ORIGINAL ARTICLE

Comparison of Marginal Logistic Model with Repeated Measures and Conditional Logistic Model in Risk Factors Affecting Hypertension

Neda Gilani[1],
Anoshirvan Kazemnejad[1],
Farid Zayeri[2],
Jamshid Yazdani[3]

[1] Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran
[2] Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3] Psychiatry and Behavioral Sciences Research Center, Department of Epidemiology and Biostatistics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

(Received May 9, 2010 ; Accepted March 19, 2011)

Abstract

Background and purpose: To analyze the data in which the correlation between observations are to be considered, a general method is using marginal model with repeated measures, yet there is another method called conditional model with random clusters. According to the binary responses, the aim of the present study is to compare the efficiency of these two models in studying the risk factors affecting hypertension in Mazandaran province.

Materials and methods: This cross-sectional study is part of research studying the risk factors of the non-infectious diseases in Mazandaran province in 2007 with a sample size of 1000 cases. First, stratified sampling was employed and then random cluster sampling was used in each stratum. In all cases, blood pressure was measured three times through standard methods. The obtained data were analyzed using χ^2 test and fitting marginal and conditional logistic models.

Results: Of the total of 1000 (500 male and 500 female) cases, 38 percent had high blood pressure. In both models four variables of age, physical activity, body-mass-index, and the consumption of fruits and vegetables were identified as factors affecting hypertension. In addition, gender and fasting blood sugar were statistically significant only in marginal logistic model. Using goodness-of-fit criteria, values of all the statistics for conditional logistic model with random effects of individuals were lower in comparison with the marginal logistic model with repeated measures.

Conclusion: It can be concluded that although one cannot decisively select a model as the most appropriate one, conditional logistic model fits were somewhat better than the marginal logistic model with repeated measures, so it can be used to analyze such data.

Key words: Hypertension, marginal logistic model with repeated measures, conditional logistic model with random clusters, goodness-of-fit criteria

J Mazand Univ Med Sci 2011; 21(82): 27-35 (Persian).

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

پژوهشی

مجله دانشگاه علوم پزشکی مازندران
دوره بیست و یکم شماره ۸۲ خرداد سال ۱۳۹۰ (۳۵-۲۷)

مقایسه مدل‌های لجستیک حاشیه‌ای با اندازه‌گیری مکرر و لجستیک شرطی در بررسی عوامل موثر بر پرفشاری خون

ندا گیلانی[1]
انوشیروان کاظم نژاد[2]
فرید زایری[3]
جمشید یزدانی[4]

چکیده

سابقه و هدف : دو تا از مهمترین روش‌های تحلیل برای داده های همبسته عبارتند از استفاده از مدل حاشیه‌ای با اندازه‌گیری مکرر و دیگری بکارگیری مدل‌های شرطی با خوشه‌های تصادفی می‌باشد. با توجه به دو حالتی بودن متغیر پاسخ، هدف از مطالعه حاضر مقایسه کارایی دو مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر و لجستیک شرطی با خوشه‌های تصادفی برای بررسی عوامل خطر موثر بر بیماری پرفشاری خون در استان مازندران می‌باشد.

مواد و روش ها : این مطالعه مقطعی قسمتی از طرح بررسی عوامل خطر بیماری‌های غیر واگیردار استان مازندران در سال ۱۳۸۶ با حجم نمونه ۱۰۰۰ نفر بود. روش نمونه‌گیری بصورت طبقه‌ای و در داخل طبقات بصورت تصادفی خوشه‌ای یک مرحله‌ای بود. در تمامی افراد فشار خون سه بار به روش استاندارد اندازه‌گیری شد. داده‌های حاصل با استفاده از آزمون Chi-square و برازش مدل‌های لجستیک حاشیه‌ای و شرطی مورد تجزیه و تحلیل قرار گرفت.

یافته ها : از مجموع ۱۰۰۰ فرد (۵۰۰ مرد و ۵۰۰ زن) مورد مطالعه، ۳۸درصد دارای بیماری پرفشاری خون بوده‌اند. در هر دو مدل، متغیرهای گروه سنی، فعالیت بدنی، نمایه توده بدنی و مصرف میوه و سبزی از عوامل تاثیرگذار بر پرفشاری خون شناخته شدند. اضافه بر این متغیرهای جنسیت و قند خون ناشتا فقط در مدل لجستیک حاشیه‌ای معنی‌دار آماری گردید. با استفاده از معیارهای نیکویی برازش، مقادیر تمامی آماره‌ها برای مدل لجستیک شرطی با اثر تصادفی افراد نسبت به مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر، کمتر بدست آمد.

استنتاج : اگرچه نمی‌توان با قاطعیت مدل مشخصی را به عنوان مناسب‌ترین مدل معرفی کرد ولی از میان دو مدل مذکور، مدل لجستیک شرطی تا حدودی برازش بهتری داشته و می‌توان برای تحلیل چنین داده‌هایی از آن بهره برد.

