

# پیش‌بینی مفاهیم اساسی یک حوزه بر اساس شاخص‌های قدمت و فراوانی استفاده از اصطلاحات موضوعی: مورد مطالعه سرطان کولون

مجتبی شمسی<sup>۱</sup>، محمد توکل‌زاده راوری<sup>۲</sup>، ابراهیم زال‌زاده<sup>۳</sup>، محمود باغبانیان<sup>۳</sup>

## مقاله پژوهشی

### چکیده

**مقدمه:** پیش‌بینی موضوعات آینده، یکی از ضرورت‌های سیاست‌گذاری در حوزه علم است که با تغییرهای گوناگونی صورت می‌گیرد. بر این اساس، پژوهش حاضر با هدف پیش‌بینی مفاهیم اساسی حوزه سرطان کولون، بر اساس شاخص‌های قدمت و فراوانی استفاده از اصطلاحات موضوعی انجام شد.

**روش بررسی:** این پژوهش از لحاظ نوع، توصیفی و از نظر هدف، تحقیقی کاربردی بود. جهت استخراج داده‌های مطالعه، دو جستجوی مجزا در دو دوره ۲۵ ساله طی سال‌های ۱۹۶۵-۱۹۸۹ و ۲۰۱۵-۱۹۹۰ در پایگاه اطلاعاتی PubMed صورت گرفت. جهت تعیین و پیش‌بینی موضوعات مهم حوزه سرطان کولون، از دو شاخص قدمت و متوسط فراوانی استفاده سالانه اصطلاحات MeSH به کار رفته در مدارک بهره گرفته شد. برای این کار، فراوانی و طول میزان استفاده از هر یک از این اصطلاحات در یک دوره ۲۵ ساله مورد توجه قرار گرفت و بر اساس کاهش و افزایش فراوانی این اصطلاحات در یک مقطع خاص، برجسته شدن یا زوال آن‌ها در مدارک آینده پیش‌بینی گردید.

**یافته‌ها:** بررسی میزان انطباق موضوعات اساسی نوپیدا در حوزه سبب‌شناسی سرطان کولون طی سال‌های ۲۰۱۵-۱۹۹۰ با موضوعات پیش‌بینی شده نشان داد که ۱۲۸ موضوع، به عنوان موضوع اساسی برای آینده پیش‌بینی گردید که از این تعداد، ۷۳ مورد (۵۸ درصد) در دوره دوم به وقوع پیوست. همچنین، بررسی میزان انطباق ۷۳ مورد با موضوعات اساسی دوره دوم حاکی از آن بود که موضوعات پیش‌بینی شده، ۲۲ درصد موضوعات اساسی را تشکیل داده است.

**نتیجه‌گیری:** نتایج پژوهش و بررسی مطالعاتی که به پیش‌بینی گرایش‌های آینده پرداخته‌اند، نشان می‌دهد که روشی جهت پیش‌بینی گرایش‌های موضوعی آینده با قطعیت و دقت بالا وجود ندارد. به نظر می‌رسد که روش‌های مختلف علم‌سنجی و نظر خبرگان باید در کنار هم مورد توجه قرار گیرد.

**واژه‌های کلیدی:** سرعنوان‌های موضوعی پزشکی؛ کتاب‌سنجی؛ سرطان کولون

پذیرش مقاله: ۱۳۹۵/۹/۷

اصلاح نهایی: ۱۳۹۵/۸/۱۶

دریافت مقاله: ۱۳۹۴/۱۱/۱۸

**ارجاع:** مختاری شمسی مجتبی، توکل‌زاده راوری محمد، زال‌زاده ابراهیم، باغبانیان محمود. پیش‌بینی مفاهیم اساسی یک حوزه بر اساس شاخص‌های قدمت و فراوانی استفاده از اصطلاحات موضوعی: مورد مطالعه سرطان کولون. مدیریت اطلاعات سلامت ۱۳: ۳۵۹-۳۵۴ (۵): ۳۵۴-۳۵۹

