

## A Marginalized Multilevel Modeling Approach for Clustered Longitudinal Binary Data with Time Dependent Covariates

Diana L. Miglioretti<sup>1,2</sup> and Patrick J. Heagerty<sup>2</sup>

<sup>1</sup>Center for Health Studies, Group Health Cooperative

<sup>2</sup>Department of Biostatistics, University of Washington

Modeling multilevel binary data continues to pose significant challenges in many situations. For short time series typical in longitudinal epidemiological studies, hierarchical or random effects models using the standard assumption of normally distributed subject-specific effects are difficult to fit and may not be reasonable. In addition, interest is often in the marginal or population-average effects, which are not directly modeled by hierarchical models. We propose a marginalized multilevel model that combines a logistic regression model to estimate the influence of covariates on the marginal mean and a separate conditional logistic regression model that captures both the serial dependence within short time series and the correlation within larger cluster. The model is estimated using a Bayesian approach. Extensions for endogenous time-dependent covariates will be discussed. The methods will be illustrated using data from the Breast Cancer Surveillance Consortium to estimate mammography accuracy in a repeatedly screened population.