探討生活習慣與飲食對SGOT和SGPT數值之影響

Assessing the Effects of Living Habits and Diets on SGOT and SGPT Levels

Student: Che-Wei Tsai / Supervisors: Shu-Chun Chen Ph.D & Prof. Charlotte Wang

Internship Department: Academia Sinica Data Science Statistical Cooperation Center

Introduction

Liver is a vital organ in the human body, which plays a crucial role in various biochemical and physiological functions. Due to the limited distribution of nerves in liver, liver diseases are challenging to detect in early phase unless severe pathological changes occur. SGOT (Serum Glutamic Oxaloacetic Transaminase) and SGPT (Serum Glutamic Pyruvic Transaminase) are enzymes primarily found in liver cells. They serve as commonly used biochemical indicators to assess liver function and health. The main purpose of this study is to utilize conditional (subject-specific) Mixed Effects Models (MEM) and marginal (population-specific) Generalized Estimating Equations (GEE) to investigate the effects of living habits and diets on SGOT and SGPT levels. Through the identification of risk factors, it is possible to prevent the onset of severe liver diseases.

Materials and Methods

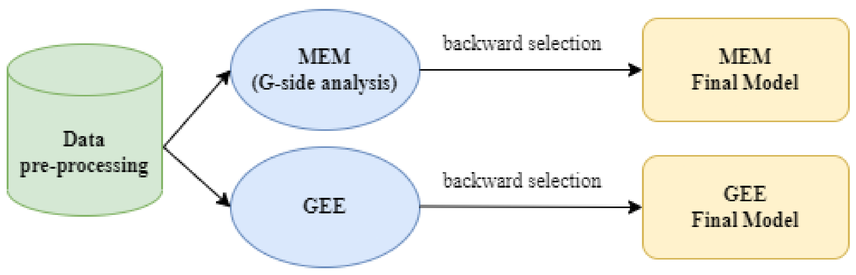

Figure 1: Project Scheme

Data

A total of 9452 samples from MJ Health Screening Center (美兆健康管理機構) underwent follow-up assessments at least three times between 2005 and 2014 (once every three years). After removing outliers based on the criterion of 3 times the interquartile range (IQR), the final sample size for building SGOT and SGPT estimation models are 9012 and 8901 respectively.

The explanatory variables related to living habits include alcohol consumption, smoking, exercise frequency [5], and sleep duration [3]; Meanwhile, the variables related to diets include vegetables [4]; processed foods [2]; sugary beverages [6]; and frying, deep-frying, grilling, and smoking foods (FDGS) [1]. All variables are coded as 0, 1, and 2 according to their respective levels in the original questionnaire. Although long-term medication doesn't belong to either the lifestyle or dietary category, it is still incorporated into the model.

Both chronic disease (hypertension, diabetes) and liver disease (cirrhosis, hepatitis, liver cancer) are regarded as potential confounding factors in the association between long-term medication and the response variables. Alcohol consumption and smoking can potentially function as confounding factors between each other and the response variables as well. Additionally, the model is adjusted for age and gender.

Mixed effects model (MEM)

$$y = X\beta + Z\gamma + \varepsilon,$$
$$\text{where } \gamma \sim \mathcal{N}(0, G),\ \varepsilon \sim \mathcal{N}(0, R)$$

MEM is particularly useful when dealing with repeated measurements data, which can captures the variations between different individuals by adding random effects (γ). This variation provides a mechanism for modeling individual-specific responses over a period of time. In this study, given the right-skewed distributions of SGOT and SGPT, applying a log-transform to the response variables is necessary to satisfy the assumptions of MEM.

Generalized estimating equations (GEE)

$$g(\mu_i) = X_i\beta,\ U(\beta) = \sum_{i=1}^{N} \frac{\partial \mu_i}{\partial \beta} V_i^{-1}\{Y_i - \mu_i(\beta)\}$$

GEE is also a statistical approach suitable for handling repeated measurements data. Different from MEM, GEE models the mean response by employing a linear function of the covariates using a transformation or link function. It only utilizes the first two moments (mean(μ), variance(V)) of the data to estimate the beta coefficients in the model, which presents a less restrictive assumption. In GEE, the response variables are log-transformed as well. The chosen working correlation matrix is the unstructured correlation matrix, determined by the smallest quasi information criterion (QIC).

Table 1: MEM v.s. GEE

	Mixed Effects Model	Generalized Estimating Equations
Focus	Individual-level effects	Population-averaged effects
Assumption	Normality for γ and ε	Less restrictive assumptions on data distribution
Usage	Personalized health suggestion	Population health suggestion

Results

Based on the findings presented in Table 2, alcohol consumption, long-term medication, FDGS, and exercise frequency are identified as risk factors associated with elevated SGOT levels in the MEM, while smoking and drinking sugary beverages exhibit negative effects on elevated SGOT levels. Similar to MEM, GEE yielded nearly identical results, except that long-term medication is not retained in the final model (not statistically significant). It is worth noting that SGOT are log-transformed initially.

