

## تحلیل فضایی و شناسایی خوشه های پرخطر و کم خطر ابتلا به بیماری لوسمی حاد لنفوبلاستیک در کودکان ایرانی طی سالهای ۱۳۸۵ الی ۱۳۹۳

سجاد رحیمی پردنجانی<sup>۱</sup>، امیر کاوسی دولانقرآ<sup>۲</sup>، کوروش اعتماد<sup>۲</sup>، عباس شاهسونی<sup>۲</sup>، بابک میرباقری<sup>۳</sup>

<sup>۱</sup> گروه اپیدمیولوژی و آمار، دانشگاه علوم پزشکی سمنان

<sup>۲</sup> دانشکده بهداشت و ایمنی، دانشگاه علوم پزشکی شهید بهشتی

<sup>۳</sup> دانشکده علوم زمین، دانشگاه شهید بهشتی

**چکیده:** در بین انواع سرطان های کودکان شایع ترین نوع سرطان لوسمی حاد لنفوبلاستیک است که تاکنون علل اصلی بروز آن شناخته نشده است. مطالعه حاضر برای مشخص کردن پاتولوژی جغرافیایی، تعیین توزیع مکانی و میزان خودهمبستگی فضایی، شناسایی خوشه های پرخطر و کم خطر ابتلا به بیماری لوسمی حاد لنفوبلاستیک در ایران انجام شده است. واژه های کلیدی: لوسمی حاد لنفوبلاستیک، خود همبستگی فضایی، خوشه پرخطر و کم خطر. کد موضوع بندی ریاضی (۲۰۱۰): 62F15، 62M30، 62H11.

### ۱ مقدمه

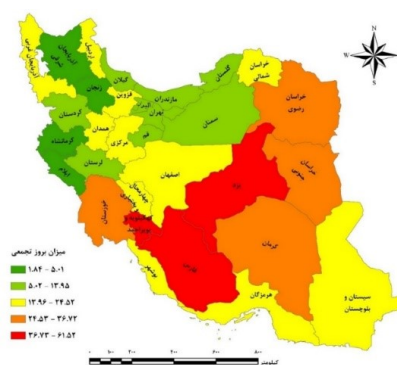
شایع ترین نوع سرطان در بین انواع سرطان های کودکان جهان و ایران لوسمی حاد لنفوبلاستیک<sup>۱</sup> (ALL) است. این بیماری حدود ۳۰٪ کل سرطان های کودکان را تشکیل می دهد (هانگر و همکاران، ۲۰۲۰). آگاهی از پهنه بندی، الگوی مکانی و شناخت مناطق در معرض خطر بالای بروز بیماری ها، یکی از مهم ترین جنبه های مطالعات اپیدمیولوژی و اتیولوژی سرطان محسوب شده و از پیش نیازهای اساسی در تصمیم گیری های مدیریت خدمات بهداشتی درمانی و سیستم مراقبت از بیماری ها است. خوشه در اپیدمیولوژی یعنی تجمع بروز موارد نادر بیماری یا هر واقعه مرتبط با سلامتی مثل سرطان ها، ناهنجاری های مادرزادی، خودکشی، سقط جنین و... که از لحاظ زمانی یا مکانی به هم نزدیک باشند و از نظر تعداد و تجمع بیش از مقداری باشد که عامل شانس باعث به وجود آمدن آن ها باشد. در خوشه ها موارد بیماری از لحاظ مکانیسم های

<sup>۱</sup> Acute Lymphoblastic Leukemia

بیولوژیکی، پاتولوژیکی، محیطی یا یک سری شرایط خاصی با یکدیگر مرتبط هستند. در اپیدمیولوژی فضایی و خصوصا مطالعات اکولوژیکال، مشاهدات از هم مستقل نبوده و برحسب مکان به یکدیگر وابسته و دارای خودهمبستگی فضایی هستند. خودهمبستگی فضایی بدین معناست که مقادیر مشاهداتی که از لحاظ مکانی به هم نزدیکتر هستند بسیار به هم شبیه‌ترند نسبت به مقادیر مشاهداتی که از لحاظ مکانی از هم دورترند بنابراین برای تشخیص خوشه‌ها بایستی از روش‌های آمار فضایی استفاده کرد. به منظور بررسی وضعیت اپیدمیولوژی و پاتولوژی جغرافیایی بیماری ALL در کودکان ایرانی و آگاهی از الگوی مکانی، برآورد خطر نسبی، شناسایی خوشه‌های پرخطر و کم خطر بروز بیماری در ایران، مطالعه حاضر انجام شد. روش‌های مختلف کشف خوشه‌ها با هم مقایسه شد تا بتوان قضاوتی صحیح‌تر از ریسک فاکتورهای دخیل در چرخه بروز ALL داشت.

