

## OLIVER BEMBOM

---

### **PhD candidate in Biostatistics**

Division of Biostatistics  
School of Public Health  
University of California, Berkeley  
140 Warren Hall, #7360  
Berkeley, CA 94720-7360

Tel: (510) 642-2313  
Fax: (510) 643-0239  
E-mail: [bembom@berkeley.edu](mailto:bembom@berkeley.edu)

---

### **EDUCATION**

**PhD in Biostatistics**, University of California, Berkeley, expected December 2007.

Grade point average: 4.0

Adviser: Dr. Mark J. van der Laan.

Dissertation: *Statistical methods for causal inference when the assumption of experimental treatment assignment is violated.*

**MPH in Epidemiology/Biostatistics**, University of California, Berkeley, 2004.

Grade point average: 4.0

Master's thesis: *Asthma hospitalization rates among young children in California.*

**BA in Molecular and Cell Biology**, University of California, Berkeley, 2002.

Grade point average: 4.0

*Highest Distinction of General Scholarship.*

*Outstanding Undergraduate Scholar Award*: The department's second-highest honor, awarded annually "for outstanding undergraduate scholarly achievements in Molecular and Cell Biology".

Honors thesis: *Use of Intracellular Single-chain Antibodies To Modulate Transcription Factor Activity.*

### **RESEARCH INTERESTS**

Causal inference in observational studies based on Marginal Structural Models.

Individualized treatment rules for sequentially randomized trials.

### **RESEARCH EXPERIENCE**

**Graduate Student Researcher**, August 2004 - Present.

Dr. Mark J. van der Laan, Division of Biostatistics, University of California, Berkeley.

Various research projects including a collaboration with Dr. Ira B. Tager aimed at estimating the causal effect of leisure-time physical activity on mortality in the elderly as well as the time-dependent modification of this causal effect by cardiovascular morbidity.

**Graduate Student Researcher**, Winter 2005.

Dr. Joseph Eisenberg, Divisions of Epidemiology and Environmental Health Sciences, University of California, Berkeley.

Research of statistical methods for data arising from a stochastic disease transmission model (in collaboration with Dr. Alan Hubbard).

**Research Assistant**, May 2003 - August 2004.

Dr. Peggy Reynolds, Environmental Health Investigations Branch, California Department of Health Services.

Assistance with data analysis and GIS tasks for two major epidemiologic research studies.

**Volunteer Researcher**, August 2000 - May 2002.

Laboratory of Dr. Mark Schlissel, Department of Molecular and Cell Biology, University of California, Berkeley.

Design and execution of various research projects in the area of molecular immunology.

**TEACHING EXPERIENCE****Teaching Assistant**, August 2000 - May 2003

Department of Statistics, University of California, Berkeley

Teaching assistant for various undergraduate statistics classes. Duties included conducting discussion sections, preparing and grading quizzes, as well as grading midterm and final exams.

*Outstanding Graduate Student Instructor Award*, 2002 - 2003.

**Substitute Lecturer**, Spring 2006

Division of Biostatistics, University of California, Berkeley

Preparing and holding the following lectures for a graduate-level biostatistics course on *Survival Analysis and Causality*:

- Derivation of the Kaplan-Meier estimator as the nonparametric maximum-likelihood estimator of the survival function*
- Derivation of Greenwood's formula for the asymptotic variance of the Kaplan-Meier estimator*
- Estimation and inference for multivariate multiplicative intensity models based on the profile partial likelihood*
- Fitting multiplicative intensity models using `coxph()` in R*

**Substitute Lecturer**, Fall 2006

Division of Biostatistics, University of California, Berkeley

Preparing and holding the following lectures for a graduate-level biostatistics course on *Censored Longitudinal Data and Causality*:

- Derivation of the double robust estimator of the marginal variable importance of a binary treatment variable*
- The causal effect of recent leisure-time physical activity on all-cause mortality among the elderly*

## COURSEWORK

- Linear Algebra, Analysis I, Analysis II, Advanced Introduction to Probability and Statistics, Theoretical Statistics.
- Statistical Analysis of Categorical Data, Introduction to Multivariate Statistics, Applications of Statistics to Genetics and Molecular Biology, Survival Analysis and Causality, Censored Longitudinal Data and Causality, Multiple Testing and Loss-Function Based Estimation: Applications in Genomics.
- Epidemiological Analysis, Epidemiologic Methods II.
- Introduction to Statistical Computing, Statistical Computing, Statistical Analysis using R, Programming and Algorithm Design for Computational Biology and Genomics Applications, C++ for Programmers.