واژه‌های کلیدی: بیماری پرفشاری خون، مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر، مدل لجستیک شرطی با خوشه‌های تصادفی، معیارهای برازش

مقدمه

در تحلیل کلاسیک آماری معمولاً فرض می‌شود که مشاهدات ازهم مستقل هستند، در حالی‌که این فرض در بسیاری از مطالعات به خصوص در حوزه پزشکی برقرار نیست و در مطالعات زیادی مشاهدات به‌هم وابسته

مولف مسئول: انوشیروان کاظم نژاد– تهران: دانشگاه تربیت مدرس، دانشکده علوم پزشکی، گروه آمارزیستی E-mail: kazem_an@modares.ac.ir
1-دانشجو کارشناسی ارشد، گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس
۲. گروه آمار زیستی، دانشکده علوم پزشکی، دانشگاه تربیت مدرس
۳. گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی
۴. مرکز تحقیات روان‌پزشکی و علوم رفتاری، گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران
تاریخ دریافت : ۸۹/۲/۱۹ تاریخ ارجاع جهت اصلاحات : ۸۹/۴/۱۲ تاریخ تصویب : ۸۹/۱۲/۲۸

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

هستند. برای مثال داده‌هایی که در طول زمان جمع‌آوری می‌شوند- مانند اندازه‌گیری فشار خون هر فرد در دوره‌های زمانی- یا داده‌هایی که در خوشه‌هایی دسته‌بندی می‌شوند. مشاهدات حاصل از یک آزمودنی در دوره‌های زمانی یا دسته بندی شده در یک خوشه شباهت بیشتری نسبت به مشاهدات آزمودنی‌ها یا خوشه‌های دیگر دارند. در نتیجه معمولاً دارای همبستگی مثبتی با هم هستند. برای تحلیل درست داده‌هایی که همبسته هستند، به یک مدل آماری نیازمندیم که در آن همبستگی میان مشاهدات لحاظ شود(۱). در مدل‌بندی حاشیه‌ای از پاسخ‌های خوشه‌بندی شده، ساختار وابستگی تؤام را به عنوان یک پارامتر مزاحم در نظر می‌گیرند. اثر تصادفی در این مدل از نوع R-side است(۳،۲). درحالی‌که در مدل‌بندی شرطی این وابستگی را به صورت عبارتی از سطح خوشه‌ای فرض کرده که دارای مقادیر یکسان برای مشاهدات داخل یک خوشه هستند که این مقادیر از مقادیر مشاهدات واقع در خوشه های دیگر متفاوت هستند. اثر تصادفی در این مدل از نوع G-side خواهد بود. بر اساس این دو مدل کلی، هرگاه متغیر پاسخ دوحالتی و مشاهدات همبسته باشند می‌توان از مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر و مدل لجستیک شرطی با خوشه‌های تصادفی استفاده کرد(۶-۴). مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر که توسط Liang و Zeger در سال ۱۹۸۶ معرفی شد؛ تعمیمی بر مدل‌های خطی تعمیم یافته[1] با لحاظ همبستگی بین داده‌ای است. در این مدل، همبستگی موجود بین پاسخ‌ها برای هر آزمودنی به عنوان پارامتر در نظر گرفته می‌شود. برآورد پارامترها و همچنین همبستگی میان مشاهدات دریافتی از هر فرد در آن با روش معادلات برآورد تعمیم یافته[2] به دست می‌آید. مدل لجستیک شرطی با خوشه‌های تصادفی در رده مدل‌های آمیخته خطی تعمیم یافته[3] قرار می‌گیرد. از این مدل‌ها زمانی استفاده می‌شود که بخواهیم اثرهای تصادفی را به عنوان بخشی از پیش بین خطی در مدل‌های خطی تعمیم یافته قرار دهیم. استفاده از تقریب لاپلاس که بر پایه درست‌نمایی تقریبی است، یکی از روش‌های استنباطی مدل‌های آمیخته خطی تعمیم یافته محسوب می‌شود. اگرچه مدل‌های با اندازه‌گیری مکرر، کاربردی‌ترین مدل در همبستگی‌های با دوره‌های زمانی است، مدل‌های شرطی با خوشه‌های تصادفی در برخی شرایط می‌توانند مناسب‌تر باشند. چون محاسبه تابع درست‌نمایی مدل‌های شرطی با خوشه‌های تصادفی به نوعی فرم بسته‌ای ندارد. در گذشته استفاده محدودی از این مدل‌ها می‌شد اما امروزه با پیشرفت‌هایی که در زمینه نرم‌افزارهای آماری روی داده است، استفاده کاربردی از این مدل‌ها در حال افزایش است(۸،۷).

فشارخون بالا عبارت است از افزایش فشار وارده از جریان خون به دیواره رگ‌های خونی. فشار خون بالا یکی از مهمترین عوامل خطر بروز بیماری‌های عروق کرونری و شایع‌ترین عامل بروز نارسایی قلبی و سکته مغزی در بسیاری از کشورها و نیز مهمترین علت ایجاد نارسایی کلیه می‌باشد. در حقیقت بعد از دیابت، مهم‌ترین دلیل شایع نارسایی مزمن کلیوی پرفشاری خون است. با تغییر الگوی اپیدمیولوژیک بیماری‌ها در کشورهای پیشرفته و در حال توسعه و افزایش امید به زندگی، افزایش استرس‌های ناشی از زندگی شهرنشینی و تغییر الگوهای غذایی، پرفشاری خون به عنوان مهمترین عامل خطرزای زمینه‌ساز بیماری‌های قلبی- عروقی، تهدیدی جدی تلقی می‌شود(۱۷-۹).

هدف این مطالعه استفاده از مدل‌های لجستیک حاشیه‌ای با اندازه‌گیری مکرر و لجستیک شرطی با خوشه‌های تصادفی برای بررسی عوامل خطر موثر بر بیماری پرفشاری خون در استان مازندران است.