## ۲ تحلیل فضایی

این مطالعه روی تمامی کودکان زیر ۱۵ سال مبتلا به لوسمی حاد لنفوبلاستیک در ایران طی سال‌های ۱۳۸۵ الی ۱۳۹۳ انجام شده است. سه نوع منبع داده در این مطالعه استفاده شد که شامل: منبع داده‌های مربوط به مشخصات بیماران که از مرکز ملی ثبت سرطان ایران طی سال‌های ۱۳۸۵ الی ۱۳۹۳ اخذ شدند. منبع داده‌های مربوط به جمعیت کودکان زیر ۱۵ سال استان‌های ایران که از مرکز آمار ایران اخذ شدند منبع داده‌های مربوط به مختصات جغرافیایی بیماران که از جدیدترین اطلاعات موجود در نرم‌افزار Google Map و لایه زمین مرجع شده استان‌های ایران بدست آمد. متوسط بروز سالانه لوسمی حاد لنفوبلاستیک طی سال‌های مورد مطالعه ۲/۲۵ در هر صد هزار کودک زیر ۱۵ سال بدست آمد. تعداد ۱۵۸۷ نفر از بیماران را دختران، ۴۲/۱۱٪، و ۲۱۸۲ نفر را پسران، ۵۷/۸۹٪، تشکیل دادند بنابراین نسبت جنسی بروز بیماری پسران به دختران ۱/۳۷ بدست آمد. میانگین و انحراف معیار سنی بیماران به ترتیب  $۳/۶۸ \pm ۵/۹۰$  و میانه سنی ۵ سال بدست آمد. به صورت کلی میزان بروز تجمعی بیماری طی سال‌های ۱۳۸۵ الی ۱۳۹۳ در ایران ۲۱/۳۱۵ در هر صد هزار نفر برآورد شد. چهار استانی که بیشترین میزان بروز تجمعی را داشتند شامل استان‌های فارس ۶۱/۵۲۴، کهگیلویه و بویر احمد ۵۴/۸۱۷، یزد ۴۶/۵۳۹ و خوزستان ۳۶/۷۲۲ بودند از سویی دیگر چهار استان دارای کمترین میزان بروز تجمعی شامل استان‌های کرمانشاه ۱/۸۳۸، آذربایجان شرقی ۳/۹۶۹، زنجان ۴/۷۵۸ و ایلام ۵/۰۰۷ بودند. در مطالعه حاضر از آماره

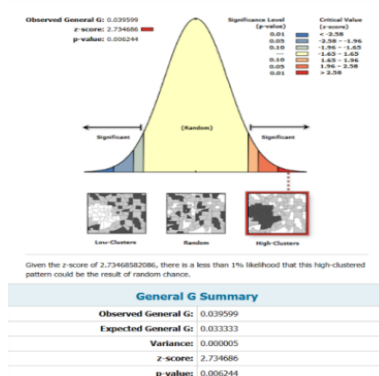


شکل ۱: میزان بروز تجمعی بیماری ALL در طی سال‌های ۱۳۸۵ الی ۱۳۹۳

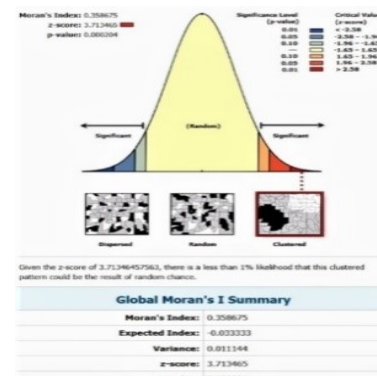
Global Moran's I و Getis-Ord General G برای ارزیابی خود همبستگی فضایی و از Anselin Local Moran's I، Optimized Hot Spot Analysis و Kuldorff's Spatial Scan Statistic نیز برای شناسایی دقیق مکان خوشه‌ها

استفاده شد.