## AWARDS AND HONORS

- Jiann-Ping Hsu Award for Excellence and Scholarship, 2006 - 2007.
- Walter G. and Julia Perker Scholarship, 2003 - 2004.
- School of Public Health University Fellowship, 2002 - 2004.
- Phi Beta Kappa, 2002.
- Siegfried Roggenbuck Scholarship, 1999-2002.
- Edmund Vey Scholarship, 2000.
- Ernst Stewner Scholarship, 1999-2000.

## PUBLICATIONS

### Refereed Publications

1. M.G. Goveia, N. Shaikh, Windham G, **O. Bembom**, K. Feldman, and R. Kreutzer (2005). Asthma awareness/environmental practices in childcare centers. *Pediatric Asthma, Allergy and Immunology*, 18(1):12-24.
2. Y. Wang, **O. Bembom**, and M.J. van der Laan (2007). Data-Adaptive Estimation of the Treatment-Specific Mean. *Journal of Statistical Planning and Inference*, 137(6):1871-1887.
3. **O. Bembom**, S. Keleş, and M.J. van der Laan (2007). Supervised Detection of Conserved Motifs in DNA Sequences with *cosmo*. *Statistical Applications in Genetics and Molecular Biology*: Vol. 6 : Iss. 1, Article 8. <http://www.bepress.com/sagmb/vol6/iss1/art8>.

4. **O. Bembom** and M.J. van der Laan (2007). Statistical methods for analyzing sequentially randomized trials. *Journal of the National Cancer Institute*, 99:1577-82.

### Book Chapters and Conference Proceedings

5. **O. Bembom**, M.L. Petersen, and M.J. van der Laan (2006). Identifying important explanatory variables for time-varying outcomes. In W. Dubitzky, M. Granzow, and D.P. Berrar (eds.), *Fundamentals of Data Mining in Genomics and Proteomics*, Springer, Chapter 11, p. 227-250.

### Technical Reports and Submitted Manuscripts

6. **O. Bembom**, M.J. van der Laan, and I.B. Tager (2007). The causal effect of recent leisure-time physical activity on all-cause mortality among the elderly. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 214. <http://www.bepress.com/ucbbiostat/paper214>.
7. **O. Bembom** and M.J. van der Laan (2007). Analyzing sequentially randomized trials based on causal effect models for realistic individualized treatment rules. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 216. <http://www.bepress.com/ucbbiostat/paper216>. (Submitted to *Statistics in Medicine*.)
8. **O. Bembom** and M.J. van der Laan (2007). Estimating the effect of vigorous physical activity on mortality in the elderly based on realistic individualized treatment and intention-to-treat rules. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 217. <http://www.bepress.com/ucbbiostat/paper217>. (Submitted to the *Electronic Journal of Statistics*.)
9. **O. Bembom**, M.L. Petersen, S.-Y. Rhee, W. J. Fessel, S.E. Sinisi, R.W. Shafer, and M.J. van der Laan (2007). Biomarker discovery using targeted maximum likelihood estimation: Application to the treatment of antiretroviral resistant HIV infection. *UC Berkeley Division of Biostatistics Working Paper Series*. Working Paper 221. <http://www.bepress.com/ucbbiostat/paper221>. (Submitted to *Statistics in Medicine*.)
10. **O. Bembom**, M.J. van der Laan, T. Haight, I.B. Tager (2007). Lifetime and current leisure time physical activity and all-cause mortality in an elderly cohort. (Submitted to *Epidemiology*.)
11. **O. Bembom** and M.J. van der Laan (2007). Data-adaptive selection of the truncation level for Inverse-Probability-of-Treatment-Weighted estimators. (Submitted to *Biometrics*.)

### California Department of Health Services Publications

12. J.K. Stockman, N. Shaikh, J. von Behren, **O. Bembom**, and R. Kreutzer (2003). California County Asthma Hospitalization Chart Book: Data from 1998-2000. California Department of Health Services, Environmental Health Investigations Branch.
13. J.K. Stockman, K. Garcia, J. von Behren, **O. Bembom**, N. Shaikh, and R. Kreutzer (2004). Asthma Disparities Among African-Americans. *California Asthma Facts, Volume 2, Issue 1*. California Department of Health Services, Environmental Health Investigations Branch.

