مدل‌های بازگشتی تصحیح واریانس در مطالعات پزشکی

آرزو باقری[1]، مهسا سعادتی[1*]

۱. مؤسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور، تهران، ایران

نشریه پایش
تاریخ پذیرش مقاله: ۱۳۹۶/۶/۱۲
سال شانزدهم، شماره پنجم، مهر ـ آبان ۱۳۹۶ صص ۵۸۵-۵۷۵
[نشر الکترونیک پیش از انتشار- ۲۱ شهریور ۹۶]

چکیده

مقدمه: در بسیاری از مطالعات پزشکی و اپیدمیولوژی هر فرد تحت شرایطی می‌تواند وقایعی را چندین بار، یکی پس از دیگری و به صورت متوالی تجربه نماید که آن‌ها را وقایع بازگشتی می‌نامند. اهداف وسیع تحلیل این‌گونه وقایع می‌تواند شامل توصیف فرآیند رخداد واقعه در افراد، پراکندگی وقایع از فردی به فرد دیگر و رابطه بین متغیرهای پیش‌بین ثابت یا وابسته به زمان با زمان رخداد واقعه باشد. از این رو، هدف اصلی این مقاله ارائه و تفسیر مدل‌های پیشرفته آماری به منظور تحلیل درست این وقایع بود که در سال‌های اخیر پیشنهاد شده‌اند.

مواد و روش کار: در این مقاله، برای مدلسازی وقایع بازگشتی روش‌های تصحیح واریانس که شامل مدل‌های AG، شرطی PWP (PWP-TT و PWP-GT) و مدل حاشیه‌ای WLW هستند، ارائه و با یکدیگر مقایسه شده اند.

یافته ها: اغلب، مدل‌های شرطی PWP زمانی کاربرد دارند که تعداد وقایع کمی برای هر فرد رخ دهد و مخاطره رخداد مجدد واقعه میان رخدادهای مختلف تغییر نماید. در حالی‌که مدل حاشیه‌ای WLW و مدل AG با پیش‌بین‌های وابسته به زمان، برای تعیین فراوانی وقایع به کار می‌روند.

بحث و نتیجه گیری: محققان حوزه پزشکی و اپیدمیولوژی برای انتخاب صحیح مدل‌های بازگشتی باید به عوامل مختلفی از جمله تعداد وقایع، ارتباط میان وقایع، متغیر بودن اثر عوامل از واقعه‌ای به واقعه دیگر، فرایند بیولوژیکی و ساختار همبستگی داده‌ها توجه نمایند.

کلیدواژه: وقایع بازگشتی، مدل‌های بازگشتی، مدل AG، مدل حاشیه‌ای WLW، مدل شرطی PWP

* نویسنده پاسخگو: تهران، خیابان شهید بهشتی، خیابان پاکستان، خیابان دوم، پلاک ۵، موسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور
تلفن: ۸۸۵۳۴۳۲۱
E-mail: mahsa.saadati@gmail.com

[Downloaded from payeshjournal.ir on 2025-07-06]

[DOR: 20.1001.1.16807626.1396.16.5.2.7]

مقدمه

در مطالعات اپیدمیولوژی و بالینی، محققان اغلب با داده‌های همبسته‌ای مواجه می‌شوند که پاسخ مورد نظر آن‌ها زمان تا رخداد واقعه موردنظر (زمان بقا) است؛ در نتیجه هدف این مطالعات تحلیل چندمتغیره داده‌های بقای همبسته است[۱]. هم گروه شدن افراد در گروه‌ها (طبقات)، تجربه چندین واقعه از یک نوع یا وقایع کاملاً متفاوت توسط یک فرد در طول زمان منجر به تولید داده‌های بقای همبسته می‌شود که به ترتیب داده‌های وقایع خوشه‌ای (Cluster events) ، بازگشتی (Recurrent events) (عودکننده) و چندگانه (Multiple events) نامیده می‌شوند و هر یک روش تحلیل‌های خاصی دارند [۲]. داده‌های بازگشتی در مطالعات پزشکی نقشی مهم و اساسی دارند زیرا اغلب افراد می‌توانند بروز یک بیماری را به صورت تکراری و بیش از یکبار تجربه کنند؛ حملات آسم، عود تومور، تشنج، حملات میگرنی، عفونت‌های گوشی و تنفسی، حملات قلبی، جراحت‌ها و بستری شدن در بیمارستان مثال‌هایی از این نوع وقایع هستند [۳-۶]. وقایع بازگشتی دارای ترتیب ذاتی بوده و در طول زمان به صورت یک دنباله مشخص و پشت سر هم رخ می‌دهند. ممکن است قبل از اتمام مطالعه، نتوان وقایع بازگشتی را مشاهده نمود که در این صورت مشاهده سانسور شده (Censored) خواهد بود. به عنوان مثال، مشاهده بستری شدن گروهی از بیماران به دفعات پشت سرهم در بیمارستان با اتمام مدت زمان مطالعه یا با از دست دادن بیمار در طول مطالعه یا مرگ او پایان می‌یابد و در نتیجه این بیمار سانسور می‌شود. شکل (۱) مثالی از رخداد وقایع بازگشتی برای ۵ فرد را نشان می‌دهد [۷].

