نقشه‌بندی تغییرات مکانی بیماری‌ها با استفاده از مدل‌های کلاسیک و بیزی
مدل‌های نقشه‌بندی تغییرات مکانی بیماری‌ها

یدالله محرابی [۱]، بهزاد مهکی [۲]، امیر کاوسی [۳]

مقاله مروری

چکیده

نقشه‌بندی بیماری شامل مجموعه‌ای از روش‌های آماری است که به تهیه نقشه‌های دقیق براساس برآوردهایی از میزان‌های بروز، شیوع و مرگ و میر منجر می‌شوند تا کاربران بتوانند با توجه به توزیع جغرافیایی بیماری‌ها، برنامه‌ریزی‌های بهداشتی لازم را با دقت بیشتری انجام دهند. اهداف اصلی نقشه‌بندی بیماری عبارت هستند از: تشریح تغییرات فضایی در شیوع بیماری به منظور فرمول‌بندی فرضیات علت‌شناسی، ارزیابی خطر و تعیین نواحی دارای خطر بالا که نیازمند جلب توجه و اعمال مداخله هستند، تهیه نقشه دقیقی از خطر بیماری در نواحی به منظور اختصاص بهتر منابع و ساخت اطلس بیماری‌ها. تاکنون مطالعات موردی متعددی انجام گرفته است که در آن از کابردهای مدل‌های نقشه‌بندی بیماری‌ها استفاده گردیده است. با این وجود جز در چند مطالعه اندک و به صورت محدود، روش‌های متعدد تحلیل نقشه‌های بیماری‌ها با هم مقایسه نشده و شرایط کلی آن‌ها مورد بررسی قرار نگرفته است. با توجه به اهمیت موضوع و کاربرد رو به گسترش نقشه‌بندی بیماری‌ها، در مطالعه حاضر مهم‌ترین مدل‌های نقشه‌بندی بیماری‌ها معرفی و مبانی تئوری، کاربردها، محاسن و معایب هر یک به اختصار بیان گردیده و از جنبه‌های مختلف مورد مقایسه قرار گرفته‌اند.

ارجاع: محرابی یدالله، مهکی بهزاد، کاوسی امیر. **نقشه‌بندی تغییرات مکانی بیماری‌ها با استفاده از مدل‌های کلاسیک و بیزی مدل‌های نقشه‌بندی تغییرات مکانی بیماری‌ها.** مجله تحقیقات نظام سلامت ۱۳۹۳؛ ۱۰(۲): ۲۱۱-۲۰۱

تاریخ دریافت: ۱۳۹۲/۰۴/۲۰ **تاریخ پذیرش: ۱۳۹۲/۰۵/۲۸**

۱. گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران
۲. گروه آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران (نویسنده مسؤول)
Email:behzad.mahaki@gmail.com
۳. گروه علوم پایه، دانشکده سلامت، ایمنی و محیط زیست، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

مقدمه

نقشه‌بندی بیماری شامل مجموعه‌ای از روش‌های آماری است که منجر به تهیه نقشه‌های دقیق براساس برآوردهایی از میزان‌های بیماری‌ها می‌شوند تا کاربران بتوانند با اطمینان توزیع جغرافیایی بیماری موردنظر را برآورد کنند (۲-۱). در سالیان اخیر موضوع نقشه‌بندی بیماری‌ها به شکل قابل ملاحظه‌ای گسترش یافته و کاربردهای متنوعی پیدا کرده است. مهم‌ترین اهداف نقشه‌بندی بیماری عبارت هستند از:

۱- تشریح تغییرات در میزان‌های شیوع، بروز و مرگ و میر بیماری‌ها
۲- پیشنهاد عوامل خطر ممکن به منظور فرمول‌بندی فرضیات علت‌شناسی
۳- تعیین نواحی دارای خطر بالا و نیازمند جلب توجه و اعمال مداخله
۴- اختصاص بهتر منابع و خدمات بهداشتی و درمانی
۵- ساخت اطلس بیماری‌ها (۵-۳)

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

دو حالت رایج در مطالعات توزیع جغرافیایی بیماری‌ها وجود دارد: ۱- ناحیه‌ای تعریف می‌گردد و مکان موارد یک بیماری خاص، درون این ناحیه در یک دوره زمانی ثابت، ثبت می‌گردد. این مکان‌ها معمولا آدرس‌های محل سکونت هستند. پس از مشخص شدن این آدرس‌ها، ادامه کار با تحلیل مستقیم مکان موارد امکان‌پذیر است. این حالت را تحلیل وقایع موردی (case-event analysis) می‌نامیم. اغلب این تحلیل نیازمند استفاده از مدل‌های روند نقطه‌ای و متدولوژی مربوط است (۸- ۶). ۲- در اغلب موارد به دلیل لزوم محرمانه بودن برخی از مسایل در پزشکی، امکان دستیابی به داده‌ها در مقیاس و سطح بالا وجود ندارد و باید دسته‌بندی مجدد برای تحلیل تعداد موارد درون نواحی کوچک مطالعه انجام شود. این نواحی کوچک، مناطقی اختیاری مانند شهرداری، منطقه، ناحیه و محله هستند که معمولا برای اهداف اجرایی و مدیریتی تعریف می‌شوند. تحلیل این نوع داده‌ها که شامل تعداد موارد در نواحی کوچک هستند را تحلیل تعداد موارد در نواحی (tract count analysis) می‌نامیم. تعداد موارد درون یک ناحیه از تجمیع همه موارد درون آن ناحیه به دست می‌آید. بدین ترتیب مکان‌های جداگانه‌ی موارد از دست رفته و اطلاعات تعداد موارد را تنها می‌توان از مکان ناحیه استخراج کرد. اغلب نقطه میانی ناحیه، برای معرفی مکان ناحیه در نظر گرفته می‌شود. در یک ناحیه انتخاب شده برای مطالعه، m ناحیه را در نظر می‌گیریم. مقادیر درون m ناحیه را با (y_i i=1,2,…,m) نشان می‌دهیم.

