

Multi-Level Micro-Randomized Trial for Detecting the Proximal Effect of a Mobile Application Messages on Physical Activity

Jing Xu, Xiaoxi Yan, Caroline Figueroa, Joseph Williams, Adrian Aguilera,
Bibhas Chakraborty

kenny.xu@duke-nus.edu.sg

Centre for Quantitative Medicine, Duke-NUS Medical School, Singapore
Cancer Data Science, Children Medical Research Institute, The University of Sydney, Westmead, NSW, Australia

January 4, 2020

PhD Research-Affiliation

Department of Mathematics and Statistics, Macquarie University, Australia



PhD Research-Topic

- Proportional hazard model estimation under dependent censoring using maximum penalized likelihood
 - Brodaty, et al. (2014). Predictors of Institutionalization in Dementia: A three Year Longitudinal Study. *Journal of Alzheimer's Disease.*
 - Xu, J. et al. (2018). Proportional hazard model estimation under dependent censoring using copulas and penalized likelihood. *Statistics in Medicine.*
 - survivalMPLdc: Fit Cox Proportional Hazard Regression Model under dependent right censoring via MPL and Archimedean Copulas

PhD Research-Group

- **Statistician:** Dr Jing Xu, A/Prof Jun Ma



- **Dementia Expert:** Dr Michael Connors, Prof Henry Brodaty



PhD Research-Novelty

- The advantages over existing likelihood method, i.e. Chen (2010)
 - Smoothness baseline hazard estimates obtained by maximum penalized likelihood (MPL) method
 - The efficient algorithm of Newton-Multiplicative and Iterative (MI) algorithm
 - The consistent and asymptotic normality properties of the MPL estimates

PhD Research-Potential

- A **statistical software paper** for the R package “survivalMPLdc”
- **Other semi-parametric** hazard models, i.e. additive or accelerated failure time hazard models
- **Precision medicine**, i.e. protein selection for predicting cancer survival outcome, i.e. Emura and Chen (2016)
 - **Statistical collaborator:** A/Prof Takeshi Emura, Graduate Institute of Statistics, National Central University, Taiwan



- **Clinical collaborator:** Prof Rosemary Balleine, ProCan Cancer Pathology Research, Children Medical Research Institute, Australia



Postdoc Research-Affiliation

Centre for Quantitative Medicine, Duke-NUS, Singapore



Postdoc Research-Topic

- Sequential multiple assignment randomized trial for periodontitis (SMARTp)
 - Xu, J. et al. (2019). SMARTp: A SMART design for non-surgical treatments of chronic periodontitis with spatially-referenced and non-randomly missing skewed outcome. *Biometrical Journal*
 - SMARTp: Sample Size for SMART Designs in Non-Surgical Periodontal Trials. Xu, J. et al, 2019 <https://cran.r-project.org/web/packages/SMARTp/index.html>
- Multi-Level Micro-randomized trial (MLMRT) for mobile health (mHealth)

Postdoc Research-Group of SMARTp

- **Statistical Expert in SMART Design:** Dr Jing Xu, A/Prof Bibhas Chakraborty



- **Statistical Expert in Spatial Modelling:** Prof Dipankar Bandyopadhyay



- **Periodontist:** Dr Bryan Michalowicz



Postdoc Research-Novelty and Potential of SMARTp

- This paper proposes some novel dynamic treatment regimes (DTRs) for periodontitis.
- These DTRs can be studied under a cluster-level SMART design.
- The proposed method considers informative missingness, spatial associations and non-normal outcomes.
- R package “SMARTp” is available in CRAN
- Sample size calculation can be extended to select the **optimal treatment regime**, e.g., Artman, et al (2017)
- **Run a study**, collect the data and perform data analysis, then update the experimental design, statistical analysis plan and sample size method

Postdoc Research-Group of MLMRT

- **Statistician:** Dr Jing Xu, A/Prof Bibhas Chackraborty, Miss Xiaoxi Yan



- **Social Welfare:** A/Prof Adrian Aguilera, Dr Caroline Figueroa



- **Computer Scientist:** Dr Joseph Williams



Outline

- 1 Mobile Health
- 2 DIAMANTE study
- 3 Multi-Level Micro-Randomized Trial Design
- 4 Analysis Plan and Sample Size Calculation
- 5 Simulation
- 6 Discussion