**الف- شاخص موران کلی<sup>۲</sup>** به منظور ارزیابی میزان خود همبستگی فضایی بیماری از شاخص Global Moran's I استفاده شد. مقادیر این شاخص بین +۱ تا -۱ است. مقدار خود همبستگی مثبت به این معنی است که مقادیر در یک ناحیه شبیه به مقادیر نواحی مجاور است (داده ها دارای خود همبستگی فضایی و نیز خوشه ایی اند)، یعنی اگر مقادیر بروز در یک ناحیه بالا باشد در نواحی مجاور هم بالاست، مقدار خود همبستگی منفی به این معنی است که اگر یک ناحیه دارای مقادیر بروز بالاست نواحی مجاور مقدار کم بروز را دارند و برعکس. مقدار صفر این شاخص هم نشان از توزیع تصادفی مقادیر بروز دارد. به طور کلی این شاخص در یک مجموعه از داده‌ها خوشه ای بودن<sup>۳</sup>، تصادفی بودن<sup>۴</sup> و پراکنده بودن<sup>۵</sup> الگوهای موجود را ارزیابی می‌کند. نتیجه حاصل از آنالیز Global Moran's I در شکل ۲ الف نشان داده شده است. همان گونه که ملاحظه می‌شود مقدار این شاخص ۰/۳۵۸ بدست آمده که حاکی از شدت خود همبستگی و تمایل بسیار بالای بیماری که ALL به خوشه‌بندی دارد. مقادیر Z score and P-value حاصله نشان می‌دهند میزان خود همبستگی فضایی بیماری معنادار است. شاخص موران کلی میزان خود همبستگی بیماری را بصورت کلی نشان می‌دهد برای اما برای تشخیص و مقایسه میزان خود همبستگی مقادیر بالا و پایین بروز بیماری از شاخص گتیس آورد عمومی استفاده شد.



(ب)



(الف)

شکل ۲: میزان خود همبستگی فضایی بیماری ALL با استفاده از شاخص الف- Moran و ب- GOGG

**ب- شاخص گتیس آورد عمومی<sup>۶</sup>** این شاخص (GOGG)، توسط گتیس و ارد (۱۹۹۲) به عنوان یکی از آماره‌های کلی سنجش الگوهای فضایی ارائه شد. شاخص گتیس آورد عمومی، درجه و مقدار خوشه‌بندی را برای مقادیر بالا و پایین نشان می‌دهد. اگر مقدار این شاخص بالاتر از مقدار مورد انتظار باشد و Z-score هم مثبت باشد نشان‌دهنده تمایل مقادیر بالا به خوشه‌بندی دارد و در صورتیکه مقدار آن پایین‌تر از مقدار مورد انتظار و Z-score نیز منفی باشد حاکی از تمایل مقادیر پایین به ایجاد خوشه دارد. فرض صفر در GOGG عدم وجود خوشه‌بندی فضایی در منطقه است. نتایج این روش در شکل ۲ ب ارائه شده است. مقدار شاخص گتیس آورد عمومی برابر ۰/۰۳۹ شده است که حاکی از تمایل بیشتر مقادیر بالای بروز بیماری (hotspots) به خوشه‌بندی دارد (Z - score = ۲/۷۳ و P - value = ۰/۰۰۶).