تاکنون مطالعات موردی متعددی انجام گرفته است که در آن از کابردها و مدل‌های نقشه‌بندی بیماری‌ها استفاده گردیده است (۹-۱۵). با این وجود جز در چند مطالعه اندک و به صورت محدود، روش‌های متعدد تحلیل نقشه‌های بیماری‌ها با هم مقایسه نشده و شرایط کلی آن‌ها مورد بررسی قرار نگرفته است. با توجه به اهمیت موضوع و کاربرد رو به گسترش نقشه‌بندی بیماری‌ها، در مطالعه حاضر مهم‌ترین مدل‌های نقشه‌بندی معرفی گردیده و مبانی تئوری، کاربردها، محاسن و معایب هر یک به اختصار بیان می‌گردد.

روش‌ها

در این بخش مهم‌ترین مدل‌های نقشه‌بندی بیماری‌ها معرفی و توضیحات لازم ارایه خواهد شد.

۱- مدل حداکثر درست‌نمایی

برای ارزیابی وضعیت یک ناحیه از نظر میزان‌های بروز و ابتلای بیماری، به عنوان گام نخست مناسب است که ارزیابی کنیم چه میزان بروزی برای آن بیماری در آن ناحیه موردانتظار است، سپس موارد مشاهده شده را با بروز مورد انتظار مقایسه کنیم. نسبت مقادیر مشاهده شده به مقادیر مورد انتظار مرگ و میر، نسبت مرگ و میر استاندارد شده (SMR یا Standardized Mortality Ratio) نامیده می‌شود. در مطالعات مربوط به ابتلا و بروز بیماری، از نسبت بروز استانداردشده (SIR یا Standardized Incidence Ratio) استفاده می‌کنیم. این نسبت‌ها برآورد حداکثر درست‌نمایی خطر نسبی در هر ناحیه هستند (۲-۱).

برای به‌کارگیری این روش، نقشه به n ناحیه مجاور نامتداخل تقسیم می‌شود (i=1,2,…,n). تعداد رویداد مورد نظر مشاهده شده و مورد انتظار در ناحیه iام به ترتیب با O_i و E_i نشان داده می‌شود که E_i طی دوره مطالعه ثابت و معلوم فرض شده و از حاصل‌ضرب بروز کلی در جمعیت هر ناحیه به دست می‌آید. خطر نسبی بیماری در ناحیه iام نیز با θ_i نشان داده می‌شود.

در نقشه‌بندی بیماری، فرض می‌شود که تعداد رویدادها، بین نواحی، مستقل از هم بوده و از توزیع پوآسن با میانگین $E_i\theta_i$ پیروی می‌کند. بنابراین با به در نظر گرفتن درست نمایی O_i ، لگاریتم گرفتن و برابر صفر قرار دادن آن، در نهایت برآورد حداکثر درست‌نمایی خطر نسبی عبات خواهد بود از:

$$\hat{\theta} = SMR_i = \frac{O_i}{E_i}$$

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

گرچه استفاده از SMR و SIR در نقشه‌بندی بیماری‌ها رایج است اما نقاط ضعفی در استفاده از این شاخص‌ها وجود دارد. از جمله‌ی این نقاط ضعف می‌توان به موارد زیر اشاره کرد.

۱- مبتنی بودن این شاخص‌ها بر برآورد نسبتی و در نتیجه تغییرات بزرگ در برآورد با اندک تغییرات در مقادیر مورد انتظار

۲- بسیار بزرگ بودن مقادیر این شاخص‌ها در صورت نزدیک بودن مقادیر مورد انتظار به صفر

۳- نشان ندادن تفاوت‌های موجود در مقادیر مورد انتظار در صورت صفر بودن مقادیر این شاخص‌ها

۴- متناسب بودن واریانس این شاخص‌ها با عکس مقدار مورد انتظار (کوچک بودن در نواحی با جمعیت بزرگ و بزرگ بودن در نواحی با جمعیت کم) و در نتیجه مشکل بودن مسأله تصمیم‌گیری براساس این شاخص‌ها

۵- برآورد اشباع برای خطر نسبی بودن و در نتیجه خلاف اصل امساک بودن استفاده از آن

از آن‌جایی که این برآوردگر دارای عیوب مذکور است می‌توان از مدل‌های دیگر برای غلبه بر این مشکلات استفاده کرد. در این راستا روش‌های متعددی پیشنهاد شده‌اند که در ادامه به این روش‌ها پرداخته می‌شود.

۲- مدل‌های هموارسازی

یک رویکرد برای بهبود برآورد خطر نسبی، استفاده از ابزارهای هموارسازی SMR و SIR به منظور کاهش اختلالات است. این ابزارها می‌توانند مبتنی بر روش‌های میان‌یابی (interpolation methods)، هموارسازهای ناپارامتری مانند رگرسیون کرنل یا روش‌های جداسازی (partition methods) باشند. چندین روش تحلیل اکتشافی داده‌ها نیز در این خصوص معرفی گردیده‌اند. عموما این روش‌ها نیازمند ثابت هموارسازی هستند که تغییرات و رفتار کلی خطر نسبی را تشریح می‌کند. برخی روش‌های موضعی نیز پیشنهاد شده‌اند که دارای حسن امکان ورود متغیرهای مستقل هستند. مدل‌های جمعی تعمیم‌یافته نیز معرفی گردیده‌اند که امکان وارد نمودن متغیرهای کمکی را فراهم می‌آورند.

