http://www.cjmb.org

Open Access Original Article

Crescent Journal of Medical and Biological Sciences
Vol. 2, No. 3, July 2015, 90–94
eISSN 2148-9696

Assessment of the Effect of Gene Polymorphisms and Environmental Risk Factors on Low HDL Over Time: Tehran Lipid and Glucose Study

Yadollah Mehrabi[1], Parvin Sarbakhsh[2*], Jeanine J. Houwing-Duistermaat[3], Farid Zayeri[4], Maryam Sadat Daneshpour[5]

Abstract

Objective: Due to existing association between high-density lipoprotein (HDL) and cardiovascular disease, detection of factors affecting this lipid is important. Environmental factors and genetic variations have an important role in HDL level. The effects of these risk factors can be time-dependent; so, study of their effects on HDL level over time is important. In this study, we used transition model to analyze binary longitudinal data to investigate single nucleotide polymorphism (SNP) and other risk factors affecting low HDL over time.

Materials and Methods: Data of 329 participants of 3 phases of Tehran Lipid and Glucose Study (TLGS) was analyzed using marginal transition model. This model has a formulation which allows first and second order Markov dependence to take into account the correlation among successive observations of the same individual in longitudinal binary response for which the marginal probability of success is modelled via a form of logistic regression.

Results: Results of first order transition model showed that the odds ratio (OR) for having low HDL in women compared to men was 1.54 (95% CI: 1.02, 2.24). High waist circumference (OR=1.67, CI 95%: 1.16, 2.39), high blood pressure (OR=0.59, 95% CI: 0.41, 0.85), high triglyceride (OR=1.85, 95% CI: 1.30, 2.65) and being homozygous for the minor allele of SRB1 (OR=0.11, 95% CI: 0.01, 0.74) were significantly associated with low HDL. Also, the OR of low HDL in phase 2 of study compared to phase 1 was 1.76 (95% CI: 1.32, 2.35). The result of second order transition model was fairly similar to first order. The parameter estimates of serial dependency are markedly significant, pointing clearly to a first and second-order serial dependence ($P<.001$).

Conclusion: Considering the identification of genetic and environmental factors affecting low HDL over time, transition model was used and the most important risk factors were identified.

Keywords: Low HDL, Markov chains, Transition regression

Introduction

Studies have shown an association between high density lipoprotein-cholesterol (HDL-C) concentrations and cardiovascular disease (1). Risk factor evaluation in order to reduce low HDL rate in the community is one solution to protect people from cardiovascular diseases.

In addition to environmental factors such as demographic characteristics and biochemical indexes, genetic variations have an important role in HDL-C level (2,3). Studies show that adenosine triphosphate-binding cassette A1 (ABCA1), apolipoprotein (Apo)A1M1, ApoA1M2, ApoB, ApoAIV, Apo CIII, SRB1, and ApoE gene polymorphisms have been associated with HDL-C concentrations (4,5).

Also, since effect of these risk factors can be time-dependent, investigation of their effects associated with HDL in a longitudinal frame work could be valuable. The present study was initially motivated by the longitudinal low HDL dataset with potential effective single nucleotide polymorphisms (SNPs) and risk factors.

In longitudinal studies, repeated measurements of subjects over time were recorded. In this situation the correlation between subjects is important and if it is neglected, estimation and results will be misleading.

There are different approaches for modeling longitudinal

Received 12 March 2014, **Accepted** 25 May 2015, **Available online** 1 July 2015

[1]Department of Epidemiology, Faculty of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [2]Women's Reproductive Health Research Center, Tabriz University of Medical Sciences; and Department of Statistics and Epidemiology, School of Public Health, Tabriz University of Medical Sciences, Tabriz, Iran. [3]Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, Netherlands. [4]Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [5]Cellular and Molecular Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. *Corresponding author: Parvin Sarbakhsh, Women's Reproductive Health Research Center, Tabriz University of Medical Sciences, Department of Statistics and Epidemiology, School of Public health, Tabriz University of Medical Sciences, Tabriz, Iran.
Tel: +984133357580, Email: p.sarbakhsh@gmail.com

data. Marginal, transition and random effect models are methods of analyzing this kind of data (6).