یک موضوع در این دو دوره مقایسه می‌گردد (۳). مطالعاتی نیز بر مبنای بسامد سالانه موضوعات انجام شده است (۴). روش دیگری که برای شناخت و پیش‌بینی گرایش‌های موضوعی مورد توجه قرار گرفته است، روش Exponential Smoothing می‌باشد. Krampen و همکاران داده‌های کتاب‌سنجی مواد منتشر شده در زمینه روان‌شناسی را با این روش از سال ۱۹۷۷ تا ۲۰۰۸ مدل‌سازی نمودند و برای ده سال بعد از سال ۲۰۰۸، پیش‌بینی کردند. این مدل پیش‌بینی بر فراوانی گذشته انتشارات استوار است (۵).

مقاله حاصل پایان‌نامه کارشناسی ارشد با شماره ۱۱۵۷۹۹۳ می‌باشد که با حمایت دانشگاه یزد انجام شده است.

۱- کارشناس ارشد، علم‌سنجی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم اجتماعی، دانشگاه یزد، یزد، ایران

۲- استادیار، علم اطلاعات و دانش‌شناسی، گروه علم اطلاعات و دانش‌شناسی، دانشکده علوم اجتماعی، دانشگاه یزد، یزد، ایران (نویسنده مسؤول)

Email: tavakoli@yazd.ac.ir

۳- استادیار، کبد و گوارش، گروه داخلی، دانشکده پزشکی، دانشگاه علوم پزشکی شهید صدوقی، یزد، ایران

### مقدمه

شناخت موضوعات آینده با اصطلاحاتی همچون Upward trends، Hot topic و Emerging trends بیان می‌شود (۱). در این راستا، روش‌های متفاوتی جهت شناخت گرایش‌های علمی و فنی در حوزه‌های سنجشی مبتنی بر اطلاعات کتاب‌شناختی مانند کتاب‌سنجی و علم‌سنجی برای شناخت موضوعات جاری و آینده صورت گرفته است که از قدیمی‌ترین آن می‌توان به اثر Price در سال ۱۹۶۵ اشاره نمود (۲). وی مفهوم جبهه تحقیق را مطرح کرد تا طبیعت تغییرپذیر و گذرای موضوعات تحقیقی را نشان دهد. آنچه را که Price مشاهده کرد و شاخص فوریت نامید، به تمایل دانشمندان به مورد استناد قرار دادن مقالات اخیر اشاره دارد. در یک حوزه علمی، یک جبهه تحقیق به مقالاتی اشاره دارد که دانشمندان به شدت مورد استناد قرار می‌دهند. بر اساس گفته او، یک جبهه تحقیق متشکل از ۴۰ تا ۵۰ درصد از مقالات می‌باشد (۲).

علاوه بر مفهوم جبهه تحقیق که بر استناد تکیه دارد، پژوهش‌هایی بر اساس داده‌های سری زمانی (مانند توزیع مدارک در یک دوره زمانی در یک موضوع خاص) روی درصد گرایش‌های موضوعی نیز صورت گرفته است. در این پژوهش‌ها، اغلب مدارک دو دوره زمانی در نظر گرفته می‌شود و درصد بسامد

پس از آن پیش‌بینی می‌شود. سپس نتایج پیش‌بینی‌ها با آنچه که چند سال بعد اتفاق افتاده است، مقایسه می‌گردد. بر این اساس، هدف ای انجام پژوهش حاضر، پیش‌بینی مفاهیم اساسی حوزه سرطان کولون بر اساس روش محاسبه شاخص‌های قدمت و فراوانی استفاده از اصطلاحات موضوعی بود.