<sup>3</sup>Cluster

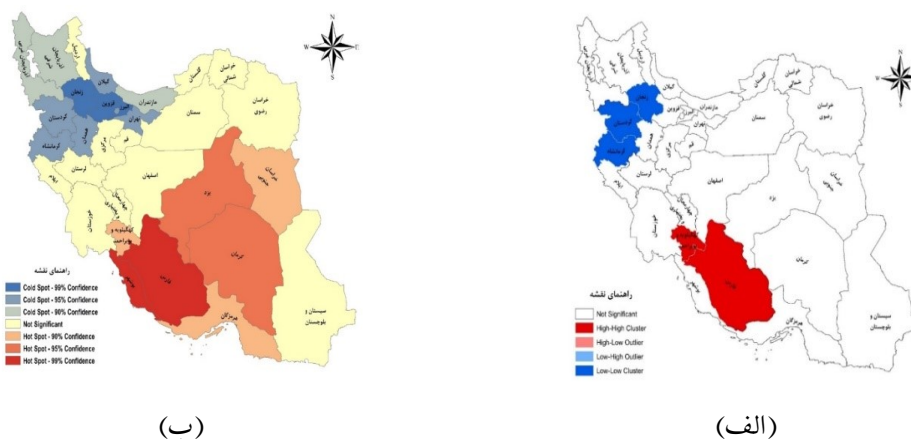
<sup>4</sup>Random

<sup>5</sup>Dispersion

<sup>6</sup>Getis-Ord General G

<sup>7</sup>Anselin local Moran's I

ج- شاخص موران محلی<sup>۷</sup> پس از اطلاع از بالابودن خود همبستگی فضایی و محرز شدن خوشه‌ای بودن توزیع فضایی بیماری تصمیم به شناسایی خوشه‌های پرخطر و کم خطر در ایران گرفته شد. توسط شاخص موران محلی، خوشه‌های پرخطر و کم‌خطر شناسایی گردید. شاخص موران محلی بلوک‌ها یا استان‌ها را در پنج دسته تقسیم می‌کند که عبارتند از: High-High(HH) یعنی بلوک‌هایی با مقادیر بالا (بلوک مرکزی) که توسط بلوک‌هایی با مقادیر بالا احاطه شده‌اند، Low-Low(LL) یعنی بلوک‌هایی با مقادیر پایین که توسط بلوک‌هایی با مقادیر پایین احاطه شده‌اند، High-Low (HL) و Low-High (LH) به ترتیب یعنی بلوک‌هایی با مقادیر کم احاطه شده توسط مقادیر بالا یا بلوک‌هایی با مقادیر بالا احاطه شده توسط بلوک‌هایی با مقادیر کم که Outlier یا نقاط پرت نیز نامیده می‌شوند و عدم معنی‌داری که نشان‌دهنده مناطقی است که توزیع بیماری دارای الگوی خاصی نیست و بصورت تصادفی پراکنده شده است. موران محلی بر روی پلیگون مرکزی متمرکز است و می‌تواند نقاط پرت (Outlier) را بخوبی شناسایی کند (هان و همکاران، ۲۰۱۹). بر اساس شکل ۳- الف، استان‌های فارس و کهگیلویه و بویر احمد به عنوان خوشه High-High شناسایی شدند که نشان‌دهنده خوشه‌ای بودن مقادیر بالا<sup>۸</sup> و تشکیل خوشه پرخطر در این مناطق است به عبارتی بروز بیماری در این خوشه بالاتر از همسایگان اطراف بوده است. در سویی دیگر استان‌های کرمانشاه، زنجان و کردستان بعنوان خوشه Low-Low شناسایی شدند که نشان‌دهنده خوشه‌ای بودن مقادیر پایین<sup>۹</sup> و تشکیل خوشه کم‌خطر در این مناطق است به عبارتی بروز بیماری در خوشه پایین‌تر از همسایگان اطراف بوده است. هیچگونه خوشه High-Low و Low-High مشاهده نگردید. اما شاخص موران محلی خوشه‌ها پرخطر و کم‌خطر را اولویت‌بندی نمی‌کند به همین خاطر از تحلیل نقاط داغ بهینه استفاده گردید.



شکل ۳: شناسایی خوشه های پرخطر و کم خطر ابتلا به ALL با استفاده از شاخص الف- Moran و ب- GOGG

یک روش دقیق‌تر، معتبرتر و جدیدتر برای تشخیص خوشه‌ها آنالیز نقاط داغ بهینه<sup>۱۰</sup>، (OHSA)، است که در واقع استفاده از آماره Getis-Ord Gi\* در حالتی است که در آن مقدار فاصله باند جهت تشخیص نقاط سرد و گرم بهینه شده باشد. OHSA به صورت خودکار مقدار فاصله باند<sup>۱۱</sup> بهینه را که در آن شدت خود همبستگی و خوشه‌بندی حداکثر است با روش Incremental Spatial Autocorrelation انتخاب کرده و بعنوان مقیاس تحلیل برای آماره Getis-Ord Gi\* مورد استفاده قرار می‌دهد. برای این کار از بالاترین مقدار Z-score استفاده می‌کند زیرا Z-score نسبت به شدت خودهمبستگی واکنش نشان داده و با افزایش مقدار خودهمبستگی افزایش و با کاهش آن نیز کاهش می‌یابد بنابراین بالاترین مقدار Z-score که