Simonoff رویکردی را برای تحلیل اکتشافی داده‌ها معرفی کرده است که اغلب در مسایل تک متغیره به کار می‌رود و در جهت هموارسازی نوسانات و اختلالات تلاش می‌کند (۱۶). رایج‌ترین روش در این کلاس، روش هموارسازی کرنل نادارایا- واتسون است که هنگام استفاده برای SMR، میانگین وزنی از SMRها (φها) است: $\hat{\theta} = \sum_{j \neq i} w_j \varphi$ که در آن $w_j = \frac{k((\varphi_i - \varphi_j)/h)}{\sum_i k((\varphi_i - \varphi_j)/h)}$ توابعی از مقادیر همسایگی‌ها، K تابع چگالی احتمال متقارن با میانگین صفر و h برای مینیمم کردن برخی از معیارهای نکویی برازش انتخاب می‌شود. رویکردهای وزن‌دهی دیگری نیز وجود دارند اما ویژگی‌های آماری این برآوردگر را ندارند. ثابت هموارسازی h را می‌توان با استفاده از وارسی اعتبار (cross-validation) انتخاب کرد.

رویکرد دیگر هموارسازی، هموارسازی جداگانه صورت و مخرج SMR کسر است. یک روش توسط Kelsall و Diggle برای برآورد خطر نسبی مکان و آدرس رویدادها پیشنهاد شده است (۱۸-۱۷).

۳- مدل‌های بیز خطی

مبتنی بر توابع خطی از SMR هستند. به‌کارگیری این روش‌ها در نقشه‌بندی بیماری‌ها، نخستین بار توسط Marshall بیان شده است (۱۱ و ۱۹). ثابت شده است که برآورد بیز خطی $\{\theta_i\}$ براساس SMR (φ_i) عبارت است از:

$$\hat{\theta}_i = E_\Theta\{\theta_i\} + \frac{\mathrm{cov}(\theta_i, \varphi_i)}{\mathrm{var}_\Theta(\varphi_i)}[\varphi_i - E_\Theta\{\theta_i\}]$$

که $E_\Theta\{\theta_i\}$ امید $\{\theta_i\}$ بر اساس توزیع پیشین محاسبه شده است. این برآوردگر بهترین برآوردگر به فرم $\hat{\theta}_i = \alpha_i + \beta\varphi_i$ است. این معادله منجر به برآوردگر به فرم

$$\hat{\theta}_i = m_i + \frac{v_i}{(m_i / v_i) + v_i}\{\varphi_i - m_i\}$$

می‌شود که در آن

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

μ_i و σ_i^2 به ترتیب میانگین و واریانس توزیع پیشین خطر نسبی هستند

Marshall دو نوع از این برآوردگر را پیشنهاد کرده است. برآورد کلی که در آن میانگین و واریانس ثابت فرض شده ($\hat{\mu}_i = \hat{\mu}$ و $\hat{\sigma}_i^2 = \hat{\sigma}^2$)و از روش گشتاوری به دست می‌آید و برآورد موضعی برآورد میانگین و واریانس پیشین صرفا با استفاده از نواحی مجاور به دست می‌آید.

۴- مدل‌های بیز تجربی

این مدل‌ها مشابه مدل‌های بیز کامل بوده اما توزیع پیشین خطرات نسبی از داده‌های مشاهده شده برآورد می‌شود (۲۱-۲۰). به عبارت دیگر این روش‌ها بینابین روش حداکثر درست‌نمایی و روش‌های بیز کامل هستند. در این مدل‌ها فرض می‌شود که خطرات نسبی (θ_i ها) از توزیع خاصی مانند ($f(\theta_i)$)- که با ابرپارامترهای γ پارامتری می‌گردد- پیروی می‌کنند. توزیع پسین حاشیه‌ای خطرات نسبی به شرط داده‌های مشاهده شده و ابرپارامترهای γ عبارت است از:

$$g(\theta|O,\gamma) \propto g(O|\theta) f(\theta|\gamma)$$

این توزیع به عنوان تقریبی برای توزیع پسین حاشیه‌ای خطرات نسبی به شرط داده‌های مشاهده شده $g(\theta|O)$ با به کارگیری برآوردهای ابرپارامترها استفاده می‌شود. معمولا ابرپارامترها از داده‌های مشاهده شده، با حداکثر نمودن درست‌نمایی حاشیه ای γ

$$g(O|\gamma) = \int g(O|\theta) f(\theta|\gamma) d\theta$$

برآورد می‌شوند. حداکثر نمودن را می‌توان از طریق الگوریتم EM یا از طریق تقریبات تحلیلی برای انتگرال از قبیل تقریب لاپلاس انجام داد. مهم‌ترین مدل‌هایی که در چارچوب مدل‌های بیز تجربی قرار دارند مدل‌های گاما- پوآسن و لگ نرمال هستند که توسط Clayton و Kaldor معرفی گردیده‌اند (۱۰).

در مدل گاما- پواسن فرض بر این است که تعداد مرگ‌ها در هر ناحیه دو به دو مستقل بوده و از توزیع پوآسن با پارامتر $E_i\theta_i$ پیروی می‌کنند. توزیع پیشین برای خطرات نسبی $Gamma(a,b)$ در نظر گرفته می‌شود. بدین ترتیب خطر نسبی دارای توزیع پسین $Gamma(a+O_i, b+E_i)$ با میانگین