Outline

1 Mobile Health

2 DIAMANTE study

3 Multi-Level Micro-Randomized Trial Design

4 Analysis Plan and Sample Size Calculation

5 Simulation

6 Discussion

Mobile Health (mHealth)-Definition

The practice of medicine and public health supported by mobile devices



can better reach areas, people, and healthcare practitioners with limited exposure to certain aspects of healthcare.

mHealth-Motivation

Life is getting increasingly digital

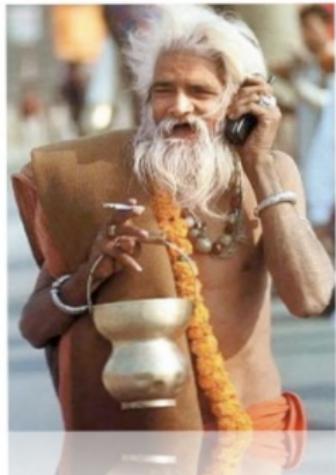


Image courtesy of Gary Bennett

Source: Duke Global Health Institute

mHealth-Motivation

Constraints of developing nations, i.e. high disease prevalence, low health care workforce and limited financial resources to support healthcare system



Malaria Clinic in Tanzania helped by **SMS for Life** program that uses cell phones to efficiently deliver **malaria vaccine**

mHealth-Motivation

Just-in-Time Adaptive Intervention (JITAI) is a special type of adaptive intervention can be delivered when and where it is needed through mobile technologies, where data are collect via a mobile device.



mHealth-Examples

- Improve HIV medication, Lewis et al (2013)
- Increase physical activity, King et al (2013)
- Supplement counseling or pharmacotherapy in treatment for substance use, Marsch (2012)
- Reduce alcohol dependence, Alessi et al (2013)

Outline

1 Mobile Health

2 DIAMANTE study

3 Multi-Level Micro-Randomized Trial Design

4 Analysis Plan and Sample Size Calculation

5 Simulation

6 Discussion

DIAMANTE App

DIAbetes and Mental Health Adaptive Notification Tracking and Evaluation



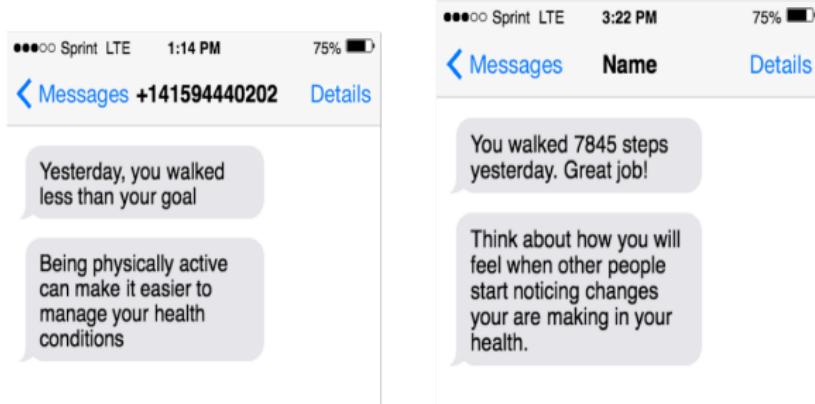
Source: <https://diamante.healthysms.org/>

Overarching Goal of DIAMANTE study

- To develop, implement and evaluate an **adaptive-learning**, clinic-integrated mobile intervention targeting **physical activity to manage co-morbid diabetes and depression** in low-income, ethnic minority patients served in the San Francisco Health Network

Developmental Goal of DIAMANTE Study

2 different messages every day, 1 minute apart



- Learn to send better messages to patients to encourage them to walk more
- Messages are multi-component interventions, with varying levels of: (a) Time Window (when to send the message), (b) Feedback Message, and (c) Motivational Message

Component 1: Time Window (when to send the message?)



