How Often Should We Collect Data in Functional Trajectory Studies Among Older People?

Yixiu Liu[1], Depeng Jiang[1,2], Robert B Tate[1], Philip D St John[3]

[1] Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba;
[2] The George Fay Yee Center for Healthcare Innovation, University of Manitoba , Winnipeg, Manitoba;
[3] Department of Internal Medicine, University of Manitoba, Winnipeg, Manitoba

Background

- Results of improper frequency of data collection
  - More dropouts and nonresponses
  - Unable to investigate complex functional forms
  - True changing pattern get masked or false trajectories are observed
- Factors affect the choice of data collection cycles
  - Research objectives, aggressive diseases vs. chronic diseases
  - Study population, younger vs. older
  - Limited funding and human resources
- **No general rule exist** – only careful considerations in each specific study context

Objective

To examine the impact of data collection cycles on estimations from joint models of physical functioning trajectories and its association with risk of death.

Methods

The Manitoba Follow-Up Study (MFUS):

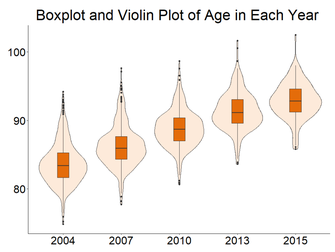
- Initiated on 1st July 1948
- 3,983 Royal Canadian Air Force air crew male recruits
- The Successful Aging Questionnaire (SAQ), with the short-form 36 (SF-36) as an important component, send out annually since 2004
- Data collected from 2004 to 2015 were used

Measures:

- The Physical Component Score (PCS): measures people's physical functioning
- The death date: has been collected through questionnaires, administrative database and direct phone calls

Sample:

- 964 males who were alive in 2004
- 5357 repeated measurements of PCS
- 20.7% right censoring (200 were alive by 2015)
- 53.7% missing repeated measurements



Joint Models:

- Advantages over traditional methods
  - Missing not at random in the longitudinal process: death → missing PCS
  - Assumptions for time-varying covariates in the extended Cox are violated
    - Measure without error, remain constant until next time point, existence does not depend on failure status
  - Separate analyses fail to capture the relationship between the two processes and may result in biased results
- Two components
  - Longitudinal submodel
    $y_i(t) = m_i(t) + \varepsilon_i(t) = x_i^T(t)\beta + z_i^T(t)b_i + \varepsilon_i(t)$, the same as growth curve model (GCM)
  - Survival submodel
    $h_i(t|\mathcal{M}_i(t), \omega_i) = h_0(t)\exp\{\gamma^T\omega_i + \alpha W_i(t)\}$, where $W_i(t)$ denotes the linking approach of the longitudinal and survival processes

Analyses:

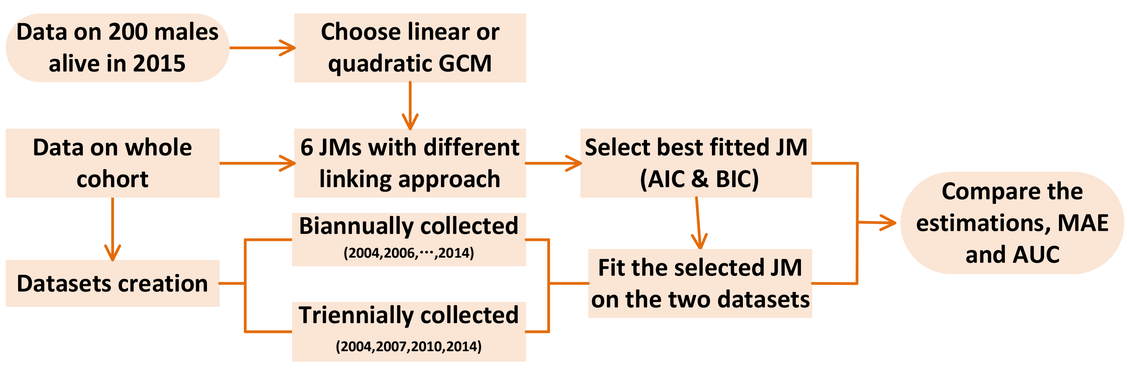


Table 1: Joint Models

	Linking Approach
Model a	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)\exp\{\alpha_1 m_i(t)\}$
Model b	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)\exp\{\alpha_1 m_i(t) + \alpha_2 m_i'(t)\}$
Model c	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)\exp\left\{\alpha_2 \int_0^t m_i(s)ds\right\}$
Model d	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)exp\left\{\alpha_1 m_i(t) + \alpha_2 \int_0^t m_i(s)ds\right\}$
Model e	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)\exp\left\{\alpha_2 \int_0^t \varpi(s) m_i(s)ds\right\}$
Model f	$h_i(t\|\mathcal{M}_i(t), \omega_i) = h_0(t)exp\left\{\alpha_1 m_i(t) + \alpha_2 \int_0^t \varpi(s) m_i(s)ds\right\}$

