ウェブサイト公開用

（様式 10）
（海外特別研究員事業）
令和元年 10 月 26 日

# 海外特別研究員最終報告書

独立行政法人 日本学術振興会 理事長　殿

採用年度　平成 30 年度
受付番号　201860006
氏　　名　江島啓介

（氏名は必ず自署すること）

海外特別研究員としての派遣期間を終了しましたので、下記のとおり報告いたします。
なお、下記及び別紙記載の内容については相違ありません。

記

1. 用務地（派遣先国名）用務地：ブルーミントン（国名：アメリカ合衆国）

2. 研究課題名（和文）人工知能を応用した体重変遷パターンと死亡リスクの関係評価及び肥満数理疫学の確立

3. 派遣期間：平成 30 年 4 月 1 日　～　令和元年 9 月 24 日

4. 受入機関名及び部局名
インディアナ大学　公衆衛生大学院

5. 所期の目的の遂行状況及び成果…書式任意　**書式任意（A4 判相当 3 ページ以上、英語で記入も可）**
（研究・調査実施状況及びその成果の発表・関係学会への参加状況等）
（注）「6. 研究発表」以降については様式 10－別紙 1～4 に記入の上、併せて提出すること。

BMI is easily calculated but still used to determine appropriate treatment in clinical settings and risk assessment in epidemiological studies. Recent studies suggested the importance of BMI history, not only current BMI status to predict obesity-related morbidity and mortality. Even one's BMI is categorized as 'normal weight', if the one experienced 'obesity' before, the risk is higher than that of ones with normal weight who never experienced obesity. However, the association between BMI history and morbidity/mortality risks has not been well investigated.

During this fellowship, I challenged this topic (1. Estimate the association between BMI trajectory and mortality) and also worked on two projects stemmed from the main project: 2. To test whether exclusion of extreme reporters of energy intake (the 'Goldberg cutoffs') reliably make results less biased in nutrition studies, 3. Propose novel sample size calculation framework with plasmode approach. Followings are summary of each project.

**1. Estimate the association between BMI trajectory and mortality**

BMI trajectory is associated with mortality or disease risk, which was suggested by previous studies based on data from only a few time points. I analyzed longitudinal data (more time points) with the latent class linear mixed model. Following is the detail:

> ***Background:*** *Obesity is associated with mortality. However, the association has been assessed focusing on a single time point BMI and mortality. Given that women's weight changes after menopausal due to change in hormone balance or chronic diseases, the BMI trajectory should be taken into account to nuance the association. I examine the association between weight*

*change and morality risk of postmenopausal population.* ***Subjects/Methods:*** *We used longitudinal weight history data of postmenopausal women from Women's Health Initiative (WHI) for our analysis. The participants whose baseline BMI is above 18.5kg/m$^2$ and who have reported at least 5 time points of BMI were included for the analysis. The participants whose BMI changed extreme (more than 20 units) in the study period were also excluded. We categorized BMI histories into five groups using growth mixture model (latent class linear mixed model), in which linear and quadratic terms of the time since the enrolment were considered as class-specific fixed effects. Subsequently we performed the survival analysis using cox proportional hazard model, in which baseline BMI categories (normal weight [18.5 to 25.0], overweight [25.0 to 30.0], class I obese [30.0 to 35.0] and class II obese [35.0 and above]) and weight trajectory groups and their interaction were used as main predictors. The model was adjusted by age groups (under 55, 55 to 60, 60 to 65, 65 to 70 and over 70) and race/ethnicity (White, Black and the others).* ***Results:*** *We classified BMI trajectory into 5 groups and named them as A) lose more weight (5 years weight change=-6.19 kg/m2; 95% CI=-6.63 to -5.76), B) lose weight (-1.97 kg/m2; -2.22 to -1.72), C) stable weight (0.21 kg/m2; -0.09 to 0.51), D) gain weight (2.1 kg/m2; 1.34 to 2.86) and E) gain more weight (5.36 kg/m2; 5.14 to 5.57). Using stable weight as a reference, we identified lose weight, lose more weight, gain more weight is statistically significantly associated with high mortality risk regardless of baseline BMI category.* ***Conclusions:*** *Losing weight and gaining more weight is associated with high mortality risk in postmenopausal women regardless of the baseline BMI.*

Although we could confirm taking into account BMI trajectory is important, the detected pattern is relatively simple. Also, I did not consider other health status (eg., blood pressure) or biomarker history. As a next step, I have to use AI to detect more flexible trajectory patterns.

**2. To test whether exclusion of extreme reporters of energy intake (the 'Goldberg cutoffs') reliably make results less biased in nutrition studies**

Energy intake is measured by various way. Two most commonly used method is self-report and doubly-labelled water (DLW) method. Each of these has inherent pros and cons. Especially, self-report has empirically shown to be biased. To mitigate the bias due to self-report, excluding extreme reporters has been proposed and widely accepted as useful method. However, there is not mathematical or empirical proof that the method can mitigate or eliminate bias. We have used three different data sources and examined whether removing extreme reporters from analysis could reliably produce less biased results. Following is the detail:

***Background:*** *The Goldberg cutoffs are used to decrease bias in self-reported estimates of energy intake ($EI_{SR}$). Whether the cutoffs decrease bias compared with objective measurements when used in regressions of health outcomes has not been assessed.* ***Objective:*** *We examined whether applying the Goldberg cutoffs to data used in nutrition studies could reliably produce less biased results.* ***Design:*** *We used data from the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE), the Interactive Diet and Activity Tracking in AARP (IDATA) Study, and the National Diet and Nutrition Survey (NDNS). Each dataset included $EI_{SR}$, EI estimated from doubly labeled water ($EI_{DLW}$) as a reference method, and health*