با توجه به این شکل، در میان این افراد، ۳ نفر حداقل دو واقعه را تجربه کرده‌اند (مربع‌های توپر) که از میان آن‌ها ۲ فرد تا پایان زمان مطالعه، (۶۰) ماه، سانسور شده‌اند (دایره‌های خالی)، سایر افراد به دلایلی نامرتبط با واقعه مورد نظر سانسور شده‌اند. فرد اول بیشترین تعداد واقعه (۶ واقعه) را در ماه‌های ۴، ۶، ۹، ۱۲، ۱۵ و ۲۸ و فرد ۳، ۲ واقعه را در ماه‌های ۱۲ و ۴۷ تجربه کرده‌اند. از آن‌جا که هدف بسیاری از مطالعات اپیدمیولوژی و بالینی ارزیابی اثر درمان در به تأخیر انداختن عود بیماری و طولانی کردن بقای بیماران است، وجود همبستگی میان این وقایع منجر به تولید مجموعه داده‌ای با ساختار پیچیده می‌گردد و اگر این همبستگی ذاتی در داده‌ها نادیده گرفته شود، فواصل اطمینان برای نرخ‌های برآوردی به غلط باریک و فرض صفر بیش از مواردی که باید، رد می‌شود. در نتیجه تحلیل داده‌های بازگشتی نیازمند استفاده از روش‌های پیشرفته و مناسب آماری برای پاسخگویی به سئوالات تحقیق است [۸]. پرکاربردترین روش برای تحلیل داده‌های بقا، مدل رگرسیون مخاطرات متناسب کاکس (Cox proportional hazards regression model) است که با توجه به پیش‌فرض مهم این مدل (استقلال مشاهدات)، از آن تنها برای مدلسازی زمان رخداد اولین واقعه می‌توان استفاده کرد که با توجه به نادیده گرفتن وقایع بعدی، برای تحلیل داده‌های بازگشتی ناکارا است [۹]. روش دیگر برای تحلیل این داده‌ها که همبستگی میان وقایع را در نظر می‌گیرند، منظور کردن تعداد وقایع برای هر فرد و برازش مدل پواسن (Poisson model) یا دوجمله‌ای منفی (Negative-Binomial model) است که در حالت کلی معادلات برآوردی تعمیم یافته (Generalized Estimating Equations) یا مدل‌های اثرات تصادفی (Random effect model) نامیده می‌شوند. با این حال این روش مدلسازی، زمان رخداد وقایع را نادیده می‌گیرد و در تحلیل داده‌های بقا، روش کارایی نیست. مدل‌های وقایع بازگشتی باید به گونه‌ای باشند که بتوانند از همه اطلاعات موجود در داده‌ها از جمله اثر وقایع قبلی بر واقعه فعلی برای برآورد دقیق خطرات نسبی این نوع وقایع استفاده نمایند [۱۰، ۶]. مدل برای داده‌های همبسته ابتدا در داده‌های خوشه‌ای به کاربرده شد که در آن ساختاری برای همبستگی داده‌ها در نظر گرفته نشد و تنها ماتریس واریانس و کواریانس برآوردی تعدیل گردید. پس از آن، انواع مدل‌ها با در نظر گرفتن الگوهای مختلف همبستگی توسط پژوهشگران مختلف ارائه شد [۶]. یکی از روش‌های متداول تحلیل وقایع بازگشتی، مدل‌های تصحیح واریانس (Random effect model) است که با توجه به تصحیحی که بر روی واریانس پارامترهای برآوردی انجام می‌دهد و کاربرد آن‌ها که ممکن است استنباط در سطح جامعه، ارائه برآورد میانگین اثر عوامل خطر یا تشخیص روند وقایع در طول زمان برای هر فرد باشد، با یکدیگر تفاوت دارند [۵].

با این‌که مباحث مختلفی در مورد تحلیل وقایع بازگشتی در پژوهش‌های مربوط به آمار زیستی، اپیدمیولوژی و پزشکی انجام گرفته است [۲۰-۱۱]، همچنان در بسیاری از مطالعات برای تحلیل داده‌های بازگشتی از روش‌های آماری ناکارا یا نامناسب استفاده می‌شود. در نتیجه، مقایسه این روش‌ها که محققان در حوزه‌های مطالعات اپیدمیولوژی یا مطالعات مرتبط را برای انتخاب مدل مناسب یاری می‌نماید، ضروری است.

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

در این مقاله چهار مدل بازگشتی مختلف تصحیح واریانس که از لحاظ پیش‌فرض‌ها و ساختار داده‌ای با یکدیگر تفاوت دارند، ارائه می‌گردد. یکی از تفاوت‌های اساسی این مدل‌ها در روشی است که برای مدلسازی وقایع تکراری استفاده می‌کنند؛ مدل‌های اندرسون- گیل ((Andersen-Gill (AG) و شرطی (Conditional models) پرنتیس- ویلیامز- پترسون (Peterson Prentice- Williams- (PWP)) که آن‌ها را مدل‌های فرایند مارکوفی (Markove process) نیز می‌نامند، فرض می‌کنند که الگوی وقایع تنها به وقایعی که بلافاصله قبل از آن‌ها رخ داده است، وابسته‌اند. مدل‌های حاشیه‌ای (Marginal models) وی- لین- ویسفیلد (Wei- Lin - Weissfield (WLW)) با میانگین‌ها (یا نرخ‌ها) به صورت فرایند شمارشی (Counting process) برخورد می‌کنند که امکان در نظر گرفتن هر نوع ساختاری را برای همبستگی میان وقایع تکراری فراهم می‌نمایند. هر یک از این مدل‌ها برای پاسخ به سئوال‌های متفاوتی در مطالعات مختلف به کار می‌روند.

هدف اصلی این مقاله، توصیف روش‌های فوق برای مدلسازی زمان رخداد وقایع بازگشتی به منظور ارائه دستورالعملی برای انتخاب روش مناسب در تحلیل این داده‌ها و تأثیر انتخاب هر یک از آن‌ها در تفسیر نتایج بود.

مواد و روش کار

اگرچه نادیده گرفتن وابستگی میان وقایع در تحلیل داده‌های بازگشتی، بر برآورد پارامترها اثرگذار نیست، اما باعث ایجاد اریبی برآورد واریانس این پارامترها و در نتیجه عدم دستیابی به استنباط صحیح در مورد اثر متغیرهای پیش‌بین بر زمان رخداد وقایع بازگشتی می‌گردد. در مدل‌های بازگشتی تصحیح واریانس، واریانس ضرایب از طریق طبقه‌بندی تابع خطر پایه یا بکارگیری روش‌هایی برای تعدیل همبستگی بین مشاهدات اصلاح می‌شود. شبیه‌سازی-های انجام شده جهت بررسی برآورد اثر درمان توسط محققان علوم پزشکی نشان می‌دهد که نادیده گرفتن پراکندگی میان افراد منجر به ایجاد اریبی منفی و در نتیجه کم برآوردی اثر درمان هستند [۲۱]. انواع مدل‌های زیرمجموعه این روش شامل مدل‌های AG، حاشیه‌ای WLW و شرطی PWP می‌باشند [۶]. که در ادامه به معرفی آن‌ها پرداخته می‌شود. در کلیه مدل‌ها فرض می‌شود که رخداد وقایع و سانسور شدن به شرط متغیرهای پیش‌بین از یکدیگر مستقلند (فرض استقلال سانسورها). یکی از مسائل مهم که منجر به تفاوت میان این مدل‌ها می‌شود، مقیاس زمانی است. مقیاس زمانی اغلب شامل زمان کل (TT)Total time)) و فاصله زمانی بین وقایع ((Gap time (GT) است. در مقیاس اول، زمان از آغاز مطالعه تا رخداد واقعه برای هر فرد و در مقیاس دوم، فاصله زمانی میان واقعه فعلی و واقعه قبلی مورد نظر است [۳].

مدل AG

که توسط اندرسون و گیل در سال ۱۹۸۲ معرفی شد یک مدل فرایند شمارشی است که از تعمیم مدل کاکس و برحسب افزایش تعداد وقایع در طول زمان مدلسازی می‌شود [۲۲]. در این مدل، پاسخ مورد نظر می‌تواند زمان شروع درمان تا رخداد واقعه باشد، در نتیجه از مقیاس زمان کل برای مدلسازی استفاده می‌شود. در این مدل همه افراد، در معرض خطر همه وقایع در کل دوره پیگیری هستند و در نتیجه می‌توانند در مجموعه خطر یک واقعه خاص بدون توجه به تعداد وقایعی که تاکنون تجربه کرده‌اند، قرار بگیرند. به عنوان مثال، مجموعه خطر می‌تواند شامل زمان واقعه اول برای یک فرد باشد در حالی‌که برای فرد دیگر این زمان، زمان رخداد واقعه دوم است. در این حالت تعداد افراد در معرض خطر با افزایش تعداد وقایع (j)، کاهش می‌یابد.این مدل یک تابع خطر پایه یکسان برای تمام وقایع در نظر می‌گیرد و یک پارامتر کلی برای عوامل مورد نظر برآورد می‌کند.

