discussion periods led by Robert Elston, Ingrid Borecki, Anne Spence, and Chris Amos. We thank these discussion leaders for their efforts in stimulating lively interactions among participants.

Many scientific reviewers provided useful comments and criticisms of the papers in this volume: Goncalo Abecasis, Alexandre Alcais, Mariza De Andrade, Allison Ashley-Koch, Larry Atwood, Julia Bailey, Michael Barmada, Jill Barnholtz, Terri Beaty, Timothy Bishop, Michael Boehnke, Shelley Bull, Gary Chase, Anthony Comuzzie, Heather Cordell, Nancy Cox, Adrienne Cupples, Stefan Czerwinski, Anita DeStefano, Ravi Duggirala, Robert Elston, Carol Etzel, Cathy Fann, Stephen Finch, Tatiana Foroud, Varghese George, Saurabh Ghosh, Katrina Goddard, Alisa Goldstein, Harald Göring, Xiuqing Guo, Sandra Hasstedt, Lorena Havill, Tony Hinrichs, Peter Holmans, John Hopper, Stefan Horvath, Wen-Chi Hsueh, Jian Huang, Sudha Iyengar, Gail Jarvik, Terri King, Alison Klein, Peter Kraft, Ethan Lange, Carl Langefeld, Daniel Levy, Michael Mahaney, James Malley, Teri Manolio, Eden Martin, Beverly Mellen, Nancy Mendell, Chantal Merette, Braxton Mitchell, Kari North, Jürg Ott, Shane Pankratz, Andrew Paterson, Elizabeth Pugh, Marylyn Ritchie, Susan Santangelo, J.H. Schick, Katrina Scurrah, Paul Sorlie, Michael Stern, Brian Suarez, Duncan Thomas, Hemant Tiwari, Kai Wang, Daniel Weeks, Bruce Weir, Marsha Wilcox, Jeff Williams, Alexander Wilson, and Heping Zhang. We are grateful for their contributions.

Vanessa Olmo has had major responsibility for all aspects of workshop organization since GAW7, in 1991. She continues to have primary responsibility for workshop logistics, including interaction with participants, organizers, editors, and publisher; data distribution; local organization; maintenance of the GAW web site and mailing list; and preparation of the proceedings. The Genetic Analysis Workshops would not succeed without her dedication and hard work. We also thank Tom Dyer, Richard Polich, Kent Polk, and Gerry Vest, who helped with data distribution; Patricia Curry, who assisted with communications with participants; René Sandoval and Rudy Sandoval, who helped prepare the pre-GAW volume; as well as April Hopstetter, Manager of Technical Publications and Printing at the Southwest Foundation for Biomedical Research, together with Maria Messenger and Shirley Schoeppel, who assisted with editing of the GAW13 proceedings.

We are grateful to Bronya Keats and the local organizing committee – Judy LaBorde, Stephanie Laurent, Diptasri Mandal, Jerlyn Rose, and Susan Stauss – for devoting countless hours to the planning and organization of a very successful GAW13 in New Orleans. Expenses of local organization were defrayed in part by generous donations from the Louisiana State University Health Sciences Center and the LSU Health Sciences Center Foundation. Souvenirs of New Orleans were provided by Baumer Foods, the Gumbo Shop, Chef Paul Prudhomme, the Zatarain Company, Audubon Nature Institute, Harrah's New Orleans Casino, Aunt Sally's Praline Shops, Inc., *WHERE New Orleans* Magazine, and the New Orleans Convention and Visitors Bureau.

Numerous organizations provided funding for scholarships to postdoctoral fellows and graduate students to help defray their expenses in attending GAW13: the National Heart, Lung, and Blood Institute provided 15 scholarships; National Institute of Mental Health, 9 scholarships; Autogen Ltd., 1.5 scholarships; and CANFOR/GmbH, Genometrix, and Genomica, 4.5 scholarships total. We are grateful for their generosity.

Long-term planning for the Genetic Analysis Workshops is the responsibility of the Genetic Analysis Workshop Advisory Committee. Its members are Chris Amos, Max Baur, Françoise Clerget-Darpoux, Cathy Falk, Lynn Goldin, Sue Hodge, Jean MacCluer (chairman), Anne Spence, Brian Suarez, and Duncan Thomas.

The National Institute of General Medical Sciences has provided continuous funding for the Genetic Analysis Workshops since 1982, through grant R01 GM31575 to Jean MacCluer. We are particularly grateful to Irene Eckstrand of NIGMS for her enthusiasm and interest in the GAWs during the past 22 years. The Genetic Analysis Workshops would not be possible without the support of Dr. Eckstrand and NIGMS.

Finally, the Genetic Analysis Workshops could not have enjoyed continued success without the ongoing, enthusiastic support of the GAW participants.