<sup>8</sup>Spatial Clusters of High Values

<sup>9</sup>Spatial Clusters of Low Values

<sup>10</sup>Optimized Hot Spot Analysis

<sup>11</sup>Distance Band or Distance Radius

نشان دهنده حداکثر مقدار خودهمبستگی است بعنوان معیاری برای انتخاب فاصله باند بهینه قرار می‌گیرد. بدین ترتیب بهترین نتایج ممکن در خروجی برای شناسایی نقاط سرد و نقاط داغ نشان داده خواهد شد چرا که خوشه‌های پرخطر و کم خطر با حداکثر مقدار خودهمبستگی فضایی و دقت بالایی تفکیک شده‌اند. همچنین مزیت دیگر OHSا این است که در نهایت پس از اجرا شدن نقاط سرد و گرم را در سطوح اطمینان ۹۰، ۹۵ و ۹۹ درصد نشان خواهد داد نتایج حاصل از OHSا در شکل ۳- ب نشان داده شده است. مقدار بهینه فاصله باند برآورد شده بر اساس این روش ۴۴۸ کیلومتر بود. این شاخص خوشه‌های پرخطر (Hot Spots) را در استان‌های فارس و بوشهر (در سطح اطمینان ۹۹٪)، یزد و کرمان (در سطح اطمینان ۹۵٪) و کهگیلویه و بویر احمد، خراسان جنوبی و هرمزگان (در سطح اطمینان ۹۰٪) شناسایی کرد. همچنین خوشه‌های کم‌خطر (Cold Spots) در استان‌های زنجان، البرز و قزوین (در سطح اطمینان ۹۹٪) و کرمانشاه، همدان، کردستان، تهران و گیلان (در سطح اطمینان ۹۵٪) و آذربایجان غربی، آذربایجان شرقی و مازندران (در سطح اطمینان ۹۰٪) مشاهده شدند. با نگاه کلی به پهنه‌بندی خوشه‌ها متوجه نکته قابل تأملی خواهد شد و آن قرار گرفتن تمامی Cold Spots در شمال و غرب کشور (طول جغرافیایی پایین‌تر و عرض جغرافیایی بالاتر)، قرار گرفتن تمامی Hot Spots در جنوب، مرکز و شرق کشور (طول جغرافیایی بالاتر و عرض جغرافیایی پایین‌تر) است. آنالیز نقاط داغ بهینه هیچ‌گونه اطلاعاتی از خطر نسبی خوشه‌های کشف شده نمی‌دهد به همین از آماره کاوشی فضایی کلدروف نیز استفاده گردید.

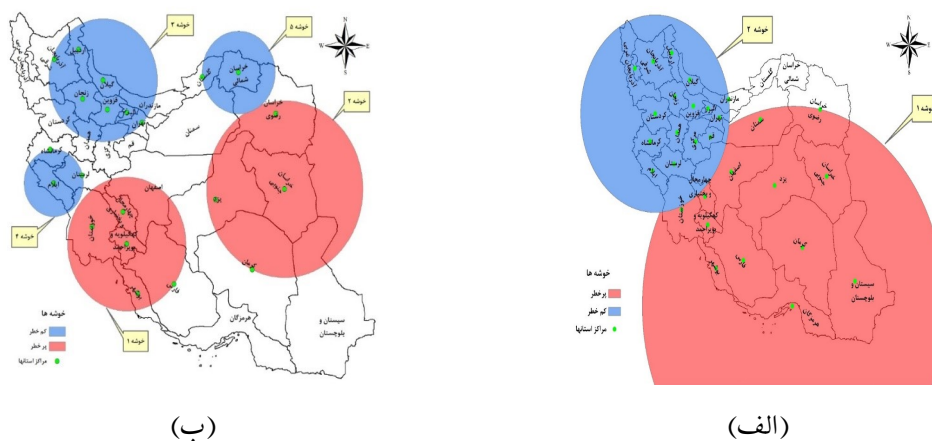
به منظور تعیین خطر نسبی در هر یک از استان‌ها تصمیم گرفته شد علاوه بر شاخص ALMI و OHSا، از روش نوین آماره کاوشی فضایی کلدروف<sup>۱۲</sup>، (KSSS)، نیز برای شناسایی خوشه‌های پرخطر و کم‌خطر استفاده گردد تا هم خطر نسبی خوشه‌ها برآورد شود و هم از صحت تشخیص خوشه‌ها اطمینان بیشتری حاصل شود. یکی از محاسن این آماره بدست آوردن میزان خطر نسبی بعنوان یکی از مهم‌ترین و پرکاربرترین شاخص‌های اپیدمیولوژیکی است که به صورت  $RR = \frac{c/E[c]}{(C-c)/(E[C]-E[c])}$  برآورد می‌شود که در آن  $c$  تعداد موارد مشاهده شده در داخل پنجره، خوشه یا منطقه جغرافیایی مورد نظر  $E(c)$  تعداد موارد مورد انتظار داخل پنجره، خوشه یا منطقه جغرافیایی مورد نظر است. بنابراین با تقسیم کردن این دو مولفه بر هم، مقدار خطر ابتلا به بیماری در داخل پنجره، خوشه یا منطقه جغرافیایی بدست می‌آید. از طرف دیگر  $C - c$  تعداد موارد مشاهده شده خارج از پنجره، خوشه یا منطقه جغرافیایی و  $E[C] - E[c]$  نیز تعداد موارد مورد انتظار خارج از خوشه را نشان می‌دهند، بنابراین برای برآورد کردن خطر در خارج از پنجره، خوشه یا منطقه جغرافیایی نیز بایستی از همین منطق پیش رفت به گونه‌ای که تعداد کل موارد مشاهده شده در خارج از پنجره، خوشه یا منطقه جغرافیایی را می‌بایست تقسیم بر تعداد موارد مورد انتظار در خارج از پنجره، خوشه یا منطقه جغرافیایی کرد تا مقدار خطر برای مناطق خارج از خوشه مورد نظر برآورد شود. موارد مورد انتظار طبق روش استانداردسازی غیر مستقیم طبق فرمول  $E[c] = p * \frac{C}{P}$  محاسبه می‌شود که در آن  $E[c]$  تعداد موارد مورد انتظار داخل پنجره، خوشه یا منطقه جغرافیایی مورد نظر،  $p$  تعداد افراد در معرض خطر بیماری در داخل پنجره، خوشه یا منطقه جغرافیایی،  $C$  تعداد کل موارد مشاهده شده در کل منطقه جغرافیای مورد مطالعه و  $P$  تعداد کل جمعیت در معرض خطر در تمام منطقه مورد مطالعه است. همچنین برای موارد مورد انتظار خارج از پنجره، خوشه یا منطقه جغرافیایی نیز مشابه آنچه در بالا اشاره شد انجام می‌شود (ما و همکاران، ۲۰۲۰).