### روش بررسی

این پژوهش از لحاظ نوع، توصیفی و از نظر هدف، یک تحقیق کاربردی بود. همچنین، از نظر ابزار جمع‌آوری داده‌های پژوهش، یک تحلیل اسنادی به شمار می‌رود که با رویکردهای علم‌سنجی صورت گرفت و مانند مطالعه Krampen و همکاران، مبتنی بر متغیرهای سری زمانی است (۵). جهت استخراج داده‌های پژوهش، دو جستجوی جداگانه از طریق PubMed انجام گرفت که در هر دو بار، اصطلاح کنترل جستجو شد. نتیجه این جستجوها به ترتیب ۳۹۴۶ و ۷۹۶۹ مدرک بود که هر کدام از این نتایج با استفاده از امکانات PubMed با فرمت Medline ذخیره گردید. مجموعه مدارک دو دوره یاد شده، به ترتیب واژگانی متشکل از ۲۵۷۴ و ۵۸۴۹ توصیفگر متمایز را دربرداشت. پس از آن، با تابعی در نرم‌افزارهای Word و Excel، زمان آغاز و پایان کاربرد توصیفگرها در مدارک مشخص گردید تا بتوان قدمت استفاده از توصیفگرها را محاسبه نمود. جهت محاسبه قدمت استفاده از توصیفگرها، به حاصل تقریق سال پایان و سال آغاز هر توصیفگر، یکی اضافه گردید [۱+ (آغاز-پایان)]. در مرحله بعد با استفاده از Ravar-Matrix 2، فراوانی استفاده از موضوعات مشخص شد تا بتوان از طریق تقسیم فراوانی هر توصیفگر بر قدمت استفاده از آن، به متوسط فراوانی استفاده سالانه از توصیفگرها دست یافت.

در پژوهش حاضر جهت تعیین و پیش‌بینی موضوعات مهم حوزه سرطان کولون، مانند روش قانون پراکندگی Bradford، ابتدا موضوعات بر اساس فراوانی و از کم به زیاد مرتب گردید؛ آن دسته از موضوعات ابتدایی که مجموع فراوانی آن‌ها یک سوم کل فراوانی‌ها بود، در دسته اول (دسته موضوعات هسته) قرار گرفتند. همچنین، موضوعاتی که از نظر مجموع فراوانی در محدوده یک سوم تا دو سوم کل فراوانی‌ها (یک سوم وسطی) قرار داشتند، به عنوان موضوعات نزدیک به هسته انتخاب گردید و یک سوم نهایی آن‌ها به عنوان موضوعات دور از هسته در نظر گرفته شد. در کنار آن، از دو شاخص قدمت استفاده و متوسط فراوانی استفاده سالانه نیز برای دسته‌بندی موضوعات در هر بازه زمانی استفاده گردید. بر اساس روش ذکر شده، سرعنوان‌های موضوعی به کار گرفته شده در مدارک مربوط به حوزه پژوهش، در صورتی می‌توانند به عنوان موضوع اساسی این حوزه محسوب شوند که هم‌زمان دو شاخص یاد شده را داشته باشند. به عبارت دیگر، سرعنوانی به عنوان موضوع اساسی قلمداد می‌شود که از لحاظ قدمت استفاده، جزء یک سوم سرعنوان‌هایی قرار گیرد که نسبت به بقیه قدمت بیشتری در مدارک حوزه مورد مطالعه دارند و از لحاظ بسامد استفاده نیز جزء بخش اول (یک سوم اول) سرعنوان‌های موضوعی پراستفاده قرار داشته باشد. همچنین، جهت تشخیص موضوعات نوپیدا (موضوعاتی که پیش‌بینی می‌شود در آینده برجسته شوند)، فرض بر این شد که آن موضوعاتی در آینده برجسته خواهند شد، علاوه بر شاخص‌های قدمت استفاده و متوسط فراوانی استفاده سالانه سرعنوان‌های MeSH، به شاخص فراوانی آن‌ها در دوره حضورشان نیز توجه گردد؛ بدین معنی که سرعنوان‌هایی که قدمت استفاده پایینی دارند، اما شاخص‌های متوسط فراوانی استفاده سالانه و

