BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Elizabeth A. Thompson

POSITION TITLE: Professor of Statistics

eRA COMMONS USER NAME (credential, e.g., agency login): ETHOMPSON

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Cambridge University, England	B.A. (hons)	1967-70	Mathematics
Cambridge University, England	Diploma	1970-71	Mathematical Statistics
Cambridge University, England	M.A.	1974	Mathematics
Cambridge University, England	Ph.D	1971-74	Statistics
Stanford University, Stanford, CA	postdoc	1974-75	Genetics

A. Personal Statement

I have worked in the area of the development of statistical and computational methods for the analysis of genetic data on related individuals both in pedigrees and in populations for 40 years. I have worked in the development of Monte Carlo methods for the estimation of linkage lod scores for over 20 years, a focus of this work being the analysis of complex quantitative genetic traits. I have supervised 28 Ph.D. students to completion, 19 of whom hold faculty positions in research universities or institutes, and many of whom are making significant methodological contributions in the area of Statistical Genetics and Genomics.

In relation to the current project, In addition to the contributions described in C5 below, recent work with students and postdocs develops population-based approaches to the inference and analysis of relatedness, and the use of relatedness measures in the analysis of quantitative genetics traits, We have addressed issues of statistical efficiency in the estimation of relatedness, heritability, dominance, and epistasis.

a. Glazner, C. G. and Thompson E. A. (2015) Pedigree-free descent-based gene mapping from population samples. Human Heredity: 80: 21—35,. PMCID: PMC4583811
b. Raffa, J. D. and Thompson E. A. (2016) Power and effective study size in heritability studies. Statistics in Biosciences, 8: 264--283. PMCID: PMC5041140
c. Sverdlov, S., and Thompson, E. A. (2017) Combinatorial methods for epistasis and dominance. Journal of Computational Biology, 24: 267--279. PMCID: PMC5372789
d. Wang, B., Sverdlov, S., and Thompson, E. A. (2017) Efficient estimation of realized kinship from SNP genotypes. Genetics, 205: 1063--1078. PMCID: PMC5340323

B. Positions and Honors

Positions and Employment

1975--78 Research fellow, King's College, Cambridge
1976--85 University Lecturer, Department of Pure Mathematics and Mathematical Statistics, Cambridge University (tenured from March 1979)
1978--81 Official Fellow and Financial Tutor, King's College, Cambridge
1981--85 Official Fellow, College Lecturer and Director of Studies in Mathematics, Newnham College, Cambridge
1985(Dec)-- Professor, Department of Statistics, University of Washington
(Chair, 1989-1994, and 2011-2014)
1988--2004 and Professor, Department of Biostatistics,

2000--2006 and Adjunct Professor of Statistics, North Carolina State University
2000-- and Adjunct Professor of Genome Sciences (until 2001, Genetics), University of Washington
2006-- and Adjunct Professor of Biostatistics, University of Washington

Other major professional experience

1975,77 Visiting Scholar, Human Genetics, Univ. of Michigan, Ann Arbor (3/75-5/75, 6/77-9/77)
1976 Visiting Research Consultant, University of Utah (6/76-8/76)
1978 Visiting Scientist, University of Utah (4/78-9/78)
1979-81 Member, Electors to Fellowships, King's College, Cambridge
1984-86 Faculty Board of Mathematics, University of Cambridge.
1988 Consultant, DMS Inc., Salt Lake City, Utah (12/87-3/88)
1992 Visiting Professor, Rutgers University (Center for Theoretical and Applied Genetics) (12/91-3/92)
1994-1997 Member, NRC Committee of Applied and Theoretical Statistics
1995-2005 Member, Program in Mathematics and Molecular Biology
1998 President, West North American Region of the International Biometric Society
1997-2000 Member, Technology Working Group of the NIJ panel on Forensic DNA
1997-2001 Member of Council, International Statistical Institute
2002-2003 Visiting Professor of Statistics, North Carolina State University
2006-2013 Member of Council, International Biometric Society
2006 Visiting Rothschild Professor, University of Cambridge, UK. (Nov-Dec)
2009-2010 Member, NRC Committee on Scientific Approaches used in the Investigation of the 2001 *Bacillus anthracis* Mailings
2014- Member, NRC Board on Mathematical Sciences and their Applications.
2015-2018 President-elect (2015) and President (2016-17) of the International Biometric Society.
2017 Carnegie Centenary Visiting Professor, Univ. StAndrews, UK (1/2017 – 6/2017)