برآوردهای این مدل ایستا (Stable) و نسبت به مدل کاکس برای زمان تا رخداد واقعه اول کاراتر هستند[۶،۳، ۲۱] . شکل (۲) نحوه تعریف مجموعه خطر، تابع خطر پایه و ضرایب رگرسیونی را در این مدل نشان می‌دهد. همان‌گونه که ملاحظه می‌شود هر فرد به شرطی که واقعه قبلی را تجربه کرده باشد در معرض واقعه بعدی قرار می‌گیرد و وابستگی میان وقایع از طریق تعداد وقایع قبلی که فرد تجربه نموده در مدل وارد می‌شود. در این مدل تنها یک تابع خطر پایه مشترک (λ_0) وارد می‌شود و ضرایب رگرسیونی (β) برای همه وقایع یکسان است. (شکل ۲. مجموعه خطر، تابع خطر پایه (λ_0) و ضرایب رگرسیونی (β) در مدل AG).

در این مدل فرض می‌شود که همبستگی میان زمان‌های رخداد واقعه برای هر فرد می‌تواند با استفاده از وقایع قبلی تبیین شود که به معنی غیرهمبسته بودن زمان رخداد وقایع به شرط پیش‌بین‌ها است. این مدل زمانی مناسب است که همبستگی میان وقایع برای هر فرد با استفاده از متغیرهای پیش‌بین قابل توصیف باشد [۲۰]. بنابراین وابستگی می‌تواند با استفاده از متغیرهای پیش‌بین وابسته

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

به زمان، مانند تعداد وقایع قبلی یا توابعی از آن تعیین شود. اما اگر این فرض برقرار نباشد، باید از ماتریس کواریانس ساندویچی استوار (Robust Sandwich covariance matrix)، برای برآوردگرهای ضرایب رگرسیونی استفاده کرد [۹]که خطاهای استاندارد استواری را تولید می‌کند. مدل AG زمانی برای تحلیل داده‌ها به کار می‌رود که کلیه وابستگی‌ها میان وقایع از طریق پیش‌بین‌های وابسته به زمان بیان شوند و هدف از انجام مطالعه، بررسی اثر کلی بر شدت رخداد یک واقعه بازگشتی باشد. از این مدل برای تحلیل عود مجدد سلول‌های سرطانی [۱۲] و بستری شدن سالمندان در بیمارستان به علت بیماری قلبی [۱۸] استفاده شده است.

مدل حاشیه‌ای WLW

را می‌توان براساس میانگین تعداد وقایع وقتی هیچ متغیر پیش‌بین وابسته به زمانی وجود ندارد، تفسیر کرد [۸، ۱۷، ۲۱-۲۳]. در این مدل اگر فردی ۴ واقعه را تجربه کند، بازه‌های خطر برای آن شامل چهار بازه $(0, t_1]$، $(0, t_2]$، $(0, t_3]$ و $(0, t_4]$ می‌شود و توابع خطر و ضرایب رگرسیونی متفاوتی برای هر واقعه به صورت جداگانه در نظر گرفته می‌شود. شکل (۳) نحوه تعریف مجموعه خطر، تابع خطر پایه و ضرایب رگرسیونی را در این مدل نشان می‌دهد. همان‌گونه که ملاحظه می‌شود هر فرد با تجربه هر تعداد واقعه از شروع مطالعه مورد بررسی قرار می‌گیرد. در این مدل برای هر طبقه که از روی تعداد وقایع ساخته می‌شود یک تابع خطر پایه مجزا (λ_{0j}) در نظر گرفته می‌شود و ضرایب رگرسیونی (β_j) برای هر طبقه متفاوت خواهد بود (**شکل ۳. مجموعه خطر، تابع خطر پایه (λ_{0j}) و ضرایب رگرسیونی (β_j) در مدل WLW**).

این روش ساختار وابستگی میان زمان‌های رخداد واقعه برای یک فرد را تعیین نمی‌کند ولی برآورد واریانس‌ها براساس همبستگی بین وقایع تعدیل می‌گردد. از آن‌جا که این مدل همه وقایع بازگشتی برای یک فرد را به صورت یک فرایند شمارشی تنها در نظر می‌گیرد و به پیش‌بین‌های وابسته به زمان برای نشان دادن تاریخچه وقایع قبلی نیاز ندارد، از مدل AG منعطف‌تر و ساده‌تر است [۱۷]. اگر هیچ متغیر پیش‌بین وابسته به زمانی که اثر وقایع قبلی بر روی رخداد مجدد وقایع بعدی را به حساب می‌آورد در مدل AG وارد نشود، برآوردهای نقطه‌ای این مدل و مدل AG با یکدیگر مشابه، اما برآورد ماتریس کوواریانس آن‌ها برای ضرایب رگرسیونی با یکدیگر متفاوت خواهند بود. این روش در بسیاری از مطالعات پزشکی که در آن به دست آوردن ساختار وابستگی مورد نظر نباشد، ساختار وابستگی پیچیده یا مجهول بوده و به خصوص با استفاده از متغیرهای وابسته به زمان در مدل AG نتوان آن‌ها را در نظر گرفت، کاربرد دارد. از آن‌جا که ممکن است وارد کردن تاریخچه وقایع قبلی اثر متغیرهای پیش‌بین را در مقایسه با اثرات حاشیه‌ای آن‌ها ضعیف‌تر برآورد کند، زمانی‌که مطالعه اثر متغیر پیش‌بین بر نرخ حاشیه‌ای وقایع در نقطه مشخصی از زمان در سطح جامعه مورد نیاز باشد، این روش مفید است. این روش نسبت به ساختارهای همبستگی درون فردی بین وقایع استوار است؛ ولی در خصوص رابطه میان زمان‌های رخداد واقعه اطلاعی فراهم نمی‌آورد. شایان ذکر است این روش به پیش‌فرض مستقل بودن سانسورها بسیار حساس است و برای جلوگیری از اریبی، این پیش فرض باید بررسی گردد [۲۴]. هزینه مراقبت‌های پزشکی و عفونت‌های چندگانه در بیماران، مثال‌هایی از کاربرد این مدل در مطالعات پزشکی هستند [۸].