1. Generalized linear models (GLM)
2. GEE
3. Generalized linear mixed models (GLMM)

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

مواد و روش ها

این مطالعه از نوع کاربردی می‌باشد و داده‌های مورد استفاده در آن که به شیوه مقطعی گردآوری شده است، قسمتی از طرح بررسی عوامل خطر بیماری‌های غیر واگیردار در استان مازندران در سال ۱۳۸۶ بود. روش نمونه‌گیری به صورت طبقه‌ای (طبقات شامل شهرها، روستاها و در شهرهای بزرگ، مناطق شهری) و در داخل طبقات بصورت تصادفی خوشه‌ای یک مرحله‌ای بود. برای بالا بردن کارایی نمونه‌گیری ترتیب قرار دادن طبقات در شهرها بر اساس موقعیت اقتصادی- اجتماعی بود. نمونه‌گیری در کلیه مراحل بصورت تخصیص متناسب[1] صورت گرفت. حجم نمونه به گونه‌ای بود که حداقل تعداد نمونه در هر خانه[2] سنی- جنسی ۱۰۰ نفر باشد. بنابراین در هر گروه سنی تعداد ۲۰۰ نفر (۱۰۰ زن و ۱۰۰ مرد) انتخاب شدند. در نهایت باتوجه به بررسی در پنج گروه سنی، حجم نمونه در استان برابر ۱۰۰۰ نفر و در هر خوشه ۲۰ نفر بود که بدین ترتیب در استان ۵۰ خوشه بررسی شد(تصویر۱).

جمع‌آوری اطلاعات در این مطالعه از سه طریق زیر انجام گرفته است: پرسشنامه همراه با مصاحبه حضوری؛ مطالعات بالینی با اندازه‌گیری قد، وزن، دور کمر و فشار خون؛ اندازه‌گیری قندخون ناشتا، کلسترول تام و HDL پلاسما.

متغیرهای مورد سنجش

ابتلا به بیماری پر فشاری خون: متغیر پاسخ عبارت است از بیماری پرفشاری خون (که دو حالتی می‌باشد: پرفشاری خون ندارد =۰ و پرفشاری خون دارد =۱) که این متغیر برای هر فرد در سه دوره زمانی با فواصل زمانی یک دقیقه اندازه‌گیری شد. این متغیر دو حالتی با توجه به تعریف JNC-7 [3] بدست آمد؛ یعنی فشارخون سیستولیک بیشتر یا مساوی ۱۴۰ میلی‌متر جیوه و یا فشار دیاستولیکی بیشتر یا مساوی ۹۰ میلی‌متر جیوه به عنوان بیماری پرفشاری خون در نظر گرفته شد(۹). همچنین کلیه افراد تحت درمان دارویی ضد فشار نیز در گروه هیپرتانسیو قرار گرفتند.

متغیرهای قند خون ناشتا، کلسترول و HDL با نقاط برش به ترتیب ۱۲۶،۲۴۰ و ۳۵ به صورت متغیرهای کیفی در نظر گرفته شدند. متغیرهای متوسط فعالیت بدنی و مصرف میوه و سبزی به ترتیب از مجموع متوسط فعالیت و ورزش یک فرد در طول هفته و متوسط تعداد مصرفی واحدهای میوه و سبزی در طول یک هفته بدست آمد. همچنین متغیرهای سن، جنس، محل سکونت (شهر، روستا) و نمایه توده بدنی (BMI [4]) نیز به عنوان متغیرهای کمکی درنظر گرفته شد.

تجزیه و تحلیل اطلاعات

مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر

مدل لجستیک با اندازه‌گیری مکرر که از نوع مدل‌های رگرسیونی شبه پارامتری است؛ برای تحلیل مشاهدات دوحالتی همبسته در مطالعات طولی از نوع اندازه‌گیری مکرر و مطالعات چند سطحی کاربرد فراوان دارد. ۳ متغیر دوحالتی فشارخون را با توجه به سه دوره زمانی با (Y_1, Y_2, Y_3) نشان می‌دهیم. قرار می‌دهیم:

$$\pi_i = P(Y_1 = i_1, Y_2 = i_2, Y_3 = i_3)$$

ملاحظه می‌شود $2^3 = 8$ جدول متقاطع از توزیع توأم (Y_1, Y_2, Y_3) بدست می‌آید. برآورد پارامترها و همچنین همبستگی میان مشاهدات دریافتی از هر فرد برای این مدل با روش معادلات برآورد تعمیم یافته بدست می‌آید. در روش GEE [5] نیازی به تعیین کامل توزیع

1. Proportional
2. Cell
3. Joint national committee

4. Body mass index
5. Generalized estimating equation

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

داده‌ای نیست. به عبارتی، سازگاری برآورد کننده GEE تنها به تعیین درست توابع میانگین یعنی $\mu_i \quad 1\le i\le n$ در معادله

$$\sum_{i=1}^{N} \dot{\mu}_i' V_i^{-1} (y_i - \mu_i) = 0$$

بستگی دارد که در آن $y_i = (y_{ij})_{1\le j\le n}$ ، $\mu_i = E(y_i)$ ، $V_i = \mathrm{var}(y_i) = (A_i^{1/2} R A_i^{1/2})\phi$ و $\dot{\mu} = \frac{\partial \mu}{\partial \beta}$ بوده و ϕ پارامتر بیش پراکنش است.