Academic Honors

1968--74 Prizes, scholarships and studentships, Newnham College, Cambridge
1973 Smith's Prize (for predoctoral research), University of Cambridge
1973--74 Sims Scholarship, University of Cambridge
1975 Stott Prize (for postdoctoral research), Newnham College
1974--78 Elected to Junior Research Fellowship, King's College, Cambridge
1978--82 Elected to Senior Research Fellowship, King's College, Cambridge
1981 Elected to **International Statistical Institute**.
1988 Awarded Doctor of Science degree, University of Cambridge.
1994 R.A.Fisher Lecturer, Joint Statistical Meetings, Toronto.
1996 Neyman Lecturer (IMS), Joint Statistical Meetings, Chicago.
1998 Elected to **American Academy of Arts and Sciences**.
2001 Jerome Sacks award for cross-disciplinary research, National Institute for Statistical Sciences
2001 Weldon Prize for contributions to Biometric Science, awarded by University of Oxford
2002-03 Guggenheim Fellow for period 09/2002- 03/2003
2005 Mary Cartwright Lecturer, London Mathematical Society, UK
2006 Distinguished Lecturer in Statistical Science, Fields Institute, Toronto.
2006 XXVII Fisher Memorial Lecture, Cambridge, UK
2008 Elected to the **US National Academy of Sciences**
2013 Elected an Honorary Fellow of Newnham College, Cambridge.

C. Contribution to Science

1. A theme of my work since 1972 has been computation of probabilities of genetic data observed jointly on multiple related individuals. These probabilities are functions of the gene identity by descent (IBD) , and the probabilities of IBD derive from a known or hypothesized pedigree structure. Together with others, I developed methods for the computation of probabilities on extended complex pedigree structures; these methods predate analogous methods now widely used for graphical models and for inference of network structures. Our methods enabled likelihood inference in segregation analysis and of allelic origins on the

large and complex pedigrees of genetic isolates. Examples of many of these inference issues and methods are presented in my 1986 monograph.

a. Thompson, E.A. (1974) Gene identities and multiple relationships. *Biometrics* **30:** 667—680.
b. Cannings, C., Thompson, E.A. and Skolnick, M.H. (1978). Probability functions on complex pedigrees. Adv. Appl. Prob. 10, 26–61.
c. Thompson, E.A. (1983) Gene extinction and allelic origins in complex genealogies*. Proc. Roy. Soc. (Lond.) B* **219:** 241—251.
d. Thompson, E.A. (1986) *Pedigree Analysis in Human Genetics.* Johns Hopkins University Press, Baltimore, MD.

2. A second theme dating from my early research but continuing to the present has been the analysis of the structure of relationship among populations. As a PhD student I developed an EM algorithm for maximum likelihood estimation of tree structures of relatedness among human populations diverging under a model of random genetic drift. This "EM algorithm" predates the formal development of EM in the statistical literature, and its current broad application. In later work, with colleagues and with students, I developed methods of likelihood inference for problems of population variation using newer genetic data types they became available: for example, rare alleles, and haplotypic variation.
    a. Thompson E.A. (1975). Human Evolutionary Trees. Cambridge University Press, Cambridge
    b. Thompson, E. A., Neel, J. V., Smouse, P. E. and Barrantes, R. (1992). Microevolution of the Chibcha-speaking peoples of lower Central America: rare genes in an Amerindian Complex. Amer. J. Hum. Genet. 51, 609-626. PMCID: PMC1682703
    c. Thompson, E. A. and Neel, J. V. (1997). Allelic association and allele frequency distribution as a function of social and demographic history. American Journal of Human Genetics 60, 197--204. PMCID: PMC1712551
    d. Chapman, N. H. and Thompson E. A. (2002) The effect of population history on the lengths of ancestral chromosome segments. Genetics 162: 449—458. PMCID: PMC1462250.

3. As new genetic data brought the possibility of inference on ever larger and more complex pedigree structures, and as computing speeds multiplied, using Monte Carlo estimates of likelihood functions and surfaces became a practical proposition. My PhD student, and later collaborator, C.J.Geyer and I developed new approaches to Monte Carlo Likelihood. Here Geyer developed the foundational ideas and theory, and I developed specific approaches for some key inference problems of statistical genetics, including ancestral inference and genetic linkage mapping.
    a. Geyer, C.J. and Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependent data (with Discussion). J. Roy. Statist. Soc. (B) 54: 657—699.
    b. Thompson, E. A. (1994). Monte Carlo likelihood in the genetic mapping of complex traits. Phil. Trans. Roy. Soc. (Lond.) Series B 344, 345-351.
    c. Geyer, C. J. and Thompson, E. A. (1995) Annealing Markov chain Monte Carlo with applications to ancestral inference. Journal of the American Statistical Association 90, 909-920.
    d. Thompson, E. A. (2000) Statistical Inferences from Genetic Data on Pedigrees. NSF-CBMS Regional Conference Series in Probability and Statistics. Volume 6. IMS, Beachwood, OH.