In this paper we used a special kind of transition model - marginal modelling of binary longitudinal data using Markov chains structure (7) - to detect and assess the effects of influential SNPs and risk factors related to longitudinal low HDL. Using this transition model, this study is aimed to assess the association between HDL-C level and Apo A1M1, ApoA1M2, APOB, ApoAIV, ApoCIII, ABCA1, SRB1, ApoE gene polymorphisms, age, sex, high waist circumference (WC), high triglyceride (TG), high blood pressure (BP) and high fasting blood sugar (FBS) and smoking. Phase of study was considered as time.

Materials and Methods

Transition Model: Marginal Modelling of Binary Longitudinal Data Using Markov Chains

This model was introduced by Goncalves in 2008 (7). In this model, for notation, $Y_{it} \in (0,1)$ is binary response variables of individual i (i = 1, ... , n) at time t (t= 1, …, T_i), with mean $P(Y_{it} = 1) = \Theta_{it}$. For each subject at each time, let X_{it} be a set of p covariates that first column of its can be a vector of ones to consider intercept term. Logistic regression model that marginally connects the probability distribution of the binary response and auxiliary variables is:

$$\log it\ P(Y_{it} = 1) = X_{it}^{T}\beta$$

Where β is a p vector of unknown parameters. To take into account the correlation among successive observations of the same individual, the model considers a Markovian type of first order (Ψ_1) or of second order (Ψ_2) dependence structure (7).

We assessed effects of some SNPs and other risk factors on having low level of HDL over time using this type of transition model with first and second order dependence structure.

Data

Subjects in this longitudinal study study were selected from among participants of the Tehran Lipid and Glucose Study (TLGS). TLGS is a prospective study to determine the risk factors and outcomes of noncommunicable disease (8). The structure of this study includes three major components; phase I is a cross-sectional prevalence study of cardiovascular disease and associated risk factors and phases II and III are prospective follow-up studies with mean follow-up of three years. The TLGS design has been explained elsewhere (9). Longitudinal data from the three phases of the TLGS study was analyzed to assess the association between the some related polymorphisms and other risk factors with low levels of HDL over time

A total of 329 subjects (127 [38.6%] men and 202 [61.4%] women) who were present in phase I, II, III of TLGS study with age ≥20 years and without any missing value in evaluated variables were included in the current study.

Low HDL-C level was defined as <40 mg/dL for men and <50 mg/dL for women. High WC was defined as WC ≥95 cm for Iranian men and women (10). High TG level was defined as TG ≥150 mg/dL, subjects who had BP ≥130/85 mm Hg or used anti-hypertensive drug, and subjects with FBS ≥110 mg/dL or users of antidiabetic drugs were considered as high BP and high FBS, respectively (11). Subjects who smoke daily or occasionally were considered as smokers. Phase of study was considered as time. Each SNP was considered as a random variable taking values 0, 1, and 2 corresponding to the nucleotide pairs. We coded each of these variables into 2 dummy binary variables corresponding to a dominant and a recessive effect.

Frequency distribution and descriptive statistics such as mean, standard deviation and percentage were calculated. To compare continues and categorical variables in three phases of study, repeated measures and Cochran test were used, respectively.