مدل لجستیک شرطی با خوشه‌های تصادفی

مدل لجستیک شرطی با خوشه‌های تصادفی در رده مدل‌های آمیخته خطی تعمیم یافته قرار می‌گیرد. یک مدل GLMM [1] را در حالت کلی می‌توان به صورت زیر در نظر گرفت:

$$y = \mu + e$$

μ برداری از مقدار مورد انتظار شرطی مشاهدات است که به پارامترهای مدل با یک تابع پیوند g مربوط می‌شود. به طوری که:

$$\mu = E(y \mid \alpha)$$
$$g(\mu) = X\beta + Z\alpha$$

X و Z ماتریس‌های طرح معلوم از اثرهای ثابت و تصادفی هستند، β و α به ترتیب برداری از پارامترهای اثر ثابت و تصادفی همانند مدل آمیخته نرمال هستند. بر اساس مدل لجستیک شرطی با خوشه‌های تصادفی، هر فرد را می‌توان به‌عنوان یک خوشه در نظر گرفت که این خوشه‌ها در واقع سطوحی تصادفی محسوب می‌گردند. بر اساس این مدل، مشاهدات داخل هر خوشه با هم همبسته بوده (اندازه‌گیری فشارخون مربوط به یک فرد) ولی با مشاهدات سایر خوشه‌ها همبستگی لحاظ نمی‌شود. در نتیجه مدل به‌صورت مدلی آمیخته لجستیک با اثر تصادفی افراد (اثر خوشه‌ها) و چندین اثر ثابت خواهد بود. اگر Y_{i1}, Y_{i2}, Y_{i3} تعداد موارد دارای بیماری فشارخون خوشه iام با توجه به سه زمان باشند؛ احتمال ابتلا به بیماری برای هر فرد با فرض ارتباط خطی لجیت احتمال‌ها و خوشه‌ها، به صورت مدل لجستیک با اثر تصادفی فرد خواهد بود. استنباط آماری این مدل با استفاده از تقریب لاپلاس بدست آمد.

معیارهای مقایسه

در این مطالعه برای مقایسه کارایی مدل‌ها از معیارهای جدول شماره ۱ استفاده شده است. در جدول شماره ۱؛ n نشان دهنده بعد مدل و n^* اندازه داده‌ای و d- تعداد پارامترهای کوواریانسی + رتبه X- است. عموما هدف این معیارهای برازش، اندازه‌گیری برازش مدل برآورد است. هر قدر مقادیر آن‌ها، کم‌تر باشد نشان دهنده مناسب بودن مدل است(۱۸،۱۹).

جدول شماره ۱ : معیارهای اطلاع

نام معیار	فرمول معیار اطلاع	مرجع
AIC[2]	$-2l + 2d$	Akaike (1974)
AICC[3]	$-2 + 2dn^*/(n^* - d - 1)$	Hurvich and Tsai (1989)
BIC[4]	$-2l + d\log n$	Schwarz (1978)
CAIC[5]	$-2l + d(\log n + 1)$	Bozdogan (1987)
HQIC[6]	$-2l + 2\log\log n$	Hannan and Quinn (1979)

برای بررسی اثر عوامل مختلف بر فشار خون بالا به گزارش نتایج آزمون Chi-square و مقایسه حاصل از برازش دو مدل لجستیک حاشیه‌ای و شرطی پرداخته شد. نتایج با استفاده از پردازه های GLIMMIX نرم افزار SAS نسخه ۹/۲ بدست آمده‌اند که در این پردازه امکان تحلیل همزمان متغیرهای مستقل کمی و کیفی برای اولین‌بار میسر شده است. همچنین سطح معنی‌داری آماری ۰/۰۵ در نظر گرفته شد(۲۲-۲۰).

یافته ها

هدف اولیه یافتن متغیرهای کیفی مرتبط با بیماری پرفشاری خون بود تا از آن‌ها بتوان در برازش مدل‌ها

1. Generalized linear mixed model

2. Akaike information criterion
3. Small sample bias corrected version of AIC
4. Bayesian inference criterion
5. Consistent Akaike's information criterion
6. Hannan Quinn information criteria

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

بهره گرفت. با استفاده از آزمون Chi-square و بر اساس نتایج جدول شماره ۲ ملاحظه می‌شود که؛ از مجموع ۹۹۹ فرد مورد مطالعه، ۳۸ درصد دارای بیماری پرفشاری خون بوده‌اند. نتایج جدول شماره ۲ نشان می‌دهد که: نوع شغل و رده سنی با این بیماری مرتبط بوده در حالی‌که چنین چیزی در رابطه با متغیر محل سکونت دیده نشد. افرادی که میزان قند خون ناشتای آن‌ها بیشتر از ۱۲۶ mg/dL بوده است (۷۴/۵ درصد) بیشتر از افرادی که قند خون ناشتای آن‌ها پایین‌تر یا مساوی ۱۲۶ mg/dL (۴۱ درصد) دچار پرفشاری خون می‌شوند و این‌که افراد با کلسترول خون کمتر از ۲۴۰ (۴۱ درصد) در مقایسه با افرادی که میزان کلسترول خون آن‌ها بالاتر از ۲۴۰ بوده است (۵۵/۹ درصد) کمتر مبتلا به فشار خون بالا می‌شوند. در حالت کلی بدون در نظر گرفتن گروه‌های سنی تفاوتی بین دو جنس از لحاظ ابتلا به بیماری پرفشاری خون وجود نداشت($p=۰/۱۲۲$).