مدل‌های شرطی PWP

وقایع چندگانه مرتب شده را با استفاده از طبقه‌بندی روی تعداد وقایع قبلی در جریان مطالعه تحلیل می‌کنند و تابع خطر پایه نیز براساس این تعریف برای هر واقعه، مقدار متفاوتی دارد [۱۳]. کلیه افراد در معرض خطر اولین طبقه هستند، اما تنها کسانی‌که در طبقه قبلی در معرض خطر واقعه بوده‌اند وارد طبقه بعدی می‌شوند [۱۱]. به عبارت دیگر در این مدل فردی در معرض خطر واقعه jام قرار می‌گیرد که قبلاً (j-1) واقعه را تجربه کرده باشد [۳، ۶]. امتیاز مدل PWP نسبت به مدل WLW این است که نرخ واقعه بعدی برای یک فرد به شرطی که قبلاً واقعه‌ای را در طی دوره مطالعه تجربه کرده باشد، محاسبه می‌شود [۲۱، ۲۵]. شکل (۴) نحوه تعریف مجموعه خطر، تابع خطر پایه و ضرایب رگرسیونی را در این مدل نشان می‌دهد. همان‌گونه که ملاحظه می‌شود هر فرد زمانی در معرض یک واقعه قرار می‌گیرد که واقعه قبلی را تجربه کرده باشد. در این مدل برای هر طبقه، یک تابع خطر پایه مجزا (λ_{0j}) در نظر گرفته می‌شود و ضرایب رگرسیونی (β_j) برای هر طبقه متفاوت خواهد بود (شکل ۴. مجموعه خطر، تابع خطر پایه (λ_{0j}) و ضرایب رگرسیونی (β_j) در مدل PWP)

در این مدل می‌توان هر دو اثر کلی و ویژه- واقعه (Event-specific) را برای هر پیش‌بین به دست آورد. نکته‌ای که در استفاده از این مدل باید مورد توجه قرار گیرد این است که در این روش به دلیل شرطی شدن مجموعه خطر، تعداد افراد با افزایش تعداد وقایع (j)

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

کاهش می‌یابند. در عمل ممکن است مجموعه خطر برای طبقات بعدی بسیار کوچک باشد، بنابراین نیاز به محدود کردن تعداد معینی از وقایع بازگشتی است که در این صورت برآوردهای ویژه-واقعه بسیار غیرقابل اعتماد خواهند شد [۳]. در این مدل علاوه بر در نظر گرفتن پاسخی مشابه با مدل AG و مدل WLW (زمان کل، TT)، می‌توان پاسخ را براساس فاصله زمانی (GT) که فاصله میان رخداد واقعه فعلی و رخداد قبلی است تعریف کرد. زمانی‌که از مقیاس فاصله زمانی یا زمان انتظار (Waiting time) استفاده می‌شود، بعد از رخداد هر واقعه بازگشتی مقیاس زمانی با در نظر گرفتن فرایند بازگشتی (Renewal process) مجدداً صفر می‌شود. مطالعه فاصله زمانی اغلب برای وقایع با فراوانی رخداد کم، هنگامی که بررسی رخداد یک واقعه بعد از واقعه قبلی یا پیش‌بینی رخداد واقعه بعدی مد نظر باشد، مفید است. بنابراین مدل‌های PWP به دو شکل می‌توانند برازش یابند؛ مدل PWP-TT که اثر یک متغیر پیش‌بین را بر روی زامین واقعه از زمان ورود به مطالعه بررسی می‌کند و مدل PWP-GT که اثر یک متغیر کاکس است و بر روی تعداد وقایع طبقه‌بندی می‌شوند تا به این ترتیب وابستگی بین وقایع در مدل لحاظ گردد. ویژگی متمایز این مدل نسبت به مدل AG آن است که تابع خطر پایه می‌تواند از واقعه‌ای به واقعه دیگر تغییر کند و در نتیجه، مدل می‌تواند براساس وابستگی واقعه کنونی به وقایع قبلی تعدیل گردد. در هر دو مدل PWP، تابع خطر در زمان t برای واقعه زام بر روی کل وقایع قبلی شرطی می‌شود [۶]. برخلاف مدل AG، اثر پیش‌بین‌ها ممکن است از واقعه‌ای به واقعه دیگر در مدل‌های PWP تغییر نمایند. بنابراین مدل‌های PWP زمانی‌که اثر پیش‌بین‌ها در جریان رخداد وقایع تغییر می‌کنند، بیماری‌هایی که مراحل آن در طول مدت پیگیری یکسان نیستند (مانند برخی بیماری‌ها از جمله عفونت‌های ویروسی که به دلیل پیشرفت ایمنی بدن، ضعیف می‌شوند) یا خطر رخداد واقعه بعدی افزایش می‌یابد (مانند تشدید عفونت‌های ریوی) بر مدل AG ترجیح داده می‌شود.

یافته‌ها

در این بخش به بررسی دو مطالعه عود مجدد سرطان مثانه [۸] و رخداد دیسشارژ در بیماران صرعی [۲۶] که برای تحلیل آن‌ها از مدل‌های بازگشتی تصحیح واریانس استفاده شده است، پرداخته شد.

مطالعه عود سرطان مثانه: در مطالعه‌ای که به منظور بررسی تأثیر درمان بر عود مجدد سرطان مثانه انجام گرفت، ۸۵ بیمار انتخاب و به مدت ۶۴ ماه پیگیری شدند. کلیه بیماران بعد از خارج کردن تومور در مطالعه وارد و بیمارانی که قبل از اتمام مطالعه، آن‌را ترک کردند یا عودی را تجربه ننمودند به عنوان سانسور در نظر گرفته شدند. در این مطالعه از مدل‌های بازگشتی AG، WLW، PWP-TT و PWP-GT استفاده شد و این مدل‌ها با یکدیگر مقایسه گردیدند. متغیرهای پیش‌بین در این مطالعه نوع درمان (تیئوتپا (Thiotepa) و دارونما)، تعداد تومورهای اولیه و اندازه بزرگترین تومور (برحسب سانتیمتر) بودند. با استفاده از مدل کاکس در این داده‌ها تنها می‌توان تأثیر درمان بر عود اول را بررسی نمود و در نتیجه ۶۳٪ از وقایع نادیده گرفته خواهند شد. به منظور تحلیل صحیح این داده‌ها از مدل‌های بازگشتی استفاده گردید که نتایج آن در جدول (۱) ارائه شده است [۸] (جدول ۱. نتایج برازش مدل‌های بازگشتی در مطالعه عود سرطان مثانه).