- Level 0: *09:00-11:30*
- Level 1: *11:30-14:00*
- Level 2: *14:00-16:30*
- Level 3: *16:30-19:00*

Component 2: Feedback Message

Level	Description	Examples
0	None	—
1	Reaching goal	<i>“Yesterday, you did not reach your goal.”</i>
2	Steps walked yesterday	<i>“Yesterday, you walked 3824 steps.”</i>
3	Walked more/less yesterday than day before	<i>“Yesterday, you walked more than day before.”</i>
4	Steps walked yesterday, + a motivational message	<i>“You walked 8000 steps yesterday. Great job!”</i>

Component 3: Motivational Message

Level	Description	Examples
0	None	–
1	Benefit	<i>“Doing more physical activity can help reduce feelings of fatigue.”</i>
2	Self-efficacy	<i>“You have made changes to improve your health before, you can do it again.”</i>
3	Opportunity	<i>“Is there a local park you have been waiting to visit? Use it as an opportunity to get out of the house and do more steps!”</i>

Outline

- 1 Mobile Health
- 2 DIAMANTE study
- 3 Multi-Level Micro-Randomized Trial Design**
- 4 Analysis Plan and Sample Size Calculation
- 5 Simulation
- 6 Discussion

Design Development for DIAMANTE study

- **Factorial designs** the gold standard when collecting data to build multi-component behavioral interventions (*Collins et al., 2005; Nair et al., 2008; Chakraborty et al., 2009*)
- **Micro-randomized trials (MRTs)** (*Klasnja et al., 2015*) the cutting-edge trial designs suitable to take care of the time-varying, sequential nature of the interventions
 - These are sequential, full-factorial designs.
 - The possible actions for each participant are randomised at each decision point.
 - They capture the “just-in-time” intervention purpose of mobile intervention.
 - Proximal outcome: **walking steps** in next 30-min after an action sent
 - Generalized Estimating Equations (**GEE**)-type analysis

Design Towards an MRT for DIAMANTE study

- The proposed **MLMRT** called “Multi-Level Micro Randomized Trial”.
 - Involving **components with more than 2 levels**
 - Not only power-based, but also **precision-based** sample size calculations
 - **Adding new component levels in sample size calculations**
 - **Learning online** about more effective intervention components while in the trial (e.g., **Thompson sampling for contextual multi-arm bandit problems**)

Outline

1 Mobile Health

2 DIAMANTE study

3 Multi-Level Micro-Randomized Trial Design

4 Analysis Plan and Sample Size Calculation

5 Simulation

6 Discussion

Data Structure

- T decision points on **a single individual**, i.e.

$$S_1, I_1, A_1, Y_2, \dots, S_t, I_t, A_t, Y_{t+1}, \dots, S_T, I_T, A_T, Y_{T+1}$$

where

- S_t : pre-treatment at the t -th decision point
- I_t : availability at the t -th decision point
- A_t : treatment at the t -th decision point
- Y_{t+1} : proximal outcome after the t -th decision point

Proximal Effect

- The effect of treatment on proximal outcome to vary with time
- Focus on one factor (e.g. motivational message) with M levels (of message)
- E.g., the proximal effect of level m at time t as

$$\beta_m(t) = E(Y_{t+1} \mid A_{mt} = 1) - E(Y_{t+1} \mid A_{mt} = 0),$$

for $m = 1, \dots, M$ and 0 is the control (none) message, and

$$Y_{t+1} = \alpha(t) + \beta_1(t)(A_{1t} - \pi_{1t}) + \cdots + \beta_M(t)(A_{Mt} - \pi_{Mt}) + \epsilon_{it}$$

Hypothesis Testing

- The null hypothesis of no proximal effect of messages, i.e.

$$H_0 : \beta_1(t) = \dots = \beta_M(t) = 0, (t = 1, \dots, T)$$

- An alternative hypothesis

$$H_1 : \beta_1(t) = \mathbf{Z}_t^\top \boldsymbol{\beta}_1 \text{ or } \dots \text{ or } \beta_M(t) = \mathbf{Z}_t^\top \boldsymbol{\beta}_M, (t = 1, \dots, T)$$

- The alternative can be a smooth, low-dimensional function, i.e.

$$\mathbf{Z}_t^\top \boldsymbol{\beta}_m = [1, (t-1), \dots, (t-1)^{p-1}] \boldsymbol{\beta}_m$$

with $\boldsymbol{\beta}_m = (\beta_{m1}, \dots, \beta_{mp})^\top$

- Nuisance parameter: $\alpha(t) = \mathbf{B}_t^\top \boldsymbol{\alpha}$, where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)^\top$