جدول شماره ۳ ابتلا به بیماری پرفشاری خون را به تفکیک گروه‌های سنی در دو جنس نشان می‌دهد. ابتلا به پرفشاری خون با افزایش سن، افزایش یافته و در مردان در گروه‌های سنی ۱۵ تا ۲۴ و ۴۵ تا ۵۴ سال این درصد بیشتر از زنان است اما در بقیه گروه‌های سنی در زنان بیشتر مشاهده شد.

جدول شماره ۴ بیانگر تحلیل مدل لجستیک حاشیه‌ای با اندازه گیری مکرر و مدل لجستیک شرطی با خوشه‌های تصادفی است. در این جدول برآورد ضرایب و نسبت شانس (OR)، خطای استاندارد و فاصله اطمینان ۹۵ درصد ضرایب به همراه آزمون مربوطه به آن -که بی‌تاثیر بودن متغیر مربوطه بر وضعیت فشارخون افراد را آزمون می‌کند- برای هر یک از متغیرها به تفکیک دو مدل آورده شده است. بر اساس این جدول در هر دو مدل، متغیرهای گروه سنی، فعالیت بدنی، نمایه توده بدنی و مصرف میوه و سبزی از عوامل تاثیرگذار بر بیماری پرفشاری خون شناخته شدند. اضافه بر این متغیرهای جنسیت و قند خون ناشتا فقط در مدل لجستیک حاشیه‌ای از لحاظ آماری معنی‌دار گردید و دیگر متغیرهای زمینه‌ای مورد بررسی از لحاظ آماری معنی‌دار نبودند. در جدول شماره ۵ با توجه به معیارهای اطلاع، آماره‌های نیکویی برازش و مقادیر لگ درست‌نمایی که از هر دو مدل حاشیه‌ای و شرطی بدست آمدند، در دو مدل با هم مقایسه شدند. تمامی این آماره‌ها برای مدل لجستیک شرطی با اثر تصادفی افراد، دارای مقادیر کمتری نسبت به مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر هستند. مقادیر کمتر این آماره‌ها دلالت بر بهتر بودن مدل برازشی دارد.

جدول شماره ۲: توزیع فراوانی افراد بر اساس ابتلا به بیماری فشارخون و متغیرهای پایه رسته‌ای

متغیر	دارای فشارخون	سالم	سطح معنی داری
	تعداد(درصد)	تعداد(درصد)	
کل افراد	۳۸۰ (۳۸)	۶۱۹ (۶۲)	
جنسیت			۰/۱۲۲
مرد	۱۸۴ (۳۶/۹)	۳۱۵ (۶۳/۱)	
زن	۱۹۶ (۳۹/۲)	۳۰۴ (۶۰/۸)	
محل سکونت			
روستایی	۲۵۹ (۳۹/۳)	۴۰۰ (۶۰/۷)	۰/۱۲۶
شهری	۱۲۱ (۳۵/۶)	۲۱۹ (۶۴/۴)	
نوع شغل			
کارمند	۲۴ (۲۳/۳)	۷۹ (۷۶/۷)	
آزاد	۹۸ (۳۶/۷)	۱۶۹ (۶۳/۳)	
در حال تحصیل-سرباز	۱۷ (۲۰/۴)	۶۶ (۷۹/۵)	۰/۰۰۰
خانه‌دار	۱۷۵ (۴۳/۶)	۲۲۶ (۵۶/۴)	
بیکار	۱۴ (۳۷/۸)	۲۳ (۶۲/۲)	
سایرمشاغل	۵۲ (۴۸/۱)	۵۶ (۵۱/۹)	
سن			
۱۵ - ۲۴	۳۵ (۱۷/۵)	۱۶۵ (۸۲/۵)	
۲۳ - ۳۴	۴۳ (۲۱/۷)	۱۵۵ (۷۸/۳)	
۳۵ - ۴۴	۶۷ (۳۲/۵)	۱۳۹ (۶۷/۵)	۰/۰۰۰
۴۵ - ۵۴	۱۰۳ (۵۳/۱)	۹۱ (۴۶/۹)	
۵۵ - ۶۴	۱۳۲ (۶۵/۷)	۶۹ (۴۳/۳)	
قندخون ناشتا			
<=۱۲۶	۲۷۳ (۴۰/۹)	۳۹۵ (۵۹/۱)	۰/۰۰۰
>۱۲۶	۳۸ (۷۴/۵)	۱۳ (۲۵/۵)	
میزان کلسترول خون			
<=۲۴۰	۲۴۹ (۴۱)	۳۵۹ (۵۹)	۰/۰۰۰
>۲۴۰	۶۲ (۵۵/۹)	۴۹ (۴۴/۱)	