نکته مهم در تحلیل داده‌ها توجه به تفسیر متفاوت پارامترها در هر یک از این مدل‌هاست. یافته‌های جدول (۱) حاکی از آن است که برآوردهای نقطه‌ای در هر دو مدل AG و WLW یکسان است، زیرا متغیرهای پیش‌بین وابسته به زمان در مدل AG وارد نشده‌اند؛ با این حال هر یک از این مدل‌ها نسبت‌های مختلفی را مدلسازی می‌کنند. مدل AG، تابع شدت (function Intensity) و مدل WLW نرخ واقعه را مدلسازی می‌نمایند. نتایج مدل AG حاکی از آن است که بیماران در گروه درمانی، کاهش ۳۷ درصدی در مخاطره عود مجدد دارند، اگرچه با تعدیل بر روی تعداد تومورهای اولیه، نوع درمان به صورت مرزی در این مدل معنی‌دار می‌شود. تعداد تومورهای اولیه، عامل تشخیصی مهم و معنی‌داری برای عود مجدد سرطان مثانه در همه مدل‌ها است. اندازه بزرگترین تومور با وجود تعداد تومورهای اولیه، در هیچ‌یک از ۴ مدل معنی‌دار نیست. برآوردهای حاصل از مدل‌های PWP براساس مجموعه خطر محدود شده به دست آمده‌اند، بدین معنی که مجموعه خطر تنها شامل افراد با تعداد مشابه عود مجدد در واقعه قبلی است. با توجه به این‌که هر یک از این مدل‌ها پیش‌فرض‌های مختلفی دارند نباید آن‌ها را مستقیم با یکدیگر مقایسه کرد. مهمترین دلیل برای انتخاب هر مدل به سئوالات پژوهش و مجموعه خطر بستگی دارد [۸].

مطالعه رخداد دیسشارژ (IED) در بیماران مبتلا به صرع: در این مطالعه، بیماران مبتلا به صرع که جهت بررسی بیماری

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

خود تحت مراقبت ۲۴ ساعته و V-EEG (Video Electroecephalogram) (روشی برای مانیتور کردن وضعیت بیماران) قرار داشتند، بررسی شدند. در نوار مغزی این بیماران نوعی IED ظاهر می‌شود که مورد توجه متخصصان مغز و اعصاب بوده و در تشخیص بیماری صرع و نوع آن نقش تعیین‌کننده‌ای دارد. از آن‌جا که IED می‌تواند در طول ثبت نوار مغزی چندین بار برای بیمار رخ دهد، لذا یک واقعه بازگشتی است. در این مطالعه از مدل‌های بازگشتی AG، PWP-TT و PWP-GT به منظور تشخیص عوامل مؤثر بر تعداد IEDها استفاده شد؛ زمان وقوع آن‌ها به عنوان پاسخ و متغیرهای سن و وضعیت جانباز بودن به عنوان پیش‌بین در نظر گرفته شد. جدول (۲) نتایج حاصل از برازش مدل‌های بازگشتی را نشان می‌دهد [۲۶]. با توجه به این‌که مدل AG شرطی نیست و در آن ضرایب به شرط وقایع قبلی برآورد نمی‌شوند، بنابراین باید تعداد وقایع قبلی در مدل وارد گردند. مدل PWP-TT یک مدل شرطی است که در آن اثر متغیرهای پیش‌بین بر واقعه فعلی به شرط وقایع قبلی محاسبه می‌شود. در این جدول ملاحظه می‌شود که در مدل AG سه عامل تعداد وقایع قبلی، سن و وضعیت جانبازی بیمار، عوامل مؤثر بر زمان کل وقوع IED هستند. در مدل PWP-TT دو عامل سن و وضعیت جانبازی بیمار از عوامل مؤثر بر زمان کل وقوع IED هستند. نتایج جدول (۲) نشان می‌دهد که خطر رخداد واقعه فعلی در مدل AG با افزایش تعداد وقایع قبلی افزایش می‌یابد. در هر دو مدل، سن رابطه معکوس با زمان کل وقوع IED دارد و نشان می‌دهد که در سنین بالاتر خطر کاهش زمان کل وقوع IED، افزایش می‌یابد. در مورد متغیر وضعیت جانبازی نیز هر دو مدل نتایج مشابهی را نشان می‌دهد به گونه‌ای که خطر کوتاه بودن فاصله زمانی وقوع IED در گروه جانبازان بیشتر است و در دو مدل تقریباً برابر ۲/۴ برابر گروه غیرجانباز می‌باشد.

در مدل PWP-GT نیز دو عامل سن و وضعیت جانبازی بیمار از عوامل مؤثر بر فاصله زمانی بین وقوع IEDها هستند. سن رابطه معکوس با فاصله زمانی رخداد IED ها دارد و نشان می‌دهد که به شرطی که IED قبلی رخ دهد، در سنین بالاتر خطر کاهش فاصله زمانی بین وقوع IEDها وجود دارد. در گروه جانبازان به شرطی که IED قبلی رخ دهد، خطر کوتاه بودن فاصله زمانی بین وقوع IEDها بیشتر است و تقریباً ۳ برابر گروه غیرجانبازان است [۲۶].

جدول ۱: نتایج برازش مدل‌های بازگشتی در مطالعه عود سرطان مثانه

	مدل‌های تصحیح واریانس			
	AG HR(95%CI)	WLW RR(95%CI)	PWP-TT HR(95%CI)	PWP-GT HR(95%CI)
نوع درمان (مرجع: دارونما)	۰/۶۳ (۰/۴۰، ۰/۹۹)	۰/۶۳ (۰/۳۸، ۱/۰۴)	۰/۷۲ (۰/۴۸، ۱/۰۵)	۰/۷۶ (۰/۵۱، ۱/۱۵)
تعداد تومورهای اولیه	۱/۱۹ (۱/۰۵، ۱/۳۵)	۱/۱۹ (۱/۰۵، ۱/۳۴)	۱/۱۹ (۱/۰۲، ۱/۲۴)	۱/۱۷ (۱/۰۶، ۱/۲۹)
اندازه بزرگترین تومور	۰/۹۶ (۰/۸۳، ۱/۱۱)	۰/۹۶ (۰/۸۳، ۱/۱۱)	۰/۹۶ (۰/۸۸، ۱/۱۲)	۱/۰۱ (۰/۸۹، ۱/۱۴)

HR : نسبت خطر، RR : خطر نسبی، CI : فاصله اطمینان

جدول ۲: نتایج برازش مدل‌های بازگشتی در مطالعه رخداد IED در بیماران مبتلا به صرع

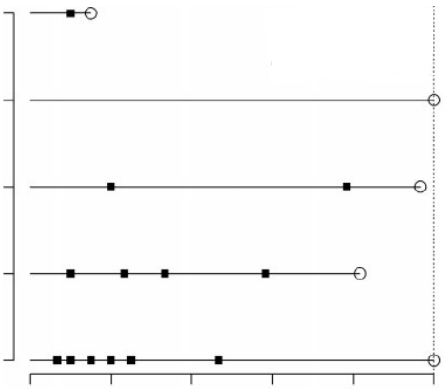
	متغیر پیش‌بین	برآورد ضریب	SE ضریب	P-value	نسبت خطر (HR)
AG	تعداد وقایع قبلی	۰/۱۷۶	۰/۰۵۹	۰/۰۰۲۹	۱/۱۹۲
	سن	-۰/۰۲۵	۰/۰۱۱	۰/۰۲۲۸	۰/۹۷۵
	وضعیت جانبازی	۰/۸۶۸	۰/۳۳۶	۰/۰۰۹۸	۲/۳۸۱
PWP-TT	سن	-۰/۰۳۱	۰/۰۱۴۲	۰/۰۳۱۴	۰/۹۷۰
	وضعیت جانبازی	۰/۸۶۸	۰/۳۵۳	۰/۰۱۳۸	۲/۳۸۳
PWP-GT	سن	-۰/۰۳۲	۰/۰۱۳	۰/۰۱۲۴	۰/۹۶۹
	وضعیت جانبازی	۱/۰۹۸	۰/۳۲۵	۰/۰۰۰۷	۲/۹۹۹