روش‌های آماری معمولی نیز در این راستا به کار گرفته شده است. Zhi و Ji تحقیقی را بر روی مقالات مرتبط با تالاب بین سال‌های ۱۹۹۱ تا ۲۰۱۱ انجام دادند. این مقالات از SCI-EXPAND بازیابی شدند و گرایش‌های جاری و آینده به روش‌های آماری مورد مطالعه قرار گرفت. با مطالعه واژگانی بر روی کلید واژه‌های پدیدآور، کلید واژه‌های عنوان و Keywords Plus، آن‌ها توانستند دریابند که آب، مواد غذایی، گیاهان و جریان از محورهای اصلی مطالعات بوده و بر اساس رشدی که موضوع خاص از سال ۲۰۰۵ داشته است، می‌تواند جزء موضوعاتی باشد که در حال ظهور در این حوزه است (۶).

استفاده از روش‌های دیداری سازی نیز از روش‌های دیگر برای پیش‌بینی است. Chen و همکاران به مطالعه گرایش‌های حوزه احیای پزشکی، بین سال ۲۰۰۰ تا ۲۰۱۱ بر اساس تحلیل علم‌سنجی پرداختند. آن‌ها یک شبکه از ۳۵۹۶۳ مقاله پژوهشی و مروری ایجاد کردند که به ۳۸۷۵ مقاله استناد کرده بودند. همچنین، از طریق تحلیل‌های دیداری‌سازی توانستند گرایش‌های موضوعی این حوزه را تشخیص دهند (۷). Glanzel نیز از فن خوشه‌بندی هیبریدی (مطالعه جفت‌های کتاب‌شناختی و پیوندهای استنادی) برای تعیین حوزه‌های در حال ظهور بهره برد که خود روش دیگری برای پیش‌بینی محسوب می‌شود (۸).

بر اساس آنچه گفته شد، روش‌های گوناگونی برای تعیین روند و گرایش‌های موضوعی و همچنین، یافتن موضوعات داغ وجود دارد که در آن‌ها، دو متغیر زمان و فراوانی (مانند فراوانی همبندی) نقش مهمی دارد. این دو متغیر در علم فیزیک تحت عنوان موج نیز مورد توجه قرار گرفته است که در آن‌ها مفاهیم فرکانس و زمان وجود پیدا می‌باشد. در بحث امواج، رفتار دوره‌ای یک موج بر اساس توابع خاصی قابل پیش‌بینی است که در آن متغیرهای زمان، فرکانس و قدرت موج (دامنه موج)، اساسی‌ترین متغیرها محسوب می‌شوند (۹). همان‌گونه که حوزه اطلاعات و ارتباطات، مفهوم گرما و انرژی گرمایی را با اطلاعات جایگزین کرده است و بر این اساس انتقال اطلاعات را تبیین می‌کند (۱۰)، پژوهش حاضر نیز به دنبال آن بود که رفتار یک موضوع را از لحاظ میزان استفاده شدن یا توجه به دو متغیر مشابه فراوانی و زمان، مورد توجه قرار دهد. بنابراین، مسأله اصلی این بود که آیا مانند فیزیک، بهره‌گیری از متغیرهای زمان (طول عمر استفاده از یک موضوع) و فراوانی (میزان استفاده از یک موضوع در مدارک به طور کلی و در زمان اخیر)، می‌تواند در پیش‌بینی وضعیت تمام اصطلاحات به کار رفته در یک حوزه، از جنبه اوج گرفتن یا پایین آمدن میزان استفاده، کمک کند. به عبارت دیگر، این پژوهش با هدف ارایه روشی جهت پیش‌بینی وضعیت استفاده از اصطلاحات موضوعی انجام شد که مبنای آن، بالا یا پایین رفتن میزان استفاده (فراوانی) از یک اصطلاح موضوعی در دوره اخیر (مثل دو ساله اخیر)، نسبت به کل فراوانی آن (مجموع فراوانی از زمان ظهور) در حوزه مورد نظر می‌باشد.