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

جدول شماره ۳ : ابتلا به بیماری به تفکیک جنس و گروه‌های سنی

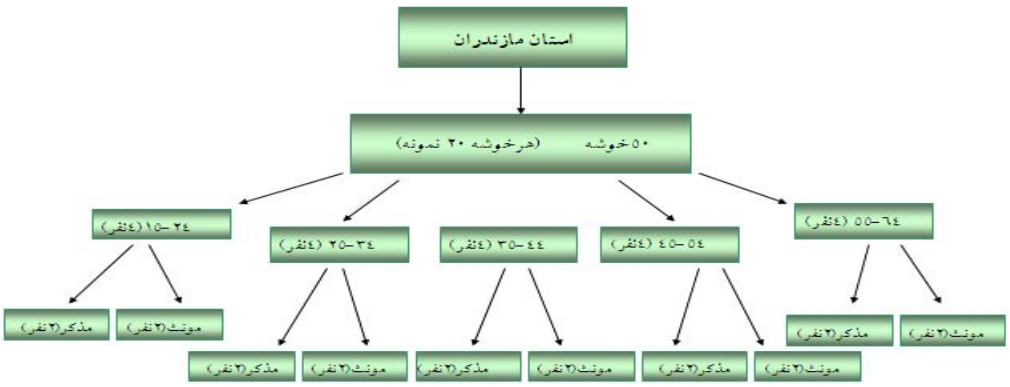
گروه سنی	ابتلا به پر فشاری خون	
	مرد تعداد (درصد)	زن تعداد (درصد)
۱۵-۲۴	۲۸ (۱۵/۲)	۷ (۳/۶)
۲۵-۳۴	۱۹ (۱۰/۳)	۲۴ (۱۲/۲)
۳۵-۴۴	۳۲ (۱۷/۴)	۳۵ (۱۷/۹)
۴۵-۵۴	۵۱ (۲۷/۷)	۵۲ (۲۶/۵)
۵۵-۶۴	۵۴ (۲۹/۳)	۷۸ (۳۹/۸)
کل	۱۸۴ (۱۰۰)	۱۹۶ (۱۰۰)

جدول شماره ۵ : مقایسه مدل‌ها با استفاده از آماره‌های نیکویی برازش

معیارهای برازش	لجستیک حاشیه‌ای	لجستیک شرطی
-2 Res Log Pseudo-Likelihood	۶۵۳۱/۸۱	۶۳۴۸/۲۶
Pseudo-AIC	۶۵۷۳/۸۱	۶۳۹۰/۲۶
Pseudo-AICC	۶۵۷۴/۵۱	۶۳۹۰/۹۵
Pseudo-CAIC	۶۶۸۲/۹۱	۶۵۰۷/۴۸
Pseudo-HQIC	۶۶۱۴/۶۹	۶۴۲۷/۴

جدول شماره ٤ : تعیین عوامل موثر بر پرفشاری خون با استفاده از مدل‌های لجستیک حاشیه‌ای با اندزه‌گیری مکرر و لجستیک شرطی با خوشه‌های تصادفی

متغیرها	لجستیک حاشیه ای با اندازه‌گیری مکرر					لجستیک شرطی با خوشه‌های تصادفی				
	$\hat{\beta}\,(SE)$	نسبت شانس (OR)	فاصله اطمینان ۹۵درصد	آماره آزمون	سطح معنی‌داری	$\hat{\beta}\,(SE)$	نسبت شانس (OR)	فاصله اطمینان ۹۵درصد	آماره آزمون	سطح معنی‌داری
جنسیت										
مذکر	—	—	—							
مونث	۰/۶۶(۰/۳۰)	۱/۹۳	(۱/۲۵ و ۰/۰۶)	۲/۱۸	۰/۰۲۹	۰/۶۷(۰/۴۸)	۱/۹۵	(۱/۶۱ و ۰/۲۸)	۲/۳۸	۰/۰۷
گروه سنی										
۱۵-۲۴	-۱/۸۶(۰/۳)	۰/۱۶	(-۱/۲۷ و-۲/۴۴)	-۶/۲۱	۰/۰۰۰	-۱/۸۹(۰/۴۶)	۰/۱۵	(-۰/۹۹ و-۲/۸۰)	-۴/۱۴	۰/۰۰۰
۲۵-۳۴	-۱/۵۳(۰/۲۱)	۰/۲۲	(-۱/۱۲ و-۱/۹۴)	-۷/۱۷	۰/۰۰۰	-۱/۵۷(۰/۳۵)	۰/۲۱	(-۰/۸۸ و-۲/۲۷)	-۴/۶۴	۰/۰۰۰
۳۵-۴۴	-۱/۱۳(۰/۱۹)	۰/۳۲	(-۰/۷۵ و-۱/۵۱)	-۵/۸۰	۰/۰۰۰	-۱/۱۶(۰/۳۳)	۰/۳۱	(-۰/۵۱ و-۱/۸۲)	-۳/۴۸	۰/۰۰۰
۴۵-۵۴	-۰/۴۲(۰/۱۹)	۰/۶۶	(۰/۰۵ و -۰/۷۸)	-۲/۲۶	۰/۰۲۴	-۰/۴۰(۰/۳۳)	۰/۶۷	(۰/۲۴ و -۱/۰۵)	-۱/۲۳	۰/۲۲
۵۵-۶۴										
نوع شغل										
کارمند	-۰/۶۹(۰/۴۹)	۰/۵۰	(-۱/۶۶و۰/۲۷)	-۱/۴۱	۰/۱۶	-۰/۱۶(۰/۶۲)	۰/۸۵	(-۱/۰۵و-۱/۳۶)	-۰/۲۵	۰/۸۰
آزاد	-۰/۱۴(۰/۴۳)	۰/۸۷	(-۰/۹۸و۰/۷)	-۰/۳۳	۰/۷۴	۰/۱۱(۰/۸۷)	۱/۱۲	(-۱/۸۲و۱/۵۹)	-۰/۱۳	۰/۹
بیکار	-۰/۱۰(۰/۶۲)	۰/۹۰	(۱/۱۲و-۱/۳۲)	-۰/۱۶	۰/۸۷	-۰/۷۲(۰/۶۹)	۰/۴۹	(-۲/۰۸و۰/۶۵)	-۱/۰۳	۰/۳۰
خانه دار	۰/۷۱(۰/۴۶)	۲/۰۳	(۱/۵۵و-۰/۱۴)	۱/۶۴	۰/۱۰	۰/۷۲(۰/۶۷)	۲/۰۵	(-۰/۶۰و۲/۰۵)	۱/۰۷	۰/۲۸
سایر	۰/۱۴(۰/۴۳)	۱/۱۵	(-۰/۷و۰/۹۹)	۰/۳۳	۰/۷۴	۰/۱۶(۰/۶۵)	۱/۱۷	(-۱/۱۳و۱/۴۴)	۰/۲۴	۰/۸۱
در حال تحصیل	—	—	—	—	—	—	—	—	—	—
قندخون ناشتا										
کمتر از ۱۲۶	-۰/۶۳(۰/۲۸)	۰/۵۳	(-۰/۱۷و-۰/۰۹)	-۲/۲۸	۰/۰۲	-۰/۶۳(۰/۵۹)	۰/۵۳	(-۱/۷۸و۰/۵۲)	-۱/۰۸	۰/۲۸
بیشتر از ۱۲۶	—	—	—	—	—	—	—	—	—	—
کلسترول										
کمتر از ۲۴۰	-۰/۱۹(۰/۲۱)	۰/۸۳	(-۰/۶و۰/۲۲)	-۰/۹۰	۰/۳۷	-۰/۱۹(۰/۴۰)	۰/۸۳	(-۰/۹۸و۰/۶۱)	-۰/۴۶	۰/۶۵
بیشتر از ۲۴۰	—	—	—	—	—	—	—	—	—	—
HDL										
کمتر از ۳۵	۰/۰۸(۰/۲۵)	۱/۰۸	(۰/۵۶ و -۰/۴۰)	۰/۳۲	۰/۷۵	۰/۰۸(۰/۴۵)	۱/۰۸	(-۰/۸۱و۰/۹۷)	۰/۱۹	۰/۸۵
بیشتر از ۳۵	—	—	—	—	—	—	—	—	—	—
نمایه توده بدنی BMI	۰/۰۸(۰/۰۱)	-	(۰/۰۵و۰/۱)	۵/۹۵	۰/۰۰۱	۰/۰۸(۰/۲۴)	-	(۰/۰۳و۰/۱۳)	۳/۳۷	۰/۰۰۰
مصرف میوه و سبزی	-۰/۰۹(۰/۰۴)	-	(-۰/۰۰۷و-۰/۱۷)	-۲/۱۴	۰/۰۳۲	-۰/۲۸(۰/۰۸)	-	(-۰/۴۵و-۰/۱۱)	-۲/۲۵	۰/۰۰۰
فعالیت بدنی	-۰/۰۰۲(۰/۰۰۰۱)	-	(-۰/۰۰۰۴و۰)	-۱/۹۷	۰/۰۴۸	-۰/۰۰۳(۰/۰۰۰۱)	-	(-۰/۰۰۰۵و۰)	-۳/۷۵	۰/۰۰۰