[Downloaded from payeshjournal.ir on 2025-07-06]

[DOR: 20.1001.1.16807626.1396.16.5.2.7]

جدول ۳. مقایسه مدل‌های تصحیح واریانس برای وقایع بازگشتی

مدل	مجموعه خطر برای واقعه jام در زمان t	مقیاس زمانی	مخاطره پایه (λ_0) برای وقایع مختلف	خطاهای استاندارد نیرومند	طبقه‌بندی براساس واقعه	سایر ویژگی‌ها
AG	وقایع را مستقل در نظر می‌گیرد.	دوره زمانی از شروع مشاهده (زمان کل)	مشترک	دارد	ندارد	- مدل کاکس حالت خاصی از این مدل است. - استقلال افزایشی وقایع فرض اصلی این مدل است که اغلب برقرار نیست؛ بدین معنی که رخداد یک واقعه، رخداد واقعه بعدی را کم یا زیاد نمی‌کند. - فرض می‌شود که تمام همبستگی‌های موجود میان وقایع از طریق پیش‌بین‌ها تبیین می‌شود، بنابراین به شرط پیش‌بین‌ها، مشاهدات مربوط به یک فرد از یکدیگر مستقلند. - یک پارامتر کلی برای عوامل مورد نظر برآورد می‌کند.
WLW	شامل کلیه افرادی که واقعه jام را در زمان t تجربه نکرده‌اند نیز می‌باشد.	دوره زمانی از شروع مشاهده (زمان کل)	متفاوت	دارد	دارد	- داده‌ها در این مدل براساس تعداد وقایع طبقه‌بندی می‌شوند، بنابراین λ_0های متفاوتی برای هر واقعه در نظر گرفته می‌شود. - برخلاف مدل AG، λ_0ها متفاوتند اما اثر پیش‌بین‌ها در هر طبقه یکسان است. - در خصوص رابطه میان زمان‌های رخداد واقعه اطلاعی فراهم نمی‌آورد. - نرخ واقعه بعدی برای یک فرد به شرطی که قبلاً واقعه‌ای را در طی دوره مطالعه تجربه کرده باشد، محاسبه می‌شود.
PWP-TT	شامل کلیه افرادی است که واقعه ۱-jام را تجربه و واقعه jام را در زمان t تجربه نکرده‌اند.	دوره زمانی از شروع مشاهده (زمان کل)	متفاوت	دارد	دارد	- همانند مدل WLW، وقایع طبقه‌بندی می‌شوند اما امتیاز این مدل ساخت طبقات براساس ترتیب وقایع است، بنابراین λ_0 های متفاوتی برای هر واقعه در نظر گرفته می‌شود. - با توجه به در نظر گرفتن ترتیب وقایع از مدل WLW کاراتر است. - در این مدل می‌توان هر دو اثر کلی و ویژه- واقعه را برای هر پیش‌بین به دست آورد. - مجموعه خطر، تعداد افراد با افزایش تعداد وقایع (j) کاهش می‌یابند.
PWP-GT	شامل کلیه افرادی است که واقعه ۱-jام را تجربه و واقعه jام را در زمان t تجربه نکرده‌اند.	دوره زمانی از واقعه قبلی مشاهده (فاصله زمانی)	متفاوت	دارد	دارد	- نرخ واقعه بعدی برای یک فرد به شرطی که قبلاً واقعه‌ای را در طی دوره مطالعه تجربه کرده باشد، محاسبه می‌شود. - همانند مدل WLW، وقایع طبقه‌بندی می‌شوند اما امتیاز این مدل ساخت طبقات براساس ترتیب وقایع است، بنابراین λ_0 های متفاوتی برای هر واقعه در نظر گرفته می‌شود. - با توجه به در نظر گرفتن ترتیب وقایع از مدل WLW کاراتر است. - در این مدل می‌توان هر دو اثر کلی و ویژه- واقعه را برای هر پیش‌بین به دست آورد. - مجموعه خطر، تعداد افراد با افزایش تعداد وقایع (j) کاهش می‌یابند.