Parameter Estimator

- Least square estimator:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^T I_{it} [Y_{it} - \mathbf{X}_{it}^\top \boldsymbol{\theta}]^2 \right\},$$

where

- $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top)^\top$
- I_{it} = availability indicator for participant i at time t
- $E(Y_{it} | I_{it} = 1, \mathbf{A}_{it}) = \mathbf{X}_{it}^\top \boldsymbol{\theta}$

- The **consistent** and **asymptotic normal** properties:

- $\hat{\boldsymbol{\beta}} \rightarrow \tilde{\boldsymbol{\beta}}$ when $N \rightarrow \infty$
- $\sqrt{N}(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) \sim N(0, \boldsymbol{\Sigma}_{\boldsymbol{\beta}\boldsymbol{\beta}})$
- $N(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})^\top \boldsymbol{\Sigma}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1} (\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) \sim \chi_{Mp}^2$

Test Statistics

- **Test Statistics:** $\hat{C}_N(\boldsymbol{\delta}) = N \hat{\boldsymbol{\beta}}^\top \boldsymbol{\Sigma}_{\boldsymbol{\beta}\boldsymbol{\beta}} \hat{\boldsymbol{\beta}}$, where $\boldsymbol{\delta} = \boldsymbol{\beta}/\sigma$

- **Large Sample:**

- $\hat{C}_N(\boldsymbol{\delta}) \sim \chi_{Mp}^2$ under H_0
- $\hat{C}_N(\boldsymbol{\delta}) \sim \chi_{Mp, C_N(\boldsymbol{\delta})}^2$ under H_1

- **Small Sample:**

- $\boldsymbol{\Sigma}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ is replaced by $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}\boldsymbol{\beta}}$
- $\hat{C}_N(\boldsymbol{\delta}) \sim T_{Mp,N}^2 = \frac{MpN}{N - Mp + 1} F_{Mp, N - Mp + 1}$ under H_0

$$\Pr(F_{Mp, N - Mp + 1} > f_{Mp, N - Mp + 1, \alpha}) = \alpha$$

- $\hat{C}_N(\boldsymbol{\delta}) \sim T_{Mp, N, C_N(\boldsymbol{\delta})}^2 = \frac{MpN}{N - Mp + 1} F_{Mp, N - Mp + 1, C_N(\boldsymbol{\delta})}$ under H_1 , i.e.

$$\Pr(F_{Mp, N - Mp + 1, C_N(\boldsymbol{\delta})} > f_{Mp, N - Mp + 1, \alpha}) = \text{Power}$$

Input Parameters for Sample Size Calculation

- For power-based calculation:
 - T (number of decision time points)
 - Effect size
 - Initial standardized effect
 - Average standardized effect
 - Trend of standardized effect over time
 - Desired power (80%)
 - Significance level (5%)
- For precision-based calculation: desired power can be replaced by desired precision (margin of error and confidence interval)

Outline

1 Mobile Health

2 DIAMANTE study

3 Multi-Level Micro-Randomized Trial Design

4 Analysis Plan and Sample Size Calculation

5 Simulation

6 Discussion

Simulation Study

- Power: 80%
- Level of significance: 5%
- Number of decision time points: one per day
- Standardized proximal effect size: small
- Trend of standardized proximal effect over time: increase until 28-day, then constant for the rest
- Availability at each decision time point: 100%

Simulation Result

Table: The significance level is 0.05. The desired power (P) is 0.80. All interventional messages are proposed initially.

M	T	Sample Size		Theoretical P		Monte Carlo P		
		0.20	0.10	0.20	0.10	0.20	0.10	
3	180	15	37	0.85	0.80	0.84	0.79	
	84	22	70	0.80	0.81	0.80	0.80	
	28	46	172	0.80	0.80	0.79	0.79	
	14	84	332	0.80	0.80	0.79	0.79	
4	180	17	42	0.82	0.81	0.80	0.80	
	84	26	77	0.82	0.80	0.82	0.80	
	28	52	190	0.81	0.80	0.80	0.80	
	14	94	367	0.81	0.80	0.79	0.80	

Simulation Result

Table: The significance level is 0.05. The desired coverage probability (CP) is 0.95. All interventional messages are proposed initially.