$$E[\theta_i|O_i,a,b] = \frac{a+O_i}{b+E_i} = w_i SMR_i + (1-w_i)\frac{a}{b}$$

$$w_i = \frac{E_i}{b+E_i}$$

است که در آن a و b پارامترهای توزیع پیشین گاما هستند. بنابراین میانگین پسین برای iامین ناحیه، میانگین وزنی SMR برای آن ناحیه و خطر نسبی بیماری در نقشه کلی است که وزن به طور معکوس با واریانس SMR در ارتباط است. اگر مقادیر مشاهده شده و مقادیر مورد انتظار هر دو بزرگ باشند آنگاه برآوردگر به سمت SMR میل می‌کند. اما اگر مقادیر مشاهده شده و مقادیر مورد انتظار هر دو کوچک باشند آنگاه به سمت میانگین پیشین (O_i / E_i) میل می‌کند (۲۱-۲۰، ۱۰) (۱۰، ۲۰ و ۲۱). با وجود این که مقادیر برآورد شده را میتوان از برآوردگر بیز خطی کامل به دست آورد، مدل گاما- پوآسن دارای چند حسن است. این مدل پسین کامل را برآورد می‌کند که می‌تواند برای محاسبه فواصل اطمینان و انجام آزمون‌های فرض استفاده شود درحالی که روش بیز خطی چنین قابلیتی را ندارد. حتی اگر توزیع واقعی خطر نسبی گاما باشد مدل گاما - پوآسن میانگین و واریانس این توزیع را از طریق حداکثر درست‌نمایی برآورد می‌کند که نسبت به روش گشتاوری استفاده شده در مدل‌های بیز خطی برتر و ارجح است (۲۱-۲۰، ۱۰) . عیب اساسی این روش نیز این است که این مدل قادر به درنظرگرفتن همبستگی فضایی نیست (۲۰).

توزیع پیشین گاما برای خطر نسبی گرچه از منظر ریاضیاتی مناسب است، می‌تواند محدودکننده باشد زیرا ورود متغیرهای کمکی به آن مشکل بوده و امکان درنظرگرفتن همبستگی فضایی در نواحی نزدیک‌تر وجود ندارد (۱). در مقابل مدل لگ

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

نرمال برای خطرات نسبی بیماری انعطاف پذیرتر از مدل گاما- پوآسن است:

$$y_i \sim poisson(E_i\theta_i)$$
$$\log\theta_i = \alpha + v_i$$
$$v_i \sim N(0,\tau_v^2)$$

در این مدل v_i برای در نظر گرفتن همبستگی فضایی و مشابهت وضعیت نواحی مجاور هم در مدل وارد شده است.

۵- مدل‌های آمیخته

مدل‌هایی که تاکنون معرفی گردیده‌اند اساسا مبتنی بر تکنیک‌های هموارسازی هستند. این مدل‌ها اغلب به هموارسازی ناپیوستگی‌ها در خطرات می‌پردازند که حایز اهمیت است. چندین روش تشخیص ساده‌ی ناپیوستگی‌ها معرفی گریده‌اند. گرچه این مدل‌ها موضوع عدم پیوستگی را مخاطب قرار داده‌اند اما امکان درنظر گرفتن همزمان هموارسازی و عدم پیوستگی در یک نقشه را فراهم نمی‌آورند. بدین منظور می‌توان مدل‌های آمیخته‌ی فضایی که اشکال متفاوت تغییرات فضایی را ارایه می‌دهند را به کار گرفت (۲۲). می‌توان یک مدل ساخت که توزیع حاشیه‌ای مقادیر، توزیعی آمیخته شامل مجموعه‌ای از اجزاء گسسته درنظر گرفته شود. این رویکرد کاملا از مدل‌های پیشین که در آن پراکندگی و تغییرات در خطر با مخلوطی از چند جزء مدل‌بندی می‌شود، متفاوت است. این مدل K سطح متفاوت از خطر را درنظر می‌گیرد $\{\theta_1,\theta_2,...,\theta_k\}$ و مقادیر مشاهده شده را از ناحیه‌ای با خطر θ_j و احتمال p_j فرض می‌کند. بدین ترتیب درستنمایی عبارت است از:

$$L=\prod_{i=1}^{n}\sum_{j=1}^{k} p_j\, pois(O_i;E_i,\theta_i)$$

در این حالت جواب به فرم بسته برای خطرات نسبی وجود ندارد اما الگوریتم‌های قابل اعتمادی مانند VEM برای این هدف پیشنهاد شده است (۲۳). راه دیگر استفاده از امید پسین است.

$$\hat{\theta}_i = \frac{\sum_{j=1}^{k}\hat{\theta}_i\hat{p}_j Pois(O_i;E_i,\hat{\theta}_i)}{\sum_{j=1}^{k}\hat{p}_j Pois(O_i;E_i,\hat{\theta}_i)}$$

در مدل آمیخته‌ی پیشنهاد شده توسط لاوسون و کلارک (۲۰۰۲) فرض می‌شود که لگاریتم خطر نسبی می‌تواند به سه مؤلفه جمع‌پذیر تفکیک شود. یک مؤلفه ثابت نشان‌دهنده‌ی ناهمگنی ناهمبسته (v) که بیش پراکندگی در یک ناحیه خاص را اندازه‌گیری می‌کند و دو مؤلفه آمیخته (u,φ) که معرف جنبه‌های متفاوت همبستگی فضایی هستند. این دو مؤلفه ترکیبی از یک مؤلفه همبستگی فضایی و یک مؤلفه که گام‌های گسسته را مدل‌بندی می‌کند می‌باشند (۲۳). مدل آمیخته عبارت است از:

$$\log\theta_i = \alpha + v_i + p_i u_i + (1-p_i)\varphi_i$$

که دارای حالات خاصی برحسب مقادیر p_i است. اگر p_i=۱ این مدل مشابه مدل بیز کامل (BYM) است که در قسمت بعد به ان اشاره می‌شود و اگر p_i=۰ این مدل تبدیل به مدل گام خالص می‌شود.