تصویر شماره ۱ : ساختار نمونه‌گیری

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

بحث

هدف این مطالعه اهمیت توجه به ساختار داده‌ها (داده‌های گرد آوری شده در طول زمان مانند داده‌های فشارخون) و برازش مدل مناسب جهت چنین داده‌هایی بود تا همبستگی میان مشاهدات در آن لحاظ شود. با توجه به ساختار داده‌ای، دو مدل لجستیک حاشیه‌ای با اندازه‌گیری مکرر و لجستیک شرطی با خوشه‌های تصادفی افراد برازش داده شد و برای مقایسه مدل‌ها از معیارهای نیکویی برازش استفاده شد. نتایج نشان دادند که مقادیر برآورد ضرایب در مدل لجستیک حاشیه‌ای با GEE و مدل لجستیک شرطی با روش تقریب لاپلاس به هم نزدیک هستند. خطای استاندارد برآورد پارامترها در مدل لجستیک شرطی نسبت به مدل لجستیک حاشیه‌ای باGEE بیشتر برآورد شده است. بعبارتی در مدل GEE ضرایب دچار کم برآوردی است. همچنین مقادیر آماره‌های نیکویی برازش برای مدل لجستیک شرطی کمتر از مدل لجستیک حاشیه‌ای بود. نتیجه آن که اگرچه در پایان نمی‌توان با قاطعیت مدل مشخصی را به عنوان مناسب‌ترین مدل معرفی کرد ولی از میان دو مدل مذکور، مدل لجستیک شرطی تا حدودی برازش بهتری داشته و می‌توان برای تحلیل چنین داده‌هایی از آن بهره برد.

مطالعات دیگر نیز در چنین مواقعی مدل شرطی را پیشنهاد کرده‌اند و نشان داده‌اند که مدل حاشیه‌ای مناسب چنین داده‌هایی نیست. بعنوان نمونه در مطالعه Wang و Devereux در سال ۲۰۰۶ بر داده‌های فشارخون، مدل‌های خطی تعمیم یافته و مدل‌های آمیخته نرمال مقایسه شد که در هر دو مدل متغیرهای سن، دیابتی بودن فرد، اضافه وزن و چاقی و الکلی بودن بر پرفشاری خون موثر بودند ولی ضرایب دو مدل با هم متفاوت بود(۲۳). در مطالعات انجام شده انواع دیگری از مدل‌های شرطی و حاشیه‌ای بر روی داده‌های مشابه، برازش شده است ولی پیشنهاد شده در صورتی که داده‌ها از نوع چند سطحی و خوشه‌ای باشند از مدل‌های شرطی استفاده گردد(۲۹-۲۴).