The effect of polymorphisms of ApoA1M1, ApoA1M2, ApoB, ApoAIV, ApoCIII, ABCA1, SRB1 and ApoE genes on low HDL over time were investigated after adjusting for sex, age, high WC, high TG, high FBS, high BP, time and smoking using mentioned transition model (marginal modelling of binary longitudinal data using Markov

Table 1. Demographic Characteristic and Clinical and Lipid Profiles of Subjects in Phases of Study

Variables	Phase 1 (n = 329)	Phase 2 (n = 329)	Phase 3 (n = 329)	*P*
Age (y)[a]	41.09 ± 15.82	44.84 ± 15.71	47.46 ± 15.62	<.001
Sex (female)[b]	202 (61.4)	202 (61.4)	202 (61.4)	-
Low HDL-c [a]	225 (68.4)	261 (79.3)	217 (66)	<.001
High WC [a]	90 (27.4)	134 (4.7)	151 (45.9)	<.001
Hypertension [a]	111 (33.7)	97 (29.5)	82 (24.9)	.002
High TG [a]	139 (42.2)	140 (42.6)	141 (42.9)	.97
High FBS [a]	39 (11.9)	39 (11.9)	41 (12.5)	.85
Smoker [a]	27 (8.2)	30 (9.1)	27 (8.2)	.66

Entries are mean±SD for age and number (%) for the rest categorical variables.

[a] Time dependent variable.

[b] Time independent variable.

Table 2. Genotype and Allele Frequencies of Apo E, Apo A1M1, Apo A1M2, Apo B, Apo AIV, Apo CIII, and SRB1 in the Study Population

Polymorphisms			
Apo E Alleles	e2	e3	e4
	34(1.3)	258(78.4)	37(11.2)
Apo A1M1 genotypes	+/+	+/-	-/-
	233(7.8)	90(27.4)	6(1.8)
Apo A1M2 genotypes	+/+	+/-	-/-
	300(91.2)	23(7)	6(1.8)
Apo B genotypes	X^+X^+	X^+X^-	X^-X^-
	28(8.5)	126(38.3)	175(53.2)
Apo AIV genotypes	TT	GT	GG
	1(0.3)	56(17)	272(82.7)
Apo CIII genotypes	CC	CG	GG
	232(7.5)	87(26.4)	10(3)
ABC A1genotypes	GG	GA	AA
	112(34)	171(52)	46(14)
SRB1genotypes	GG	GA	AA
	268(81.5)	58(17.6)	3(0.9)

chains) with first and second order dependence structure. Statistical analysis was done through R (3.1.0) software and "The R Package bild for the Analysis of Binary Longitudinal Data" (12).

Results

Table 1 pictures the summary of demographic characteristic and clinical and lipid profiles of these subjects in three phases of study. Highest prevalence of having low HDL (79.3%) was seen in phase 2 of study. Allele frequencies given in Table 2 show genotype distributions. The +/+ genotype of Apo A1M2 gene had the highest prevalence (91.2%) and TT genotype of Apo AIV gene had the lowest frequency (0.3%).

The results of transition model with first order Markov chain (Table 3) show that the odds ratio (OR) for having low HDL in women compared to men was 1.54 (95% CI: 1.02, 2.24). The effect of high WC (OR = 1.67, $P = .005$) and high TG (OR = 1.85, $P < .001$) on low HDL were significant with an increased OR. Also, high BP (OR = 0.59, 95% CI: 0.41, 0.85), and being homozygous for the minor allele of SRB1 (OR = 0.11, 95% CI: 0.01, 0.74) were significantly associated with low HDL. Also, the odds of low HDL in phase 2 of study was 1.76 (95% CI: 1.32, 2.35) times more

Table 3. Results of Transition Models to Study Risk Factors Affecting Low HDL Over Time