۶- مدل بیز کامل (BYM)

در این مدل فرض بر این است که تعداد مر گ ها در هر منطقه دوبه دو مستقل بوده و از توزیع پوآسن پیروی می کنند. این مدل توسط Kaldor و توسط Besag، York و Mollie بسط داده شد. (۱۰ و ۲۴). فرمول بندی این مدل به شرح زیر است:

$$\log(\theta_i) = \alpha + u_i + v$$

در این مدل پارامتر خطر نسبی بیماری به سه مؤلفه تجزیه می‌شود α مؤلفه روند که سطح کلی خطر نسبی است. u_i نیز بیش‌پراکندگی فضایی یا ناهمگنی همبسته‌ی فضایی است برای درنظر گرفتن امکان همبستگی بین خطرات نسبی در نواحی نزدیک هم، در مدل وارد می‌شود. برای این مؤلفه، ساختار همبستگی فضایی استفاده می‌شود که در آن برآورد خطر نسبی در هر ناحیه وابسته به نواحی همسایه است. مدل

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

اتورگرسیو شرطی (Conditional Autoregressive) پیشنهاد شده توسط Besag و همکاران عبارت است از:

$$\left[u_i \middle| u_j, i \neq j, \tau_u^2\right] \sim N(\bar{u}_i, \tau_i^2)$$

$$\bar{u}_i = \frac{1}{\sum_j w_{ij}} \sum_j u_j w_{ij} \quad \text{و} \quad \tau_i^2 = \frac{\tau_u^2}{\sum_j w_{ij}}$$

که $w_{ij} = 1$ (اگر i و j مجاور باشند، $w_{ij} = 0$ و در غیر این صورت) با فرمول‌بندی مدل برای بیش‌پراکندگی همبسته فضایی، واریانس وابسته به تعداد همسایگی‌ها است و استقلال به خوبی تعریف نمی‌شود. برای حل این مشکل، عبارت دیگری (v_i) به مدل وارد می‌شود که پارامتر بیش‌پراکندگی غیر فضایی یا ناهمگنی ناهمبسته‌ی فضایی است (۲۵).

۷- مدل‌های چندسطحی

روش دیگر نقشه‌بندی بیماری‌ها فرم لگ خطی را به کار می‌گیرد. $\theta_i = \exp\{\beta_o + v_i\}$

که در آن عبارت تصادفی $v_i \sim N(0, \sigma_v^2)$ است و σ_v^2 واریانس اثرات تصادفی v است. این مدل را می‌توان برحسب تعداد موارد نیز نوشت (به جای میزان‌ها):

$$y_i \sim Poisson(\mu_i) \quad \text{و} \quad \log(\mu_i) = \log(E_i) + \beta_0 + v_i$$

در حالت کلی مدل‌های چندسطحی به داده‌هایی برازش داده می‌شوند که در ساختار خود دارای سطوح خوشه‌بندی هستند. در نقشه‌بندی بیماری‌ها و کاربردهای جغرافیایی، این سطوح، می‌توانند مناطق جغرافیایی مانند بخش‌های شهرستان‌ها، شهرستان‌های استان‌ها یا استان‌های کشور تعریف شوند که برای هر سطح جغرافیایی می‌توان اثرات تصادفی دارای توزیع نرمال را در نظر گرفت (۲۶). به عنوان مثال برای شهرستان‌های استان‌ها، می‌توان مدل

$$y_{ij} \sim Poisson(\mu_{ij}) \quad \text{و} \quad \log(\mu_{ij}) = \log(E_{ij}) + \beta_0 + v_{ij} + u_{ij}$$

را تعیین کرد که در آن اثرات تصادفی شهرستان‌ها و استان‌ها دارای توزیع نرمال هستند.

$$v_j \sim N(0, \sigma_v^2) \quad \text{و} \quad u_{ij} \sim N(0, \sigma_u^2).$$

مدل‌های چندسطحی با پاسخ پوآسن را می‌توان هم با رویکردهای فراوانی‌گرایانه و هم با رویکردهای بیزی برازش داد. رویکردهای فراوانی‌گرایانه معمولا نیازمند استفاده از تقریبات هستند. به عنوان مثال نرم‌افزار MLwiN از روش‌های شبه‌درست‌نمایی استفاده می‌کند که نیازمند استفاده از تقریب‌های سری‌های تیلور برای تبدیل مسأله به حالتی است که بتوان با استفاده از الگوریتم حداقل مربعات مکرر عمومی برازش را انجام داد (۳۰-۲۷). دیگر رویکردهای رایج فراوانی گرایانه شامل تقریبات لاپلاس و مربع نرمال است. برآورد بیزی شامل بسط مدل‌ها به مدل‌هایی است که شامل اطلاعات پیشین بوده و از برآوردهای MCMC برای برازش استفاده می‌کنند (۳۲-۳۱).

بحث

با وجود مطالب و مقالات فراوان درخصوص طراحی مدل‌های نقشه‌بندی بیماری‌ها، توجه اندکی بر ارزیابی و مقایسه مدل‌ها به ویژه با استفاده از معیارهای نکویی برازش مدل‌ها معطوف شده است. Lawson و همکاران در مطالعه‌ای با استفاده از شبیه‌سازی داده‌ها از برخی از مدل‌های ممکن، به ارزیابی چگونگی پوشش پراکندگی فضایی توسط هر کدام از مدل‌ها و بررسی شایستگی نسبی هر کدام از مدل‌های برازش یافته با استفاده از چندین معیار نکویی برازش پرداختند (۲). Richardson یک مطالعه شبیه‌سازی جامع را برای روشن نمودن میزان هموارسازی خطرات در مدل‌های بیزی نقشه‌بندی بیماری‌ها به منظور ارزیابی عملکرد آن‌ها در تشخیص نواحی با خطر بالای واقعی انجام داد (۳۳). Marshall روشهای بیز تجربی و برخی روش‌های ابتدایی بیز کامل در نقشه‌بندی بیماری‌ها را مرور و مورد مقایسه قرار داد (۳۴). Bithell درخصوص رویکردهای ناپارامتری و مبتنی بر مدل بحث نمود (۳۵). Lawson و Richardson استفاده‌ی مدل‌های فضایی در کاربردهای اپیدمیولوژیک را مورد بحث قرار دادند (۱۴ و ۳۶). Manda و Leyland روش‌های فراوانی‌گرایانه (حداکثر درست‌نمایی) را با مدل‌های فضایی بیزی مقایسه کرده‌اند (۳۷). Bernardinelli و

