1.a. **Full Title**: Joint modeling of longitudinal data and recurrent events in the presence of informative terminal event

**b. Abbreviated Title (Length 26 characters)**: Joint modeling

2. **Writing Group**:

   Writing group members: Se Hee Kim, Donglin Zeng

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. _SK_ [please confirm with your initials electronically or in writing]

**First author**: Se Hee Kim

Address: 3105D McGavran-Greenberg Hall
School of Public Health, CB #7420
University of North Carolina
Chapel Hill, NC 27599-7420

Phone: 919-265-4693    Fax:  
E-mail: sehee@email.unc.edu

**ARIC author** to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (this must be an ARIC investigator).

   Name: 
   Address: 

   Phone:    Fax:  
   E-mail: 

3. **Timeline**: This work could be done by June 30, 2009.

4. **Rationale**: In many cohort studies, we observe that a patient may experience the same event at multiple times (e.g., repeated hospital admissions or recurrent strokes in the ARIC study) before death or drop-out by the end of studies. At the same time, longitudinal markers (e.g., LDL, HDL, blood pressure in the ARIC study) are collected during patient’s time course. Recurrent event, death and longitudinal markers are


dependent and informative of one other. Analyzing one process but ignoring the dependence from the other two processes may lead to biased or result in inefficient inference. Therefore, it is important to jointly model recurrent events, death and longitudinal markers altogether. By this way, we will be able to make use of all data information and identify the effects of risk factors after correctly controlling the interplay among these processes.

5. **Main Hypothesis/Study Questions**: The primary objective of this paper is to develop a methodology analyzing the longitudinal measures, recurrent event times, and terminal event time simultaneously. Using real data from the ARIC study as an application of the proposed method, we can characterize the relationship between clinical markers, recurrent events and death, assess the effects of (time-varying) covariates on clinical markers, recurrent events and death, and utilize the final models for the accurate prediction of risk of recurrent events and death.

6. **Design and analysis (study design, inclusion/exclusion, outcome and other variables of interest with specific reference to the time of their collection, summary of data analysis, and any anticipated methodologic limitations or challenges if present).**

The study design is of a longitudinal cohort study with clinical markers (lipids levels - total cholesterol, LDL, HDL; blood pressure, BMI), recurrent events (repeated hospital admissions, recurrent strokes, MI, CVD), and terminal event (death) as outcome variables. The covariates will include traditional risk factors (site, age, gender, race, smoking status, alcohol use, diabetes, hypertension, family history, etc), through the end of study. A joint model will be fit with the subject-specific random effects shared between Gaussian process for repeated measures and survival distribution and between counting process for recurrent events and survival distribution. Box-Cox and logarithm transformation for the cumulative intensity function of counting process and the cumulative hazard function of survival will be considered and compared. Nonparametric maximum likelihood estimation will be used via EM algorithm to estimate the regression parameters and the cumulative intensity/hazard functions.

7.a. **Will the data be used for non-CVD analysis in this manuscript?**  ____ Yes  _x_ No

b. **If Yes, is the author aware that the file ICTDER03 must be used to exclude persons with a value RES_OTH = “CVD Research” for non-DNA analysis, and for DNA analysis RES_DNA = “CVD Research” would be used?**  ____ Yes  ____ No

(This file ICTDER03 has been distributed to ARIC PIs, and contains the responses to consent updates related to stored sample use for research.)
8.a. Will the DNA data be used in this manuscript?  ____ Yes  ___x__ No

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the file ICTDER03 must be used to exclude those with value RES_DNA = “No use/storage DNA”?  
____ Yes  ____ No

9. The lead author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the web site at:  http://www.cscce.unc.edu/ARIC/search.php

___x___  Yes  _______ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)? None.

11. a. Is this manuscript proposal associated with any ARIC ancillary studies or use any ancillary study data?  ____ Yes  __x__ No

11.b. If yes, is the proposal
____  A. primarily the result of an ancillary study (list number* ____________)
____  B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables; list number(s)* ____________ ____________ ____________)

*ancillary studies are listed by number at http://www.cscce.unc.edu/aric/forms/

12. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.