Variable	**First Order Transition Model**		**Second Order Transition Model**	
	OR (95% CI)	***P***	**OR (95% CI)**	***P***
Sex (women compared to men)	1.51 (1.02, 2.24)	.04	1.58 (1.04, 2.239)	.03
Age	1 .00(0.98, 1.01)	.59	0.99 (0.98, 1.01)	.62
Smoking	0.89 (0.48, 1.63)	.71	1.003 (0.54, 1.84)	.99
High WC	1.67 (1.16, 2.39)	.005	1.57 (1.10, 2.23)	.01
High FBS	1.27 (0.7, 2.28)	.43	1.33 (0.73, 2.43)	.34
High BP	0.59 (0.41, 0.85)	.004	0.6 (0.42, 0.86)	.005
High TG	1.85 (1.3, 2.65)	<.001	1.76 (1.24, 2.50)	<.001
Dominant ApoA1M1	0.82 (0.53, 1.25)	.35	0.84 (0.53, 1.33)	.46
Recessive ApoA1M1	1.64 (0.36, 7.42)	.52	1.67 (0.31, 9.1)	.55
Dominant ApoA1M2	0.54 (0.27, 1.08)	.08	0.56 (0.26, 1.19)	.13
Recessive ApoA1M2	1.85 (0.42, 8.25)	.42	1.5 (0.3, 7.57)	.62
Dominant APOB	0.9 (0.61, 1.34)	.6	0.87 (0.57, 1.32)	.51
Recessive APOB	1.14 (0.57, 2.29)	.71	1.24 (0.58, 2.63)	.57
Dominant ApoCIII	0.94 (0.61, 1.44)	.77	0.88 (0.56, 1.39)	.58
Recessive ApoCIII	2.02 (.55, 7.38)	.29	1.97 (0.48, 8.18)	.34
Dominant ABCA1	0.93 (0.62, 1.41)	.75	0.97 (0.63, 1.5)	.88
Recessive ABCA1	1.15 (0.65, 2.05)	.63	1.12 (0.6, 2.07)	.72
Dominant SRB1	0.69 (0.43, 1.13)	.14	0.64 (0.38, 1.07)	.09
Recessive SRB1	0.11 (0.01, 0.74)	.02	0.14 (0.02, 1.01)	.05
ApoE: e2 or e4 compared to e3	0.69 (0.44, 1.08)	.1	0.7 (0.43, 1.12)	.13
Phase 2 compared to phase 1	1.76 (1.32, 2.35)	<.001	1.73 (1.29, 2.31)	<.001
Phase 3 compared to phase 1	0.79 (0.57, 1.10)	.16	0.79 (0.59, 1.05)	.11

than phase 1. The first order Markov chain dependence between adjacent observations of response was estimated 2.27 indicating the strong dependence between successive observations ($P < .001$).

The result of second order transition model was fairly similar to the result of the first order model (Table 3). The parameter estimates of serial dependency were markedly significant, pointing clearly to a first ($\Psi_1 = 2.24$, $P < .001$) and second order serial dependence ($\Psi_2 = 1.58$, $P < .001$.

Discussion

Longitudinal studies have been increased in medical research in recent years. An issue that often arises in analysis of longitudinal data is correlation among successive observations of the same individual. Therefore using longitudinal models is common as an approach for taking to account this correlation. For analyzing these types of data 3 methods including transition models, marginal models and random effect models are common in longitudinal studies (6).

In this study each patient was a cluster and the repeated measures of them in three phases of study were correlated. So, for analyzing this data, we used a kind of transition model for marginal modeling of binary longitudinal response which is able to capture correlations between longitudinal correlated responses and allows for inference about risk factors affecting phenotypes over time. Mentioned models were applied to the data from TLGS study and some SNPs and risk factors related to low HDL were identified.

Similar to other studies (2,3,13-15), in our study TG, WC, BP and SRB1 were significantly associated with low HDL, too. Results of first order model show that high TG and high WC can increase the odd of having low HDL during time and consequently increase the risk of cardiovascular disease. Also, people with SRB1=AA compare to other people have lower odd for having low HDL (OR = 0.38).

On the other hand time had a significant positive effect on rate of low HDL. This means that with increasing time the rate of low HDL increased. It could be related with change in people lifestyle such as nutrition and physical activity due to technology. In addition to regression parameter in this model, parameters of the first and second order serial dependence were estimated. The estimated parameters were strongly significant indicated the longitudinal correlation between the subjects.