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

Montomoli مقایسه‌ای بین مدل‌های بیز کامل و بیز تجربی و مدل‌هایی که با پراکندگی بیماری‌ها سروکار دارند انجام داده‌اند (۳۸). Best جدیدترین مدل‌های بیزی که در نقشه‌بندی بیماری‌ها از مدل‌های بیز کامل استفاده می‌کنند را خلاصه کرده است (۳۹). Clayton و Kaldor مقایساتی درباره مدل‌های بیز تجربی و بیز کامل انجام داده‌اند (۱۰). تلاش‌های دیگری برای ارزیابی حساسیت مدل بیز کامل نیز انجام شده است که این ارزیابی‌ها نیز تنها محدود به دامنه محدودی از مدل‌های بیزی است (۴۰، ۲۶). بر اساس این مطالعات نتایج زیر حاصل شده است.

الف)مقایسه مدل‌های بیزی و غیربیزی: با وجود تشابه برآورد پارامترها در دو مدل، برآورد واریانس اثرات تصادفی در رویکرد فراوانی‌گرایانه عموما کمتر از روش بیزی است. هم‌چنین اجرای روش بیزی آسان‌تر است و برآوردهای قابل اطمینان‌تری ارایه می‌دهند. میزان‌های بروز بیماری‌ها که با استفاده از روش بیزی به دست می‌آیند دقیق‌تر (دارای فاصله اطمینان کوتاه‌تر) هستند (۴۲-۴۱). ثابت شده است که روش‌های بیزی مبادل‌های بین اریبی و کاهش واریانس برآوردها ایجاد می‌کنند. هم‌چنین ثابت شده است که به ویژه هنگامی که حجم نمونه کم است این روش‌ها ویژگی‌های خوبی برحسب کمینه‌کردن زیان مربع خطا دارند (۱۴).

ب)مدل‌های هموارسازی: هموارسازی با حذف نوسانات، امکان مشاهده الگوهای ذاتی و واقعی فضایی را فراهم می کند و برآوردهای دقیق‌تر و فواصل اطمینان کوتاه‌تری نسبت به حالت عدم هموارسازی را نتیجه می‌دهد. در حالت کلی، اثر هموارسازی کاهش تعداد واحدها در بالاترین و پایین‌ترین گروه‌ها و بدین ترتیب افزایش تعداد واحدها در گروه‌های میانی است. در حالت کلی این مدل‌ها عملکرد مطلوبی نداشته و استفاده از هموارساز کرنل توصیه نمی‌شود اما این روش در برآورد لگاریتم خطر نسبی عملکرد بهتری دارد (۲). در صورت وجود خود همبستگی هموارساز کرنل عملکردی ضعیف دارد.

ج)مدل‌های بیزخطی: در بین روش‌های بیز خطی، مدل کلی Marshall نتایج بهتری را ارایه می‌دهد گرچه خودهمبستگی در باقیمانده‌ها را به جا می‌گذارد. هم‌چنین این مدل‌ها رفتاری مشابه مدل گاما - پوآسن از خود نشان می‌دهند (۲).

د)مقایسه مدل‌های بیز تجربی و بیز کامل: مدل بیز تجربی از این لحاظ مورد انتقاد است که برآوردهای تولید شده بسیار زیاد به سمت میانگین کلی میل می‌کنند. به همین دلیل برآوردهای بیز تجربی مقید (constrained) معرفی گردیده‌اند که برآوردهایی تولید می‌کنند که پراکندگی توأمشان مشابه داده‌های خام است (۴۳). مدل بیز کامل برتر از مدل بیز تجربی است زیرا عدم قطعیت در پارامترهای مدل را درنظر می‌گیرد در حالی‌که برآوردهای مدل بیز تجربی، برآوردها را نسبت به برآورهای نقطه‌ای پارامترهای مدل شرطی می‌کند. بنابراین ممکن است برآوردهای مدل بیز تجربی صحت کمتری داشته باشند اما از سوی دیگر روش بیز کامل نیازمند برنامه‌های رایانه‌ای پیچیده‌تر بوده و اجرای آن مشکل‌تر است (۳۹). با این وجود مدل بیز تجربی برای ارزیابی اولیه و مقدماتی خطر بیماری در سیستم‌های اطلاعات جغرافیایی (GIS) که تعداد زیادی داده مدیریت می‌شوند و الگوریتم مورد استفاده باید سریع و آسان باشد، مفید است. در برآورد خطر نسبی نیز با وجود تفاوت برآورد خطر نسبی تحت دو روش، رتبه خطر نسبی در کل ناحیه نقشه‌بندی مشابه است (۱۰). بررسی مجلات مهم اپیدمیولوژیک در دهه ۱۹۹۰- ۲۰۰۰ نشان داد که مدل BYM تنها مدل فضایی بیز کامل مورد استفاده در کاربردهای منتشرشده‌ی نقشه‌بندی بیماری‌ها بوده است. مقایسات انجام شده در مطالعه اخیر و مطالعات دیگر نشان می‌دهد که این مدل هنوز هم یک مدل مناسب در نقشه‌بندی بیماری‌های نواحی کوچک است اما مدل‌های فضایی بیزی دیگر که در سالیان اخیر پیشنهاد شده‌اند، در برخی کاربردها توانایی بهبود برآوردها را دارند (۳۹).