Table 2: Estimated effects of significant variables in the final SGOT estimation model after backward selection (adjusted for age, gender, and confounding factors)

	MEM			GEE		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Alcohol consumption	0.01514	0.003991	0.0001	0.0154	0.0041	0.0002
Smoking	-0.00956	0.002787	0.0006	-0.0096	0.0031	0.002
Exercise frequency	0.01104	0.002087	<0.0001	0.011	0.0022	<0.0001
Sugary beverages	-0.0058	0.001771	0.0011	-0.0061	0.0018	0.0005
FDGS	0.006543	0.0031	0.0348	0.0069	0.0032	0.0296
Long-term medication	0.00686	0.003359	0.0411	-	-	-

Referring to Table 3, long-term medication is identified as a risk factor in both models, whereas processed foods and FDGS are recognized as a risk factor in MEM and GEE respectively. In addition, alcohol consumption, sleep duration, and exercise frequency act as protective factors against elevated SGPT levels. SGPT are log-transformed at first as well.

Table 3: Estimated effects of significant variables in the final SGPT estimation model after backward selection (adjusted for age, gender, and confounding factors)

	MEM			GEE		
	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Alcohol consumption	-0.02155	0.006722	0.0013	-0.0212	0.0069	0.0022
Sleep duration	-0.01195	0.004733	0.0116	-0.0115	0.0048	0.0158
Exercise frequency	-0.01078	0.00346	0.0018	-0.0114	0.0036	0.0015
Processed foods	0.01312	0.004851	0.0068	-	-	-
FDGS	-	-	-	0.0107	0.0054	0.0458
Long-term medication	0.01305	0.005577	0.0193	0.0126	0.0057	0.0281

Discussion and Conclusion

Most of the results are consistent with previous studies, except for exercise frequency and sugary beverages in relation to SGOT, and alcohol consumption in relation to SGPT. In fact, SGOT is not exclusively related to liver , it can also be elevated by muscle injury, such as intense exercise or accidents. Although the effect of sugary beverages on SGOT did not align with our expectations, its effect size was too small that it can be negligible. As for alcohol consumption, it's important to note that the relationship between alcohol consumption and SGPT levels is complex and can vary based on factors such as individual susceptibility, overall liver health, genetics, etc.

According to the results above, it is advisable to recommend patients to reduce alcohol consumption; limit the intake of frying, deep-frying, grilling, smoking, and processed foods; and minimize unnecessary medication use in order to prevent liver damage and inflammation, while prolong the sleep duration and increase exercise frequency are also good for liver health. With MEM, we can additionally utilize random effects to identify high and low-risk samples, enabling the precision of personalized health suggestions.

There exist some limitations about the data in this study. All the explanatory variables were derived from questionnaires, which introduces the potential for several biases, including response bias, recall bias, and recency bias. It means that the validity and reliability of the finding may be questioned. Furthermore, it is important to consider that the data sourced from a private health screening center may have limitations in terms of its extrapolation to the broader Taiwanese population. We did the chi-square homogeneity test to compare the distribution of age, gender, family income, and education between MJ data and general Taiwanese population. All of them reached statistical significance, which means that the results of this study are only applicable to the sampled population and cannot be generalized to the entire Taiwanese population.

References

[1] E. Costello et al. Exposure to per- and polyfluoroalkyl substances and markers of liver injury: A systematic review and meta-analysis. volume 130(4), page 46001, 2022.

[2] A. E. Henney et al. Ultra-processed food intake is associated with non-alcoholic fatty liver disease in adults: A systematic review and meta-analysis. *Nutrients*, 15(10):2266, 2023.

[3] Y. Ilan et al. Prolonged sleep-deprivation induced disturbed liver functions serum lipid levels, and hyperphosphatemia. volume 22(11), pages 740–743, 1992.

[4] M. Mollahosseini et al. The association between fruit and vegetable intake and liver enzymes (aspartate and alanine transaminases) in tehran, iran. *Ethiopian journal of health sciences*, 27(4):401–410, 2017.

[5] U. Nivukoski et al. Impacts of unfavourable lifestyle factors on biomarkers of liver function, inflammation and lipid status. *PLOS ONE*, 14(6):1–15, 06 2019.

[6] M. K. Shimony et al. The relationship between sugar-sweetened beverages and liver enzymes among healthy premenopausal women: a prospective cohort study. *European journal of nutrition*, 55(2):569–576, 2016.