یکی از حوزه‌هایی که تغییرات موضوعی را به وضوح تجربه کرده است، سرطان کولون می‌باشد که اغلب به صورت ایجاد چند پولیپ (توده‌های خوش‌خیم بر روی سطح داخلی روده) آغاز می‌شود (۱۱). شیوع این بیماری بیشتر در سنین بالای ۵۰ سال است و با افزایش سن، ابتلا به بدخیمی آن بیشتر می‌شود (۱۲). بنابراین، روش پیشنهادی روی این حوزه از پزشکی سنجش می‌شود؛ به این صورت که فراوانی و قدمت ظهور اصطلاحات MeSH در این حوزه، از آغاز تا یک زمان خاص و همچنین، فراوانی به کارگیری از آن‌ها در چند سال منتهی به سال پایانی محاسبه و وضعیت استفاده از آن‌ها برای چند سال

در دوره زمانی سال‌های ۱۹۹۰ تا ۲۰۱۵، در سطح بالای ۶۷ درصد، آن موضوعاتی که حداقل ۲۳ سال قدمت داشتند و حداقل متوسط فراوانی استفاده سالانه آن‌ها ۰/۶۸ بار در سال بود، به عنوان موضوعات اساسی این دوره شناخته شدند که تعداد آن‌ها ۴۶۱ مورد برآورد گردید. برای رعایت اختصار، ۲۰ مورد از این موضوعات در جدول ۲ آمده است.

جهت آزمایش میزان توانایی تشخیص موضوعات برجسته در آینده، ابتدا داده‌های مرتبط با دوره اول مورد بررسی قرار گرفت تا مشخص گردد که این روش پیش‌بینی تا چه حد توانسته است در تشخیص موضوعات اساسی دوره بعد از خود موفق باشد. از میان موضوعات دوره اول، ۱۲۸ سرعنوان دارای متوسط فراوانی استفاده سالانه بیشتر از ۱ بار در سال و فراوانی حضور ۴ بار در دوره اول بودند، اما قدمت استفاده آن‌ها کمتر از ۱۵ سال بود. به عبارت دیگر، این تعداد سرعنوان شرایط یک موضوع نوپیدا را داشتند.

در ادامه، بر اساس شاخص‌های قدمت استفاده و متوسط فراوانی استفاده سالانه در سطح بالای ۰/۶۷، موضوعات اساسی دوره دوم تعیین شد تا بتوان به بررسی تحول در این حوزه پرداخت. بر این اساس، موضوعاتی که حداقل ۲۳ سال قدمت استفاده داشتند و متوسط فراوانی استفاده سالانه آن‌ها نیز حداقل ۰/۶۸ بود، به عنوان موضوعات اساسی این دوره در نظر گرفته شدند. از این طریق مشخص گردید که ۴۶۱ سرعنوان موضوعی MeSH از میان ۵۸۴۹ موضوع به کار رفته در مدارک مربوط به حوزه سبب‌شناسی سرطان کولون، شرایط حضور در جمع موضوعات اساسی دوره دوم این حوزه را دارند. از این تعداد موضوع اساسی، ۱۳۴ مورد همان موضوعات اساسی دوره اول بودند. به عبارت دیگر، ۶۷ درصد (۱۳۴ مورد) از موضوعات اساسی دوره قبل با موضوعات اساسی دوره جدید همپوشانی داشت و تنها ۳۳ درصد (۶۶ مورد) از این فهرست حذف شدند. همچنین، مشخص گردید که از ۱۲۸ مورد پیش‌بینی جهت تشخیص موضوعات اساسی دوره بعد (نوپیدا)، ۷۳ مورد درست بود. می‌توان گفت که روش پیشنهاد شده، توانسته است ۲۲ درصد (۷۳ مورد) از موضوعات اساسی پیدا شده در دوره جدید را شناسایی نماید.