ه)مدل‌های آمیخته: مدل BYM و مدل آمیخته دارای ویژگی‌های خوب و مطلوبی در مدل‌بندی یک بیماری هستند. ویژگی اختصاصی مدل آمیخته، توانایی اقتباس اطلاعات هم از نواحی مجاور و هم از نواحی غیر مجاور است. این ویژگی

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

منجر به اریبی کمتر برآوردهای نقطه‌ای خطرات، نسبت به مدل‌های دیگر برای خطرات بالا می‌گردد. مدل BYM گرایش به بیش‌هموارسازی برآوردهای نقطه‌ای دارد اما در طبقه‌بندی کلی نواحی، اندکی بهتر از مدل آمیخته است. مدل‌های آمیخته حتی هنگامی که خطرات نسبی واقعی آمیخته باشند، خطرات نسبی را به خوبی پوشش نمی‌دهند (۳۹). این مدل‌ها در تحلیل مقادیر به خوبی عمل می‌کنند اما این تحلیل‌ها معمولا با مدل‌های BYM و گاما بهتر انجام می‌شوند (۲). چندین مدل برای تشریح توزیع فضایی میزان‌ها معرفی شده‌اند. این مدل‌ها، مدل‌های پوآسن امیخته (Poisson Mixed Models) بوده که اثرات فضایی اختصاصی نواحی را وارد مدل می‌کنند. این اثرات می‌توانند معرف عوامل محیطی باشند و برخی اوقات همبستگی فضایی را نیز نشان دهند (۴۴، ۳۸، ۲۴، ۱۰).

نتیجه‌گیری

به طور کلی می‌توان گفت مدل‌های متعددی برای نقشه‌بندی بیماری‌ها مطرح شده است که هرکدام دارای مبانی تئوری، مزایا و معایب و نقاط قوت و ضعف خاص خود بوده و در مواردی خاص نسیت به دیگر مدل‌ها ارجح هستند. در این مقاله مجموعه‌ای از مهم‌ترین مدل‌های نقشه‌بندی بیماری‌ها تشریح و مورد مقایسه قرار گرفته و تفاوت‌ها و شباهت‌های آن‌ها مورد بررسی قرار گرفت اما از آنجا که در سالیان اخیر تلاش‌های فراوانی در جهت رفع نقاط ضعف مدل‌های موجود و گسترش و بسط بیش از پیش و روزافزون این مدل‌ها به عمل آمده است، لزوم مقایسه‌ی دامنه‌ی گسترده‌تری از مدل‌ها، هم از منظر تئوری و هم از منظر نتایج در یک مطالعه شبیه‌سازی جامع ضروری به نظر می‌رسد. ضمن آن‌که در این مقاله تنها در خصوص نقشه‌بندی یک بیماری بحث گردیده و با توجه به اهمیت و کاربرد فراوان مدل‌های نقشه‌بندی توأم و چندمتغیرهی بیماری‌ها، مطالعه و مقایسه‌ی این مدل‌ها نیز می‌تواند مفید واقع شود.

References

1. Lawson AB, Browne WJ, Rodeiro CLV. Disease mapping with WinBUGS and MLwiN. Washington :Wiley Online Library; 2003.
2. Lawson A, Biggeri A, Boehning D, Lesaffre E, Viel J, Clark A, et al. Disease mapping models: an empirical evaluation. Disease Mapping Collaborative Group. Statistics in Medicine2000;19(17-18):2217-41.
3. Mahaki B, Mehrabi Y, Kavousi A, Mohammadian Y, Khodakarim S. Applying Gamma-Poisson, LogNormal and BYM models in comparing relative risk of suicide among counties of Ilam, Journal of Health System Research 2012; 9(1): 1-10. [In Persian]
4. Pickle LW, Mungiole M, Jones GK, White AA. Exploring spatial patterns of mortality: the new atlas of United States mortality. Statistics in Medicine 1999;18(23):3211-20.
5. Cancer E-Atlas [http://www.ncin.org.uk/eatlas/]. Available at Feb 2013.
6. Congdon P. Applied Bayesian Modeling. Clinchester: John Wiley&Sons; 2003
7. Lawson AB. Statistical Methods in Spatial Epidemiology. Chichester: John Wiley & Sons. 2001.
8. Elliott P, Wakefield JC, Best NG, Briggs DJ. Spatial Epidemiology: Methods and Applications. London: Oxford University Press; 2000.
9. Asmarian NS, Kavousi A, Salehi M, Mahaki B. Mapping of Stomach Cancer Rate in Iran Using Area-to-Area Poisson Kriging. Journal of Health System Research 2012; 9(3): 681-7. [In Persian]
10. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping Biometrics 1987;43(3):671-81.
11. Marshall RJ. Mapping disease and mortality rates using empirical Bayes estimators. J R Stat Soc Ser C Appl Stat 1991;40(2):283-94.