## Manuscripts in Preparation

14. **O. Bembom** and M.J. van der Laan. Data-adaptive selection of the adjustment set in variable importance estimation.
15. S. Morshed, T. Miclau III, **O. Bembom**, and J.M. Colford. Delayed femoral shaft fracture internal fixation in multi-system trauma patients decreases mortality in the National Trauma Data Bank: An application of inverse-probability-of-treatment-weighted analysis to polychotomously defined treatment time.

## Software

16. **O. Bembom**, S. Keleş, and M.J. van der Laan (2006). `ccosmo`: A stand-alone C program for the supervised detection of conserved motifs in DNA sequences (<http://cosmoweb.berkeley.edu/software.html>).
17. **O. Bembom**, S. Keleş, and M.J. van der Laan (2006). `cosmoweb`: A web application for the supervised detection of conserved motifs in DNA sequences (<http://cosmoweb.berkeley.edu>).
18. **O. Bembom**, F. Gallusser, and S. Dudoit (2006). `cosmo`: An R package for the supervised detection of conserved motifs in DNA sequences (<http://www.bioconductor.org>, <http://cosmoweb.berkeley.edu/software.html>).
19. F. Gallusser, **O. Bembom**, and S. Dudoit (2007). `cosmoGUI`: An R package for the interactive construction of constraint sets as used by the `cosmo` R package (<http://www.bioconductor.org>, <http://cosmoweb.berkeley.edu/software.html>).
20. **O. Bembom** (2007). `seqLogo`: An R package for plotting DNA sequence logos (<http://www.bioconductor.org>, <http://cosmoweb.berkeley.edu/software.html>).

## PRESENTATIONS

### Invited Presentations

- **O. Bembom** and M.J. van der Laan. Analyzing sequentially randomized trials based on causal effect models for realistic individualized treatment rules. *2007 Joint Statistical Meetings*. July 29 - August 2, 2007. Salt Lake City, Utah, USA.
- **O. Bembom**. Supervised detection of conserved motifs in DNA sequences with `cosmo`. *UC Berkeley Statistics and Genomics Seminar*. September 7, 2006. Berkeley, California, USA.
- **O. Bembom**, M.J. van der Laan, A.E. Hubbard, and N.P. Jewell. Estimation of Treatment Effects in Randomized Trials with Non-compliance and a Dichotomous Outcome. *23<sup>rd</sup> International Biometrics Conference*. July 16-21, 2006. Montréal, Quebec, Canada.
- **O. Bembom**. An application of History-Adjusted Marginal Structural Models. *UC Berkeley Biostatistics Seminar*. November 29, 2005. Berkeley, California, USA.

- 
- **O. Bembom.** COMODE - A web application for constrained motif detection. *UC Berkeley Statistics and Genomics Seminar*. April 14, 2005. Berkeley, California, USA.
  - **O. Bembom.** Implementing COMODE as a web application. *UC Berkeley Biostatistics Seminar*. February 1, 2005. Berkeley, California, USA.

### **Contributed Presentations**

- S. Morshed, T. Miclau III, **O. Bembom**, and J.M. Colford, Jr.. Delayed femoral shaft fracture internal fixation in multi-system trauma patients decreases mortality in the National Trauma Data Bank. *Orthopaedic Trauma Association Annual Meeting*. October 18-20, 2007. Boston, Massachusetts, USA.
- J. von Behren, N. Shaikh, **O. Bembom**, and R. Kreutzer. Predictors of asthma hospitalization and prevalence rates in California. Poster presentation. *American Public Health Association 131<sup>st</sup> Annual Meeting and Exposition*. November 15-19, 2003. San Francisco, California, USA.

### **REFEREE SERVICE**

- *The Annals of Statistics.*
- *Bioinformatics.*
- *Journal of the Royal Statistical Society, Series B.*
- *Journal of Statistical Planning and Inference.*
- *Statistical Applications in Genetics and Molecular Biology.*
- *Statistics in Medicine.*

### **COMPUTING EXPERIENCE**

Operating systems: UNIX, Linux, Windows.

Languages: R, SAS, C, C++, Java, Perl, HTML,  $\LaTeX$ .

Applications: Arcview 3.x, ArcGIS 8, MS Access, SQL.

### **LANGUAGE SKILLS**

Fluent in German and English.

Beginning French and Spanish.