### ۳ تعیین شعاع بهینه پنجره‌ها برای تشخیص خوشه‌ها

قبل از آنالیز خوشه‌ها نیاز به تعریف شعاع پنجره دایره‌ای اسکن کننده می‌باشد که با توجه به تعداد جمعیت در معرض خطر ناحیه تحت مطالعه تعیین می‌شود. بدیهی است اندازه خوشه‌ها هرگز از حداکثر اندازه تعریف شده برای شعاع پنجره نمی‌تواند بزرگتر باشد. در حالت پیش فرض روش Hierarchical شعاع دایره روی حداکثر مقدار ممکن خود یعنی ۵۰٪

<sup>12</sup>Kuldorff's Spatial Scan Statistic

جمعیت در معرض خطر در نظر گرفته شده است سپس به منظور شناسایی خوشه‌ها شعاع دایره به تدریج از صفر تا حداکثر مقدار تعیین شده افزایش می‌یابد. اما در حالت پیش فرض ممکن است خوشه‌های تکی بزرگ و وسیعی کشف شوند که از لحاظ اپیدمیولوژیکی خیلی کارایی نداشته باشند. برای حل این موضوع دو راه حل وجود دارد: اول اینکه در همان روش Hierarchical کاربر یک تحلیل حساسیت<sup>۱۳</sup> روی مقادیر درصد جمعیت در معرض خطر انجام دهد و با توجه به نتایج حاصله از شناسایی خوشه‌ها بهترین اندازه شعاع دایره را انتخاب کند. راه حل دوم و جدیدتر استفاده از روش Gini Optimized Cluster Collection است که با یک روش دقیق و به طور خودکار اندازه بهترین شعاع یعنی شعاع بهینه دایره اسکن کننده را برای شناسایی خوشه‌ها برآورد می‌کند. اخیرا نشان داده شده است ضریب جینی می‌تواند به جای شناسایی یک خوشه بزرگ، چندین خوشه ناهمپوشان را به صورت خالص تر، محلی تر، کوچکتر و واقعی تر شناسایی کند (کیم و جونگ، ۲۰۱۷). بنابراین جهت تعریف اندازه شعاع دایره و انجام تحلیل خوشه‌ها در مطالعه حاضر، ابتدا از روش پیش فرض Hierarchical با انتخاب کردن ۵۰٪ جمعیت در معرض خطر استفاده شد سپس در مرحله بعد به منظور شناسایی خوشه‌ها در مقیاسی محلی تر و تفکیک خوشه‌هایی که ممکن بود در درون خوشه‌های تشخیص داده شده با روش پیش فرض Hierarchical باشند از روش Gini Optimized Cluster Collection با مقدار بهینه برآورد شده یعنی ۳۱٪ جمعیت در معرض خطر استفاده شد. با در نظر گرفتن روش Hierarchical خوشه‌های شناسایی شده در ایران در شکل



شکل ۴: شناسایی خوشه‌های پرخطر و کم خطر ابتلا به ALL با استفاده از شاخص الف- Moran و ب- Gini