M	T	Sample Size		Theoretical CP		Monte Carlo CP	
		0.25	0.15	0.25	0.15	0.25	0.15
3	180	13	22	0.96	0.96	0.95	0.97
	84	18	35	0.95	0.95	0.96	0.95
	28	32	75	0.95	0.95	0.96	0.96
	14	55	139	0.95	0.95	0.95	0.95
4	180	17	27	0.97	0.96	0.95	0.96
	84	23	43	0.96	0.95	0.96	0.95
	28	40	93	0.95	0.95	0.95	0.94
	14	68	171	0.95	0.95	0.96	0.95

Simulation Result

Table: The significance level is 0.05. The desired power (P) is 0.80. Two interventional messages are proposed initially and rest are proposed halfway through the study period.

M	T	Sample Size		Theoretical P		Monte Carlo P	
		Average	proximal	standardised	effect	0.10	0.20
3	180	15	38	0.84	0.81	0.82	0.80
	84	23	70	0.83	0.80	0.82	0.80
	28	51	189	0.81	0.80	0.80	0.78
	14	93	366	0.81	0.80	0.79	0.79
4	180	17	42	0.81	0.80	0.81	0.80
	84	26	79	0.81	0.80	0.79	0.80
	28	60	222	0.81	0.80	0.79	0.80
	14	109	429	0.80	0.80	0.78	0.79

Simulation Result

Table: The significance level is 0.05. The desired coverage probability (CP) is 0.95. Two interventional messages are proposed initially and rest are proposed halfway through the study period.

M	T	Sample Size		Theoretical CP		Monte Carlo CP	
		Margin	error of average proximal standardised effect	0.25	0.15	0.25	0.15
3	180	13	22	0.96	0.96	0.95	0.96
	84	18	35	0.95	0.95	0.95	0.96
	28	35	83	0.95	0.95	0.95	0.95
	14	60	153	0.95	0.95	0.95	0.96
4	180	17	27	0.97	0.95	0.96	0.97
	84	23	44	0.96	0.95	0.96	0.97
	28	46	108	0.96	0.95	0.96	0.96
	14	78	200	0.95	0.95	0.95	0.96

Outline

1 Mobile Health

2 DIAMANTE study

3 Multi-Level Micro-Randomized Trial Design

4 Analysis Plan and Sample Size Calculation

5 Simulation

6 Discussion

Conclusion

- Mobile/digital interventions can potentially reduce health disparity, and thus appealing from a global health perspective!
- Multi-Level Micro-Randomized Trial is a cutting-edge trial design in mobile/digital health, i.e.
 - very relevant for behavioral, lifestyle interventions,
 - extension of factorial designs in sequential decision-making context,
 - GEE-type longitudinal data analysis,
 - sample size formulas are available.

Future Plan

- 2020-2021:
 - Run additional simulation studies, e.g., outcome with zero inflated distributions
 - Extend the MLMRT design to consider, e.g., interactions between time window and motivational message
 - Complete the manuscript and aim to submit it to, e.g., Statistics in Medicine or Annal of Applied Statistics
 - Complete the R package “MLMRT” and the corresponding R shiny app
- 2020-2023:
 - Continue to contribute statistical support (i.e. analysis) for the DIAMANTE study
 - Develop novel statistical methods based on analysing DIAMANTE data
 - Complete a manuscript for the proposed analysis method and submit it to a high rank statistical journal
 - Complete a manuscript for the R package “MLMRT” and aim to submit it to Journal of Statistical Software or R Journal
- 2020-2025:
 - In the similar way, look for collaborative opportunities (China, USA or Australia, e.g., NDARC UNSW) in other mHealth domain, e.g., alcohol or tobacco cessation

Acknowledgement

Financial Support:

- Duke-NUS Medical School, Singapore
- Institute of Mathematical Sciences, NUS, Singapore
- Ministry of Education (MOE), Singapore
- National Institutes of Health (NIH), USA

Questions, comments, criticisms, request for slides to: kenny.xu@duke-nus.edu.sg



thank
you