شکل ۱. مثالی از رخداد وقایع بازگشتی برای ۵ فرد

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

شکل ۲. مجموعه خطر، تابع خطر پایه (λ_0) و ضرایب رگرسیونی (β) در مدل AG

شکل ۳. مجموعه خطر، تابع خطر پایه (λ_{0j}) و ضرایب رگرسیونی (β_j) در مدل WLW

شکل ۴. مجموعه خطر، تابع خطر پایه (λ_{0j}) و ضرایب رگرسیونی (β_j) در مدل PWP

بحث و نتیجه‌گیری

مدل‌های وقایع بازگشتی به درک روند بیماری‌ها کمک می‌کنند، بنابراین به کار بردن چنین داده‌هایی برای مطالعه روند بیماری‌ها حائز اهمیت است که برای تحلیل معتبر آن‌ها باید از روش‌های پیشرفته و مناسب آماری استفاده کرد. بررسی اثر عوامل خطر بر روی تعداد وقایع بازگشتی در طول زمان تنها با درنظر گرفتن واقعه اول، نتایج ناکارایی به دنبال خواهد داشت (۱۱، ۲۷). متأسفانه بسیاری از پژوهشگران با وجود محدودیت‌های روش‌های کلاسیک آماری و ایجاد روش‌های معتبر برای تحلیل این نوع داده‌ها، همچنان از روش‌های نادرست مدلسازی مانند رگرسیون لجستیک استفاده می‌کنند [۲۸، ۱۹]. روش‌های مختلفی در تحلیل بقا برای وارد کردن همبستگی درون- فردی که از رخداد چندین واقعه برای یک فرد ایجاد می‌شود، پیشنهاد شده‌اند. انتخاب مدل مناسب از میان مدل‌های موجود به توزیع زمینه‌ای زمان‌های رخداد واقعه، همبستگی بین وقایع درون فردی، تعداد وقایع رخ داده و هدف مطالعه که بررسی اثر متغیرهای پیش‌بین به شرط سابقه فرد، متوسط اثر جامعه یا اندازه‌گیری وابستگی وقایع یک فرد است، بستگی دارد. همچنین فرآیند بیولوژیکی بیماری نیز پایه‌ای در انتخاب مدل برای زمان رخداد وقایع فراهم می‌کند. با توجه به عدم توافق در زمینه انتخاب مدل مناسب برای تحلیل وقایع بازگشتی بقا، در این مقاله انواع مدل‌های تصحیح واریانس بررسی شدند. مهمترین تفاوت میان مدل‌های بازگشتی، نحوه تعریف مجموعه خطر و همچنین ثابت یا متغیر فرض کردن خطر پایه است. در مدل AG کلیه افراد در دوره مطالعه در معرض خطر مجدد رخداد واقعه قرار می‌گیرند. اگر فرض ثابت ماندن مخاطره رخداد مجدد واقعه بدون توجه به وقایع قبلی برقرار باشد، همچنین اگر بتوان خطر زمینه‌ای مشترکی را در فرآیندی مانند بیماری، در نظر گرفت و تنها برآورد نرخ کلی از عود مجدد مد نظر باشد، مدل AG پیشنهاد می‌شود [۱۵]. در این حالت توجه به منطقی بودن در نظر گرفتن خطرات پایه مشترک بسیار مهم است. این مدل فرض می‌کند که زمان‌های میان وقایع به شرط پیش‌بین‌ها غیرهمبسته هستند و حذف یک متغیر پیش‌بین مهم می‌تواند وابستگی را کاهش دهد. در این حالت خطاهای استاندارد کم‌برآورد می‌شوند که منجر به تورم خطای نوع اول می‌گردد. یک روش برای اصلاح این مدل برازش آن با پیش‌بین‌های وابسته به زمان برحسب تعداد وقایع است. امتیاز مدل AG قابلیت آن در وارد نمودن پیش‌بین‌های وابسته به زمان و فواصل غیرپیوسته مخاطرات است [۲۹].

اگر این فرض که رخداد اولین واقعه احتمال رخداد مجدد واقعه را افزایش می‌دهد برقرار باشد یا هرگاه وابستگی به وقایع گذشته در افراد قوی باشد و زمانی‌که برآورد اثرات برای هر واقعه به صورت جداگانه مد نظر است، استفاده از مدل PWP با تابع خطر پایه غیرمشترک مناسب است. این مدل‌ها فرض می‌کنند که افرادی می‌توانند در معرض خطر واقعه فعلی قرار گیرند که واقعه قبلی را تجربه کرده باشند. از آن‌جا که در این مدل‌ها مجموعه خطر برای

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

وقایع بعدی کوچک می‌شود، برآوردهای غیرپایدار (Unstable) تولید می‌کنند که از محدودیت‌های این مدل‌ها است و برای رفع آن، اغلب باید داده‌ها را برید (Truncate) [۶].

در مدل WLW کلیه افراد در طول دوره مطالعه، بدون توجه به رخداد وقایع قبلی، در معرض خطر رخداد مجدد وقایع قرار می‌گیرند اما تابع خطر پایه برای هر واقعه به‌صورت جداگانه در مدل وارد می‌شود. زمانی‌که هدف مطالعه مدلسازی تعداد مورد انتظار وقایع یا نرخ رخداد مجدد آن‌ها به شرط متغیرهای پیش‌بین باشد از مدل WLW استفاده می‌شود. افزودن پیش‌بین‌های وابسته به زمان به این مدل‌ها، منجر به تفسیرهای مختلفی با توجه به روش به کار رفته می‌شود. با این‌که اغلب مقادیر برآوردی برای دو مدل AG و WLW با یکدیگر مشابه هستند، این مدل‌ها نسبت‌های متفاوتی را مدلسازی می‌کنند؛ مدل‌های AG تابع شدت و مدل WLW نرخ رخداد وقایع را مدلسازی می‌نمایند.

تعداد وقایع در داده‌ها تأثیر بسیاری در انتخاب مدل مناسب دارد، اگر تعداد وقایع کم باشد می‌توان فرض کرد که خطر رخداد هر یک از وقایع با یکدیگر تفاوت دارد و در نتیجه مدل‌های PWP مناسب هستند، ولی وقتی تعداد رخداد وقایع در فرد زیاد باشد خطر، کمتر بین وقایع تغییر می‌کند و استفاده از مدل AG مناسب است.

وقتی هدف مطالعه برآورد اثر کلی یک متغیر پیش‌بین یا اثر آن روی هر واقعه به صورت جداگانه باشد، مدل WLW مناسب نیست زیرا فرد در معرض خطر بیش از یک واقعه است و در این حالت مدل‌های PWP و AG مناسب هستند. در کل با توجه به ویژگی‌های مدل WLW از جمله مجموعه خطر آن، این مدل برای تحلیل اغلب داده‌های وقایع بازگشتی مناسب نیست؛ به عنوان مثال افراد در معرض خطر واقعه چهارم، حتی اگر تنها یک واقعه را تجربه کرده باشند، می‌توانند در مجموعه خطر قرار گیرند که این مسئله منجر به بیش برآورد اثر درمان می‌گردد. به دلیل این ویژگی، تفسیر ضرایب این مدل برای وقایع بازگشتی نیز مشکل است، زیرا ضریب یک متغیر پیش‌بین نشان دهنده اثر متوسط آن در سطح جامعه بر زمان رخداد وقایع است. این مدل بیشتر برای داده‌هایی مناسب است که یک فرد می‌تواند انواع مختلف واقعه را تجربه کند. همچنین چون در برازش این مدل از اطلاعات بخش اولیه مشاهدات استفاده می‌گردد لذا وزن بیشتری به وقایع اولیه می‌دهد. به عنوان مثال اگر مداخله‌ای در پیشگیری از واقعه اول مؤثر باشد ولی بر روی وقایع بعدی اثر نداشته باشد، مدل WLW نشان می‌دهد که این مداخله بر روی همه وقایع مؤثر است، زیرا اگر درمان، واقعه اول را به تأخیر بیاندازد در نتیجه وقایع بعدی نیز خود به خود عقب می‌افتند و حتی اگر درمان مؤثر نباشد مدت رخداد واقعه آن بیشتر می‌شود [۶].

جدول (۳)، به صورت خلاصه ویژگی‌های هر یک از مدل‌های تصحیح واریانس را نشان می‌دهد. با توجه به مباحث مطرح شده، اهداف و سئوالات پژوهش و تعریف مجموعه خطر نقش بسزایی در انتخاب مدل دارند، زیرا براساس آن‌ها نحوه وارد کردن وابستگی میان وقایع در مدل و در نتیجه تفسیرها متفاوت خواهند بود.