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

12. Mahaki B, Mehrabi Y, Kavousi A, Akbari ME, Waldhoer T, Schmid VJ, et al. Multivariate Disease Mapping of Seven Prevalent Cancers in Iran using a Shared Component Model. Asian Pac J Cancer Prev 2011;12(9):2353-8.
13. Asmarian NS, Kavousi A, Salehi M, Mahaki B. Comparison of Poisson Kriging and Empirical Bayesian Methods in mapping of gastrointestinal cancer incidence rate in Iran. Journal of Health System Research 2013; 9(3): 277-85. [In Persian]
14. Lawson AB, Biggeri AB, Boehning D, Lesaffre E, Viel J-F, Bertollini R. Disease Mapping and Risk Assessment for Public Health. New York: Wiley;1999.
15. Asmarian NS, Ruzitalab A, Amir K, Masoud S, Mahaki B. Area-to-Area Poisson Kriging Analysis of Mapping of County-Level Esophageal Cancer Incidence Rates in Iran. Asian Pac J Cancer Prev 2013;14(1):11-3.
16. Simonoff JS. Smoothing methods in statistics. New York: Springer Verlag; 1996.
17. Kelsall JE, Diggle PJ. Non-parametric estimation of spatial variation in relative risk. Stat Med 1995; 14(21-22):2335-42.
18. Kelsall J, Diggle P. Spatial variation in risk of disease: a non-parametric binary regression approach. Journal of the Royal Statistical Society: Series C (Applied Statistics) 1998; 47(4): 559-73.
19. Lee PM. Bayesian Statistics: An Introduction. 2th End. New York: Wiley; 1997.
20. Wolpert RL, Ickstadt K. Poisson/gamma random field models for spatial statistics. Biometrika 1998;85(2):251-67.
21. Carlin BP, Louis TA. Bayes and empirical Bayes methods for data analysis. Statistics and Computing 1997;7(2):153-4.
22. Böhning D, Schlattmann P, Lindsay B. Computer-assisted analysis of mixtures (C.A.MAM): statistical algorithms. Biometrics 1992; 48(1):283-303.
23. Lawson AB, Clark A. Spatial mixture relative risk models applied to disease mapping. Stat Med 2002; 21(3):359-70.
24. Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 1991; 43(1):1–59.
25. Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M. Bayesian analysis of space-time variation in disease risk. Stat Med 1995;14(21-22):2433-43.
26. Goldstein H. Multilevel Statistical Models. 2th End. London: Edward Arnold; 1995.
27. Goldstein H. Multilevel mixed linear model analysis using iterative generalized least squares. Biometrika 1986; 73(1), 43–56.
28. Goldstein H. Nonlinear multilevel models with an application to binary response data. Biometrika 1991; 78(1), 45–51.
29. Goldstein H, Rasbash J. Improved approximations for multilevel models with binary response. Journal of the Royal Statistical Society1996; 159 (3), 505–13.
30. Rasbash J, Browne W, Goldstein W, Yang M, Plewis I, Healy M, et al. A User's Guide to MLwiN. London: Institute of Education; 2009.
31. Raudenbush S, Yang M, Yosef M. Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate Laplace approximation. Journal of Computational and Graphical Statistics 2000; 9(1): 141–57.
32. Rabe-Hesketh S, Toulopoulou T, Murray R. Multilevel modeling of cognitive function in schizophrenic patients and their first degree relatives. Multivariate Behavioral Research 2001; 36(2): 279–98.
33. Richardson S, Thomson A, Best NG, Elliott P. Interpreting posterior relative risk estimates in disease mapping studies. Environ Health Perspect 2004; 112(9):1016-25.

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

34. Marshall RJ. A review of methods for the statistical analysis of spatial patterns of disease. Journal of the Royal Statistical Society Series A (Statistics in Society). 1991; 154(3): 421–41.
35. Bithell JF. A classification of disease mapping methods. Stat Med 2000; 19(17-18):2203-15.
36. Richardson S. Spatial models in epidemiological applications. In Green PJ, Hjort NL, Richardson S, eds. highly structured stochastic systems. Oxford: Oxford University Press; 2003. P. 237–59.
37. Manda SOM, Leyland A. An empirical comparison of maximum likelihood and Bayesian estimation methods for multivariate spatial disease modeling. South African Statistical Journal 2007; 41(1):1-21.
38. Bernardinelli L, Montomoli C. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. Stat Med 1992; 11(8):983-1007.
39. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. Statistical Methods in Medical Research 2005; 14 (1), 35–59.
40. Best N, Waller L, Thomas A, Conlon E, Arnold R. Bayesian models for spatially correlated disease and exposure data. In:, Bernardo J M., Berger J. O., Dawid A. P., Smith A. F. M. editors. Bayesian Statistics 6. London; Oxford University Press: 1998.
41. Downing A, Forman D, Gilthorpe MS, Edwards KL, Manda SOM. Joint disease mapping using six cancers in the Yorkshire region of England. International Journal of Health Geographics 2008;7(1):41.
42. Assunção RM, Castro MSM. Multiple cancer sites incidence rates estimation using a multivariate Bayesian model. International Journal of Epidemiology 2004;33(3):508-516.
43. Devine OJ, Louis TA. A constrained empirical bayes estimator for incidence rates in areas with small populations. Stat Med 1994; 13(11):1119-33.
44. Inskip H. Standardization methods. In: Armitage P, Corton T. editors. Encyclopedia of Biostatistics. New York; Wiley: 1998; 4237–50.

[Downloaded from hsr.mui.ac.ir on 2026-11-24]

Mapping Spatial Variation of Disease using Classic and Bayesian Models

Yadollah Mehrabi [1], Behzad Mahaki [2], Amir Kavousi [3]

Review Article

Abstract

Disease mapping includes a set of statistical techniques that lead to provide clean maps based on estimation of the incidence, prevalence and mortality rates for the users to be able to estimate the distribution of disease reliably. The main aims of disease mapping are to: describe the spatial variation in disease incidence for the formulation of etiological hypotheses; identify areas of unusually high risk so that action may be taken; provide a clean map of disease risk in a region to allow better resource allocation and risk assessment; disease atlas construction. So far, several case studies have been conducted in which the applications of disease mapping models have been used.

However, except in a few small studies and in limited situations, disease mapping models have not been compared and their general terms have not been studied. According to the importance of disease mapping and its growing use, in the present study the most important models have been introduced and their theoretical foundations, applications, advantages and disadvantages have been mentioned briefly and their various aspects have been compared.

Citation: Mehrabi Y, Mahaki M, Kavousi A. **Mapping Spatial Variation of Disease using Classic and Bayesian Models.** J Health Syst Res 2014; 10(2):201-211

Received date: 11.07.2013 **Accept date: 19.08.2013**

1. Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2. Department of Biostatistics, School of Public Health, Isfahan University of Medical Sciences, Isfahan, Iran (Corresponding Author) Email:behzad.mahaki@gmail.com
3. Department of Sciences, School of Health, Safety and Environment, Shahid Beheshti University of Medical Sciences,Tehran, Iran

[Downloaded from hsr.mui.ac.ir on 2026-11-24]