References

1. Diggle P, Liang K, Zeger S. Analysis of Longitudinal Data. 2nd ed. New York: Oxford University Press, 2002.
2. Dobson A. An Introduction to generalized linear Models. 2nd ed. London: Chapman & Hall/CRC, 2001.
3. Kevin H, Kim N. General Linear Models: Theory and applications using SAS software. 2nd ed. London: Chapman & Hall/CRC, 2007.
4. Charles E, McCulloch C, Seare S. Generalized Linear and mixed models. 1st ed. London: Wiley Interscience Publication; 2001. P 220-246.
5. Brady W, Kathleen B, Andrzej T. Linear mixed models: A practical guide using statistical software. 1st ed, London: Chapman & Hall/CRC; 2007. P 9-13.
6. Muller K. Introduction to mixed modeling Univariate, Multivariate and mixed models. 1st ed. London: Willey Interscience Publications, 2006.
7. Jiang J. Linear and Generalized Linear Mixed Models and Their Applications. 1st ed. NewYork: Springer Science; 2007. P 119-122.
8. Breslow NE, Clayton DG. Approximate Inference in generalized linear mixed models. J Am Stat Assoc 1993; 88(421): 9-25.
9. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. The

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure: The JNC7 report. JAMA 2003; 289(19): 2560-2572.

10. Kriby M. NICE and Hypertension. Sage 2006; 6(4): 177-179.
11. Cohen JD. Hypertension Epidemiology and Economic Burden: Refining Risk Assessment To Lower Costs. Manag Care 2009; 18(10): 51-58.
12. Karimloo M. Application of General Linear Models in Blood pressure. Tehran Univ Med Sci 1991; 12(3): 19-27 (Persian).
13. Amirkhizi F, Siyasi F, Minaei S, Jalali M, Dorosti A, Chamri M. Blood pressure status of rural women in Kerman and its relationship with anthropometric indexes. Lorestan Univ Med Sci 2008; 10(2): 31-38 (Persian).
14. Sadeghi M, Roohafza H, Sadry Gh, Bahonar A, Saeidi M, Asgari S, et al. Prevalence of high blood pressure and its relation with cardiovascular risk factors. Qazvin Univ Med Sci 2003; 26(3): 46-52 (Persian).
15. Azizi A, Abasi M, Abdoli Gh. The prevalence of Hypertension and its Association with age, sex and BMI in a population being educated using community based medicine in Kermanshah. IJEM 2008; 10(4): 323-329 (Persian).
16. Ahmadi A, Hasanzade J, Rajaei fard A. Risk Factors of Hypertension in Koohrang city of Charmahal o bakhtiyari in 2007. IREA 2008; 4(2): 19-25 (Persian).
17. Goodarzi M, Badakhsh M, Masinaei N, Abbas zade M. Hypertension prevalence in over 18 years old. JIUMS 2004; 11(43): 821-828 (Persian).
18. Jiang J. Goodness-of-fit tests for mixed model diagnostics. Ann Statist 2001; 29(4): 1137-1164.
19. Brown H, Prescott R. Applied Mixed Models in Medicine. 2nd ed, London: John Wiley and Sons; 2006. P 33-59.
20. Pinheiro JC, Chao EC. Efficient Laplacian and Adaptive Gaussian Quadrature Algorithms for Multilevel Generalized Linear Mixed Models. J Comput Graph Stat 2006; 15(1): 58-81.
21. Brady T, West Kathleen B, Welch Andrzej T. Linear mixed models A practical guide using statistical software. Linear mixed models A Practical Guide Using Statistical Software. London: Chapman & Hall/CRC: Taylor & Francis Group, 2007.
22. Little R, Milliken G. SAS/STAT 9.2 User's Guide,Chapter 38:The GLIMMIX Procedure. 1st ed. Vancouver: SAS Publishing, 2006.
23. Wang W, Devereux R. A Longitudinal Study of Hypertension Risk Factors and Their Relation to Cardiovascular Disease. Hypertension 2006; 47(3): 403-409.
24. Fouriaud C, Jacquint-Salord MC, Degoulet P, Aime F, Lang T, Laprugne J, et al. Influnce of socio professional condition on blood pressure levels and hypertension control. Am J Epidemiol 1984; 120(1): 7286-7292.
25. Suarez LM, Criqui H, Barrett-Connor E. Spouse Concordance for systolic and diastolic Blood pressure. Am J Epidemiol 1983; 118(3): 345-351.
26. Lane D. Ethnic differences in Hypertension and blood pressure control in the UK. QJM 2001; 94(7): 391-396.
27. Kurt R, Maurer C. Blood pressure levels for person ages 6 months-74 years, 1976-80.

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]

National Center for Health Statistics. DHHS 1983; 11(233): 10-69.

28. Yadav S, Boddula R, Genitta G, Bhatia V, Bansal B, Kongara S, et al. Prevalence and risk factors of pre-hypertension and hypertension in an affluent north Indian population. Indian J Med Res 2008; 128(6): 712-720.

29. Roca C, Suarez A, Ceballos J. Control of hypertension in patients at high risk of cardiovascular disease. QJM 2005; 98(8) 581-588.

[Downloaded from jmums.mazums.ac.ir on 2024-03-29]