The results of first and second order Markov chain were fairly similar to each other but Akaike information criterion (AIC) of transition model with second order dependency (1003.06) were less than the model with first order (1027.16) so it can be concluded that second order model is able to fit the data better than first order.

Conclusion

Considering the identification of genetic and environmental factors affecting low HDL over time, marginal transition model was used and the most important risk factors were identified.

Ethical issues

This study is an observational research was approved by the research ethics committee of Shahid Beheshti University of Medical Sciences, Iran.

Conflict of interests

The authors declare no conflict of interest.

Acknowledgments

We would like to thank the staff and participants in the TLGS study for data collection. The research presented in this paper was carried out on the High Performance Computing Cluster supported by the computer science department of Institute for Research in Fundamental Sciences (IPM). This article has been extracted from PhD thesis of Parvin Sarbakhsh in Biostatistics at School of Para-medicine at Shahid Beheshti University of Medical Sciences.

References

1. Toth PP. Cardiology patient page. The "good cholesterol": high-density lipoprotein. Circulation 2005;111(5):e89-e91.
2. Greenow K, Pearce NJ, Ramji DP. The key role of apolipoprotein E in atherosclerosis. J Mol Med (Berl) 2005;83(5):329-42.
3. Schaefer EJ. Lipoproteins, nutrition, and heart disease. Am J Clin Nutr 2002;75(2):191-212.
4. Brown CM, Rea TJ, Hamon SC, et al. The contribution of individual and pairwise combinations of SNPs in the APOA1 and APOC3 genes to interindividual HDL-C variability. J Mol Med (Berl) 2006;84(7):561-572.
5. Frikke-Schmidt R. Context-dependent and invariant associations between APOE genotype and levels of lipoproteins and risk of ischemic heart disease: a review. Scand J Clin Lab Invest Suppl 2000;233:3-25.
6. Diggle PJ HP, Liang KY, Zeger SL. Analysis of Longitudinal Data. Oxford: Oxford University Press; 2002.
7. Goncalves MH, Azzalini A. Using markov chains for marginal modelling of binary longitudinal data in an exact likelihood approach. Metron 2008;LXVI(2):157-181.
8. Azizi F, Rahmani M, Emami H, et al. Cardiovascular risk factors in an Iranian urban population: Tehran Lipid and Glucose Study (Phase 1). Soz Praventivmed2002;47(6):408-26.
9. Azizi F, Ghanbarian A, Momenan AA, et al. Prevention of non-communicable disease in a population in nutrition transition: Tehran Lipid and Glucose Study phase II. Trials 2009;10:5. doi: 10.1186/1745-6215-10-5.
10. Azizi F, Khalili D, Aghajani H, et al. Appropriate waist

circumference cut-off points among Iranian adults: the first report of the Iranian National Committee of Obesity. Arch Iran Med 2010;13(3):243-4.
11. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation 2005;112(17):2735-2752.
12. Gonçalves MH, Cabral MS, Azzalini A. The R package bild for the analysis of binary longitudinal data. J Statistical Software 2012;46(9):1-17.
13. Daneshpour MS, Hedayati M, Eshraghi P, Azizi F. Association of Apo E gene polymorphism with HDL level in Tehranian population. Eur J Lipid Sci Technol 2010;112(7):810-816. doi: 10.1002/ejlt.200900207.
14. Saidi S, Slamia LB, Ammou SB, Mahjoub T, Almawi WY. Association of apolipoprotein E gene polymorphism with ischemic stroke involving large-vessel disease and its relation to serum lipid levels. J Stroke Cerebrovasc Dis 2007;16(4):160-6.
15. Zaman MM, Ikemoto S, Yoshiike N, Date C, Yokoyama T, Tanaka H. Association of apolipoprotein genetic polymorphisms with plasma cholesterol in a Japanese rural population: the Shibata study. Arterioscler Thromb Vasc Biol 1997;17(12):3495-3504. doi: 10.1161/01.ATV.17.12.3495.