Data Exploration

Figure 1: Box and Violin Plot of PCS

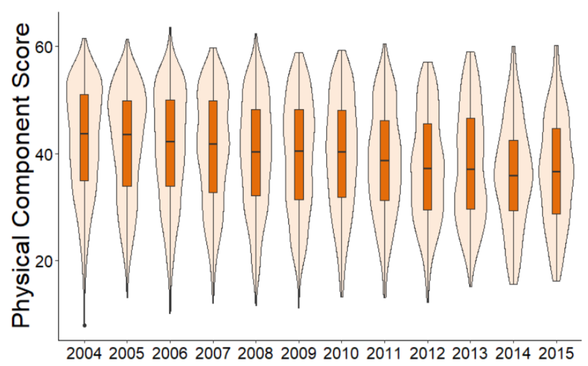
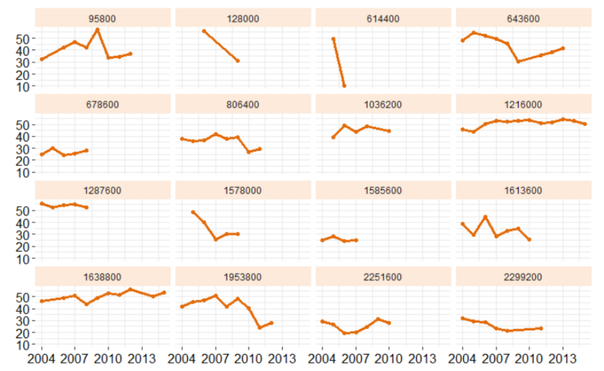


Figure 2: Individual Trajectory



Results

Table 2: Goodness-of-fit Indices of the GCMs

Model	Log-Likelihood	AIC	BIC
Intercept only	-13004	13010	13020
Linear	-12503	12515	12535
Quadratic	**-12466**	**12486**	**12519**

Table 3: Goodness-of-fit Indices of the 6 JMs

Model	Log-Likelihood	AIC	BIC
Model a	-20491	41018	41106
Model b	**-20329**	**40696**	**40789**
Model c	-20563	41162	41249
Model d	-20478	40994	41086
Model e	-20553	41142	41230
Model f	-20472	40982	41075

Table 4: Parameter Estimations on the Data from the Three Study Designs

	Longitudinal submodel							Survival submodel	
	$\hat{\sigma}_\varepsilon$	$\hat{\sigma}_{b_{0i}}$	$\hat{\sigma}_{b_{1i}}$	$\hat{\sigma}_{b_{2i}}$	$\hat{\beta}_{00}$	$\hat{\beta}_{01}$	$\hat{\beta}_{02}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$
Annual	5.35	8.76	1.38	0.09	42.52(0.28)***	-1.18(0.10)***	-0.01(0.01)	-0.06(0.005)***	-0.09(0.05)†
Biennial	5.21	8.52	1.05	0.08	42.50(0.21)***	-0.98(0.11)***	-0.007(0.01)	-0.05(0.005)***	0.03(0.05)
Triennial	5.04	8.86	1.23	0.06	42.56(0.33)***	-0.93(0.13)***	-0.011(0.02)	-0.06(0.006)***	0.05(0.08)

Figure 3: Mean Absolute Error of Predicted PCS in Each Year of the Three Study Designs

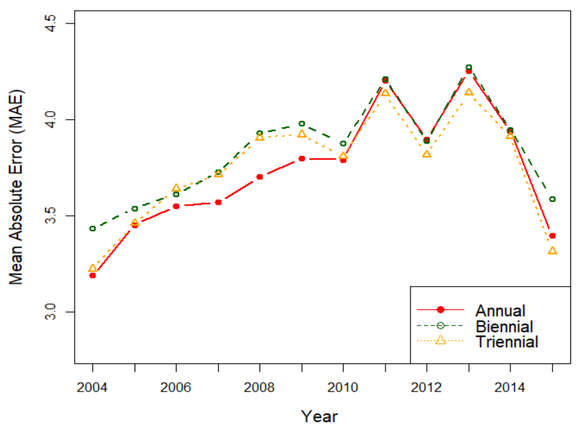
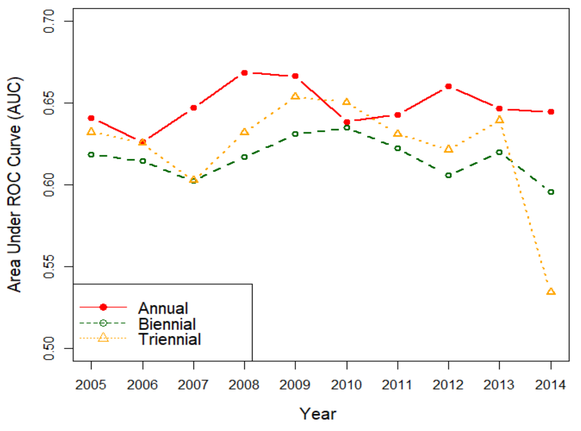


Figure 4: AUCs of the Survival Process in Each Year of the Three Study Designs



Conclusions

- The marginal-significant effect of the change rate of the physical functioning on the hazard of death cannot be captured in a study design that adopts longer data collection interval.
- The inter-subject variation in the trajectories of physical functioning over time could be substantially underestimated based on a less frequent data collection strategy.
- The impact of study design on estimations of parameters for describing the longitudinal trajectory is minimal as long as we have enough data points to estimate the individual shape of trajectory (e.g., 3 points for linear and 4 points for quadratic).
- The predictions of mortality risk obtained using annual measurements of physical functioning were better than using biennial or triennial measurements, while the predictions obtained using biennial or triennial measurements were almost equivalent.
- The selection of data collection cycles depends on the complexity of longitudinal process and its linkage to survival outcomes.
- This study provides a reference for selecting the follow-up strategy in an aging longitudinal study when focusing on the functional trajectories and its linkage to the survival probability using joint models.

Acknowledgements

Data: Robert Tate (PI), The Manitoba Follow-Up Study
Funding: 1. Robert Tate (PI), Predictions and consequences of long-term functional trajectories of older men: The Manitoba Follow-Up Study. CIHR (2017-2020)
2. Depeng Jiang (PI), Person-oriented Statistical Methods for Health Services Research. CIHR (2017 – 2019).
3. The VADA Program