سهم نویسندگان

مهسا سعادتی: مجری طرح

آرزو باقری: همکار اصلی طرح

تشکر و قدردانی

این مقاله مستخرج از طرح «روش‌های تحلیل سابقه رخداد وقایع در مطالعات جمعیتی» (ابلاغ شماره ۱۱/۲۴۳۳۲ مورخ ۹۵/۰۵/۳۰) است که با حمایت مالی مؤسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور در سال ۱۳۹۵ انجام شده است.

منابع

1.Abrahantesa JC, Burzykowskia T, Legrandb C, Janssena P, Ducrocqc V, Duchateaud L. Comparison of Different Estimation Procedures for Proportional Hazards Model with Random Effects. Computational Statistics & Data Analysis 2007; 51: 3913 - 30

2.Manda S, Meyer R. Bayesian Inference for Recurrent Events Data Using Time-Dependent Frailty. Statistics in Medicine 2005; 24: 1263-74

3.Kelly PJ, Lim LLY. Survival Analysis for Recurrent Event Data: An Application to Childhood Infectious Diseases. Statistics in Medicine 2000; 19: 13-33

4.Pennell ML, Dunson DB. Bayesian Semi parametric Dynamic Frailty Models for Multiple Event Time Data. Biometrics 2006; 62: 1044–52

5.Yau KKW, McGilchrist CA. ML and REML Estimation in Survival Analysis with Time Dependent Correlated Frailty. Statistics in Medicine 1998; 17: 1201-13

6.Lim HJ, Liu J, Melzer-Lange M. Comparison of Methods for Analyzing Recurrent Events Data: Application to the Emergency Department Visits of

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

Pediatric Firearm Victims**.** Accident Analysis and Prevention 2007; 39: 290–99

7.Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. International Journal of Epidemiology 2015; 144:324-33

8.Amorim LD, Cai J, Zeng D, Barreto ML. Regression Splines in the Time-Dependent Coefficient Rates Model for Recurrent Event Data. Statistics in Medicine 2008; 27: 5890-906

9.Cox DR. Regression models and life-tables (with Discussion). Journal of Royal Statistics Social 1972; 34:182–220

10. Du P**.** Nonparametric Modeling of the Gap Time in Recurrent Event Data**.** Lifetime Data Analysis 2009; 15: 256-77

11. Pandeya N, Purdie DM, le Green A, Williams G. Repeated occurrence of basal cell carcinoma of the skin and multifailure survival analysis: follow-up data from the Nambour Skin Cancer Prevention Trial. American Journal of Epidemiology 2005; 161:748–54

12. Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. Annals Statistics 1982; 10:1100–20

13. Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. Biometrika 1981; 68: 373–79

14. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of American Statistic Association Assoc 1989; 84:1065–73

15. Therneau TM, Grambsch PM. Modelling Survival Data: Extending The Cox Model. 2st Edition, Springer: New York, 2000

16. Pepe MS, Cai J. Some graphical displays and marginal regression analysis for recurrent failure times and time dependent covari-ates. Journal of American Statistic Association 1993;88:811–20

17. Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of Royal Statistics Social 2000; 62:711–30

18. Kennedy BS, Kasl SV, Vaccarino V. Repeated hospitalizations and self-rated health among the elderly: a multivariate failure time analysis. Journal of Epidemiology 2001; 153: 232–41

19. Gill DP, Zou GY, Jones GR, Speechley M. Comparison of regression models for the analysis of fall risk factors in older veterans. Annals Epidemiology 2009; 19:523–30

20. Moulton LH, Dibley MJ. Multivariate time-to-event models for studies of recurrent childhood diseases. International Journal of Epidemiology 1997; 26: 1334–39

21. Andersen P.K, Borgan Ø, Gill, R.D, Keiding N. Statistical Models Based on Counting Processes. 1st Edition, New York: Springer-Verlag, 1993

22. Lawless JF. The analysis of recurrent events for multiple subjects. Journal of Clinical Epidemiology 1995; 44:487–98

23. Cai J, Schaubel D. Analysis of recurrent event data. In: Balakrishnan N, Rao CR (eds). Handbook of Statistics: Advances in Survival Analysis. Vol. 23. Amsterdam: Elsevier, 2004

24. Strawderman RL. The Accelerated Gap Times Model. Biometrika 2005; 92: 647-66

25. Cui J, Forbes A, Kirby A, Marschner I, Simes J, West M, Tonkin A. Parametric Conditional Frailty Models for Recurrent Cardiovascular Events in the Lipid Study**.** Clinical Trials 2008**;** 5**:** 565–574

26. Hosseinzadeh S. Modeling gap times between recurrent events by time dependent frailty model with application to medical data, Ph.D Thesis; Tarbiat modares University, 2011

27. Dacourt V, Quantin C, Abrahamowicz M, Blinquet C, Alioum A, Faivre J. Modelling recurrence in colorectal cancer. Journal of Clinical Epidemiology 22004; 57:243–51

28. Purroy F, Caballero PEJ, Gorospe A et al. Recurrent transient is chaemic attack and early risk of stroke: data from the PROMAPA Study. Journal of Neurology, Neurosurgery, and Psychiatry 2013; 84: 596–603

29. Castan˜ eda J, Gerritse B. Appraisal of several methods to model time to multiple events per subject: modelling time to hospitalizations and death. Revista Colombiana de Estadística 2010; 33:43–61

[Downloaded from payeshjournal.ir on 2025-07-06]
[DOR: 20.1001.1.16807626.1396.16.5.2.7]

ABSTRACT

Variance-corrected recurrent models in medical studies

Arezoo Bagheri [1], Mahsa Saadati [1*]

1. National Population & Comprehensive Management Institute, Tehran, Iran

Payesh 2017; 5: 575- 585
Accepted for publication: 3 September 2017
[EPub a head of print-12 September 2017]

Objective (s): In many medical and epidemiological studies, each person can experience several events one after the other sequentially under some circumstances, known as recurrence events. Broad analytical objectives of recurrent events consist of describing event process in an individual, dispersion of events from one person to another, and determination of relations between constant or time-dependent predictors with event time. So the main purpose of this article is to present and interpret advanced statistical models to analyze such events which have been proposed in recent years.
Methods: In this article, variance-corrected models including AG, conditional PWP (PWP-TT and PWP-GT) and marginal WLW models for analyzing recurrent events are presented and compared with each other.
Results: When the number of events for each individual are small and the risk of a reoccurrence of an event varies between different events, usually conditional PWP models are used, while marginal WLW and the time-dependent AG models are applied to determine frequency of events.
Conclusion: Medical and epidemiology studies should consider various factors such as the number of events, relationship between events, the variability of the effect of factors from one event to another, the biological process, and the correlation structure of data to select the valid recurrent models.

Key Words: Recurrent events, AG model, Conditional PWP model, Marginal WLW model

[Downloaded from payeshjournal.ir on 2025-07-06]

[DOR: 20.1001.1.16807626.1396.16.5.2.7]

* Corresponding author: National Population & Comprehensive Management Institute, Tehran, Iran
Tel: 88534321
E-mail: mahsa.saadati@gmail.com