4. Markov chain Monte Carlo (MCMC) is a specific form of Monte Carlo, enabling realizations from probability distributions known only up to a normalizing constant. The latent-variable structure of genetic linkage data is such that this unknown constant is itself the probability of observed data that is the core of the linkage LOD score for data observed on pedigree structures. My research group was focused on MCMC methods for inference in linkage analysis from an early stage. Together with PhD and postdoctoral students, I developed ever more effective methods for MCMC realization of latent inheritance conditional on genetic marker data, and hence MCMC estimates of linkage LOD scores. In the mapping of genetic traits, these methods enabled full use of the power of multiple markers and extended pedigrees with many unobserved individuals. Our MORGAN software makes these methods available to all.

a. Guo S.W. and Thompson, E. A. (1992) A Monte Carlo method for combined segregation and linkage analysis. Amer. J. Hum. Genet. 51, 1111-1126.
b. George, A. W. and Thompson, E. A. (2003) Discovering disease genes: Multipoint linkage analysis via a new Markov Chain Monte Carlo approach. Statistical Science: 18: 515--531.
c. Tong, L. and Thompson, E. A. (2008) Multilocus lod scores in large pedigrees. Human Heredity 65**:** 142—153. PMCID: PMC2701716
d. Thompson, E.A. (2011) The structure of genetic linkage data: from LIPED to 1M SNPs. Human Heredity, 71: 86--96. PMCID: PMC3136382

5. While data on extended pedigrees contain much information, even more can be gained by considering the unspecified coancestry of members of a population. My recent work has been in inference of IBD in individuals not known to be related, and in the development of methods for use of this inferred IBD in the resolution of complex genetic traits. Together with students, I have developed methods for inference of IBD across a chromosome among multiple individuals, using SNP data at much greater density than can be used in pedigree-based linkage analyses. With colleagues, I have shown how this population-based IBD can be used effectively in genetic mapping. Still more power may be gained by combining population-based and pedigree-based iIBD inferred from marker data.
   a. Brown, M. D., Glazner, C. G., Zheng, C., and Thompson, E. A. (2012) Inferring coancestry in population samples in the presence of linkage disequilibrium. Genetics, 190: 1447--1460. PMCID: PMC3316655
   b. Browning S. and Thompson E. A. (2012) Detecting rare=variant associations by identity-by-descent mapping in case-control studies. Genetics 190: 1521-1531. PMCID: PMC 316661
   c. Thompson, E. A. (2013) Identity by descent: Variation in meiosis, across genomes, and in populations. Genetics 194: 301--326. PMCID: PMC3664843
   d. Zheng, C., Kuhner, M. K., and Thompson, E. A. (2014) Joint inference of identity by descent along multiple chromosomes in population samples. Journal of Computational Biology 21: 185--200. PMCID: PMC 3948483

List of Published Work in MyBibliography:
https://www.ncbi.nlm.nih.gov/myncbi/browse/collection/41087487/
(This is a subset of total publications of 4 books and 234 papers, not including abstracts.)

D. Research Support
Current
R37 GM 46255 (Current Years 22-26) (PI Thompson) 9/01/1991—1/31/2018
NIH/NIGMS
Methods for the Genetic Epidemiology of Complex Traits
Development of computational and statistical methods, including Markov chain Monte Carlo methods, for the genetic analysis of complex traits using multilocus data on extended pedigrees. Current approaches include the use of dense genomic data in the inference of coancestry and descent. Implementation and distribution of software for these methods.
Role: PI

P01 GM099568 (PI Weir) 5/01/2012 —4/30/2017
NIH/NIGMS
Statistical and Population Genetics.
This program project supports five senior investigators in the development and application of statistical methods for genomic data.
Project 4: Resolving Complex Traits through Inferred Coancestry of Genome Segments
Development of computational and statistical methods for the inference of coancestry and shared genome in individuals not know to be related. Development of methods for using inferred coancestry in the analysis of complex quantitative phenotypes (including longitudinal data) observed in population samples. Implementation and distribution of software for these methods.
Role: PI of Project 4.

Completed Research Support: none in last 3 years.