۴- الف نشان داده شده‌اند. همان‌گونه که ملاحظه می‌شود ۲ خوشه کشف شده است که خوشه کم خطر در شمال غرب و خوشه پرخطر از مرکز تا جنوب ایران شناسایی شده است. مقدار خطر نسبی در خوشه پرخطر برابر با ۲/۷ معنادار بدست آمد که بیان می‌کند خطر بروز لوسمی حاد لنفوبلاستیک در نواحی درون خوشه نسبت به نواحی بیرون از خوشه ۲/۷ برابر است (۱۷۰٪ بیشتر). همچنین مقدار خطر نسبی در خوشه کم خطر برابر با ۰/۳۸ و معنادار بدست آمد که بیان می‌کند خطر بروز لوسمی حاد لنفوبلاستیک در نواحی درون خوشه ۰/۳۸ برابر نواحی بیرون خوشه است (۶۲٪ کمتر). خوشه‌های شناسایی شده توسط روش جینی در شکل ۴- ب نشان داده شده است. همان‌گونه که مشخص است این روش خوشه‌ها را با دقت بالاتر و مقیاس کوچکتر و محلی تر شناسایی نموده است که از لحاظ اپیدمیولوژیکی بسیار کاربردی تر هستند. در مجموع ۵ خوشه مهم فضایی کشف شد که شامل ۲ خوشه پرخطر و ۳ خوشه کم خطر بود. مهم ترین خوشه پرخطر<sup>۱۴</sup> تشخیص داده در این روش با مقدار حداکثر لگاریتم نسبت درستمایی ۳۲۷ و مقدار خطر نسبی ۲/۵۶ در استان‌های فارس، کهگلویه و بویراحمد،

<sup>13</sup>Sensitivity Analysis

<sup>14</sup>Most Likely Cluster

خوزستان، چهارمحال و بختیاری و بوشهر شناسایی شد. بدین معنا که خطر بروز لوسمی حاد لنفوبلاستیک در داخل این خوشه بالاتر از خارج خوشه و ۲/۵۶ برابر آن است. در سویی دیگر مهم‌ترین خوشه کم‌خطر تشخیص داده در این روش با مقدار حداکثر لگاریتم نسبت درستمایی ۲۲۷ و مقدار خطر نسبی ۰/۴ در استان‌های اردبیل، آذربایجان شرقی، گیلان، البرز، تهران، زنجان و قزوین شناسایی شدند. خصوصیات کامل اپیدمیولوژیکی، آماری، جمعیتی و جغرافیایی خوشه‌های کشف شده توسط روش Gini Optimized در جدول ۱ ارائه شده است.

جدول ۱: خصوصیات خوشه‌های پرخطر و کم‌خطر با استفاده از روش Gini Optimized و  $P - value < ۰/۰۰۱$

شماره خوشه	خطر خوشه	نوع خوشه	تعداد استان	مختصات مرکز خوشه	شعاع خوشه	مشاهده شده	مورد انتظار	نسبت خطر نسبی	LLR
۱	بالا	اولیه	۵	(۳۰/۷۷, ۵۰/۸۳)	۲۹۴/۹	۱۲۷۶	۶۲۸/۹	۲/۰۳	۳۲۷/۴
۲	بالا	ثانویه	۴	(۳۲/۶۷, ۵۹/۲۲)	۳۸۶/۵	۸۵۶	۵۵۰/۱	۱/۵۶	۸۷/۶
۳	پایین	اولیه	۷	(۳۷/۲۵, ۴۹/۴۹)	۲۷/۰۳	۵۱۷	۱۰۶۵/۷	۰/۴۹	۲۲۷/۱
۴	پایین	ثانویه	۳	(۳۳/۱۲, ۴۶/۹۲)	۱۴۷/۵	۶۸	۲۱۷/۷	۰/۳۱	۷۳/۶
۵	پایین	ثانویه	۲	(۳۷/۴۰, ۵۷/۱۳)	۱۸۱/۳	۹۲	۱۴۵/۵	۰/۶۳	۱۱/۷