*outcomes including baseline anthropometric, biomarker, and behavioral measures and fitness test results. We conducted 3 linear regression analyses using $EI_{SR}$, $EI_G$ (plausible $EI_{SR}$ based on the Goldberg cutoffs), and $EI_{DLW}$ as an explanatory variable for each analysis. Regression coefficients were denoted $\hat{\beta}_{SR}$, $\hat{\beta}_{G,}$ and $\hat{\beta}_{DLW}$, respectively. Using the jackknife method, bias from $\hat{\beta}_{SR}$ compared to $\hat{\beta}_{DLW}$ and remaining bias from $\hat{\beta}_G$ compared to $\hat{\beta}_{DLW}$ were estimated. Analyses were repeated using Pearson's correlation coefficients.* ***Results:*** *Among those variables with significant bias, using $EI_G$ instead of $EI_{SR}$ significantly decreased the bias for weight (56.1%; 95% CI: 28.5, 83.7) and waist circumference (59.8%; 95% CI: 33.2, 86.5) with CALERIE; for weight (20.8%; 95% CI: -6.4, 48.1) and waist circumference (17.3%; 95% CI: -20.8, 55.4) with IDATA; and for waist circumference (-9.5%; 95% CI: -72.2, 53.1) with NDNS. The reduction in bias was not statistically significant for the other outcomes. Results obtained with Pearson's correlation coefficient analyses were qualitatively consistent.* ***Conclusions:*** *Some, but not all, associations between $EI_G$ and outcomes were biased compared with $EI_{DLW}$. Use of the Goldberg cutoffs was not a reliable method for reducing bias.*

In literature search, we also found the Goldberg cutoffs are used to exclude extreme reporters not only in terms of energy intake, but also dietary information. We will apply this approach to test whether the Goldberg cutoff reliably reduce reporting bias of dietary consumption (such as sodium and potassium). The results of this study were presented at Obesityweek at Nashville (USA) in 2018, and published in American Journal of Clinical Nutrition in 2019.

**3. Propose novel sample size calculation framework with plasmode approach. Followings are summary of each project**

Murine (i.e., mouse and rat) is common animal model used for preclinical research. Due to technical and financial limitation as well as for animal welfare, using minimal number of animals is recommended. To compute sample size, usually power calculation is conducted, however, there are several assumptions behind it. For example, data distribution is assumed to follow normal distribution and the variances are assumed to be identical between control and treatment groups. However, most of the data may not meet these assumptions. We proposed a plasmode approach, in which the population distribution is identical to empirical distribution. Following is the detail:

***Background:*** *Genetic obesity is frequently studied in murine (i.e., mouse and rat) animal models to elucidate potential explanations, causes, and correlates of obesity in humans and other animals. However, these studies often use small sample sizes, and the data may violate some of the assumptions of common statistical tests. Plasmode-based simulation using empirical animal data can provide more real-world answers to whether and to what extent these factors affect type I error rates and statistical power.* ***Objective:*** *To evaluate type I error rates and statistical power of commonly used statistical analyses with small ($n \leq 5$) to moderate sample sizes, by using plasmode-based simulation from existing weight data from murine genetic models of obesity.* ***Methods:*** *We compared 7-11 week old weight data from five distinct, homozygous, monogenic, murine models of obesity with non-mutant controls of both sexes. To examine whether type I error rates could be affected by choice of statistical tests, we adjusted the empirical distributions of weights to ensure the null hypothesis (i.e., no mean difference) in*

*two ways: Case 1) center weight distributions of both groups on the same mean weight; Case 2) combine data from control and mutant groups into one distribution. From these cases, 3 to 20 mice were resampled to create a 'plasmode' dataset. We performed five common tests (Student's t-test, Welch's t-test, Wilcoxon test, permutation test and bootstrap test) on the plasmode datasets and computed type I error rates. Power was assessed using plasmode datasets in which the mean of the control group was shifted by adding a constant value as in Case 1 to create different effect sizes while keeping other aspects of the data distributions (e.g., variance) intact.* ***Results:*** *For Case 1, Type I error rates were significantly higher than the nominal significance level of 0.05, which we name 'type I error rate inflation', for Student's t-test, Welch's t-test and permutation test, especially when sample size was small, whereas conservative error rates were noted for bootstrap with small samples. Because of heterogeneity of variance, Wilcoxon test was generally inappropriate because it tests for differences in distribution, not just differences in central tendency. For Case 2, consistent type I error inflation was observed only for permutation test with small samples. In both Cases, increasing sample size mitigated inflation and deflation. Patterns were generally similar when significance was set to 0.01, 0.005, and 0.001, with a couple marked differences. Power was markedly higher or lower from the theoretical power, particularly with small samples. Compared with the other tests, bootstrap was underpowered with small samples as a tradeoff for maintaining type I error rates, and permutation was overpowered in conjunction with inflated type I error rates.* ***Conclusions:*** *With small samples ($n \leq 5$), bootstrap avoided type I error rate inflation, but often at the cost of lower power. To avoid type I error rate inflation for other tests, sample size should be increased. Wilcoxon should be avoided because of heterogeneity of weight distributions between mutant and control mice.*

Overall, our study suggested bootstrap test is the best statistical test in terms of type I error rate, although there is an issue of power. However the conclusion is reasonable only for this specific case we studied (i.e., murine genetic model of obesity). To make our claim more general, and also these approach more feasible for non-statisticians, we are planning to develop a software for sample size calculation. The results of this study were presented at American Society for Nutrition at Boltimore (USA) in 2019, and the manuscript was submitted.