## بحث و نتیجه‌گیری

آنالیز داده‌ها به وسیله شاخص موران عمومی، شاخص موران محلی، شاخص گتیس عمومی، آنالیز نقاط داغ بهینه، آماره کاوشی فضایی کلدروف با استفاده از نرم افزارهای ArcGIS، SatScan، SPSS Stata، انجام شد. با نگاه کلی به پهنه‌بندی تمامی شاخص‌های تشخیص خوشه‌ها متوجه نکته قابل تأملی خواهیم شد و آن قرار گرفتن تمامی Cold Spots در شمال و غرب کشور (عرض جغرافیایی بالاتر و طول جغرافیایی پایین‌تر)، قرار گرفتن تمامی Hot Spots در جنوب و شرق کشور (عرض جغرافیایی پایین‌تر و طول جغرافیایی بالاتر) و توزیع تصادفی بیماری در مسیر کمربند جنوب غربی تا شمال شرقی ایران است. این طرز خوشه‌بندی و همبستگی حاصل بسیار خاص، قابل توجه و معنی‌دار بوده و ممکن است در اثر اینترکشن بین ریسک فاکتورهای محیطی و جغرافیایی همانند طول و عرض جغرافیایی محل سکونت، مواجهه شدید با اشعه فرابنفش نور خورشید و زندگی در مناطق نفت و گاز خیز جنوب ایران ایجاد شده باشد. یکی از دلایل بروز بالای بیماری در عرض جغرافیایی پایین می‌تواند تابش بیشتر و مستقیم‌تر نور خورشید باشد که متعاقباً منجر به مواجهه طولانی‌تر و شدیدتر با اشعه فرابنفش نور خورشید (UV) بعنوان یک کارسینوژن در بروز ALL باشد. در مطالعه کاسته و همکاران بسیار واضح شد با افزایش تابش نور خورشید در عرض جغرافیایی پایین، مواجهه با UV و متعاقباً بروز لوسمی نیز افزایش خواهند یافت (کاسته و همکاران، ۲۰۱۵). یکی دیگر از دلایل بروز بالای بیماری و پیدا شدن خوشه‌های پرخطر در مناطق جنوب ایران می‌تواند نفت‌خیز بودن، وجود منابع گازی فراوان و غلظت بالای بنزن و سایر هیدروکربن‌های نفتی در این مناطق باشد که توسط آژانس بین‌المللی تحقیقات روی سلامت بعنوان ریسک فاکتور قوی در بروز لوسمی گزارش شده‌اند. در بخش روش‌های آماری فضایی، آماره کاوشی فضایی کولدروف در تشخیص خوشه‌ها از انعطاف و کارایی بیشتری نسبت به سایر روش‌ها برخوردار است و نتایج حاصل از آن از دیدگاه اپیدمیولوژی منطقی‌تر و تفسیر پذیرتر است.

## تقدیر و تشکر

بدین وسیله از معاونت تحقیقات و فناوری دانشگاه علوم پزشکی شهید بهشتی، وزارت بهداشت درمان و آموزش پزشکی ایران، مرکز آمار ایران، قطب علمی تحلیل داده‌های همبسته مکانی و زمانی دانشگاه تربیت مدرس، مرکز مطالعات سنجش از دور و GIS دانشگاه شهید بهشتی که ما را در رسیدن به اهداف مطالعه یاری نمودند، تشکر می‌شود. این مقاله حاصل بخشی از رساله مقطع دکترای تخصصی (Ph.D) اپیدمیولوژی با کد اخلاق IR.SBMU.PHNS.REC.۱۳۹۸.۱۴۳ است.

## مراجع

- Hunger, S. P., Teachey, D. T., Grupp, S., & Aplenc, R. (2020), Childhood leukemia, *Abeloff's Clinical Oncology*, 1748-1764, Elsevier.
- Getis, A., & Ord, J. K. (1992), The Analysis of Spatial Association by Use of Distance Statistics, *Geographical Analysis*, **24**, 189-206.
- Han, X., Li, H., Liu, Q., Liu, F., & Arif, A. (2019), Analysis of Influential Factors on Air Quality from Global and Local Perspectives in China, *Environmental Pollution*, **248**, 965-979.
- Lu, P., Bai, S., Tofani, V., & Casagli, N. (2019), Landslides Detection through Optimized Hot Spot Analysis on Persistent Scatterers and Distributed Scatterers, *ISPRS Journal of Photogrammetry and Remote Sensing*, **156**, 147-159.
- Ma, R., Liang, L., Kong, Y., Zhai, S., Gu, J., Zhang, G., & Wang, T. (2020), Hotspot detection and socio-ecological factor analysis of asthma hospitalization rate in Guangxi, China, *Environmental Research* **183**, 109201.
- Kim, J., & Jung, I. (2017), Evaluation of the Gini Coefficient in Spatial Scan Statistics for Detecting Irregularly Shaped Clusters, *PloS One*, **12**, e0170736.
- Coste, A., Goujon, S., Boniol, M., Marquant, F., Faure, L., Doré, J. F., ... & Clavel, J. (2015), Residential Exposure to Solar Ultraviolet Radiation and Incidence of Childhood Hematological Malignancies in France, *Cancer Causes & Control*, **26**, 1339-1349.