برای پیش‌بینی موضوعات اساسی دوره جدید، میزان شباهت موضوعات اساسی دوره دوم با موضوعات نوپیدای دوره اول مورد بررسی قرار گرفت. همانگونه که بیان شد، تعداد موضوعات اساسی دوره دوم، ۴۶۲ موضوع بود که از این تعداد، ۱۳۴ موضوع به صورت مشترک در موضوعات اساسی دوره اول نیز حضور داشتند. بنابراین، ۳۲۸ موضوع از ۴۶۲ موضوع اساسی دوره دوم، به عنوان موضوع جدید مشخص شد که بعد از مقایسه این تعداد موضوع اساسی با ۱۲۸ موضوع نوپیدای دوره اول، تعیین گردید که تعداد ۷۳ موضوع به صورت مشترک در دو گروه حضور دارند. جهت تعیین میزان انطباق موضوعات اساسی نوپیدای حوزه پژوهش در دوره ۱۹۹۰ تا ۲۰۱۵ با موضوعات پیش‌بینی شده، از رابطه ۲ استفاده گردید.

فراوانی آن‌ها بالا می‌باشد، می‌تواند در آینده برجسته شود و به عنوان یک موضوع اساسی برای حوزه پژوهش حاضر محسوب گردد. وضعیت موضوعات باقی‌مانده به گونه‌ای است که در حال حاضر اهمیت زیادی ندارند و نقش آن‌ها در آینده غیر قابل پیش‌بینی است.

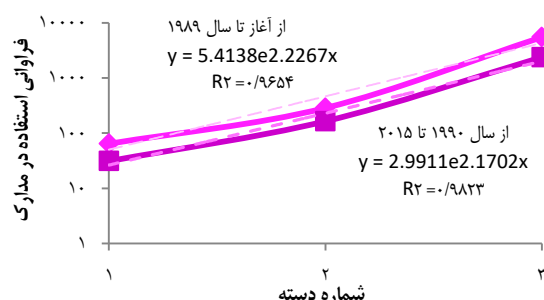
شاخص دربردارندگی، شاخصی است که به منظور بررسی کلی میزان شباهت هر حوزه نسبت به حوزه‌های دیگر به کار می‌رود. این شاخص شباهت یک حوزه را نسبت به حوزه دیگر به صورت نامتقارن و یک طرفه نشان می‌دهد (۱۳). رابطه ۱ نحوه محاسبه شاخص دربردارندگی را نشان می‌دهد.

$$\text{IncA,B} = 100 \cdot N/M \quad \text{رابطه ۱}$$

در این رابطه، N برابر با تعداد اصطلاحات مشترک بین دو مجموعه A و B می‌باشد و M نماینده تعداد اصطلاحاتی است که در مجموعه A وجود دارد.

### یافته‌ها

جهت تعیین موضوعات هسته، ابتدا از توزیع Bradford استفاده شد؛ بدین صورت که فراوانی سرعنوان‌های MeSH در مدارک مربوط به سبب‌شناسی سرطان کولون تا سال ۱۹۸۹ و همچنین، از سال ۱۹۹۰ تا ۲۰۱۴ محاسبه گردید. شکل ۱ نحوه برآزش توزیع سرعنوان‌های موضوعی را از نظر فراوانی کاربرد در مدارک مورد مطالعه نشان می‌دهد.



شکل ۱: توزیع استفاده از سرعنوان‌های موضوعی در دو دوره با توزیع Bradford (مقیاس محور عمودی لگاریتمی است)

جهت شناسایی موضوعات اساسی این دوره، آن دسته از موضوعاتی که دارای قدمت استفاده بیش از ۱۵ سال و نیز متوسط فراوانی سالانه بیش از ۱ بار در سال (در سطح بالای ۶۷ درصد) بودند، از بین موضوعات استخراج گردید که تعداد آن‌ها در دوره اول، ۲۰۰ موضوع بود. جهت رعایت اختصار، تعداد ۲۰ مورد از این موضوعات در جدول ۱ ارایه شده است.

جدول ۱: موضوعات اساسی سرطان کولون در دوره زمانی اول (سال‌های ۱۹۸۹-۱۹۶۵)

موضوعات اساسی			
Abo blood-group system	Chronic disease	Immunolectrophoresis	Neoplasms, multiple primary
Abscess	Colectomy	Immunoglobulin a	Neoplasms, radiation-induced
Acute disease	Colitis	Immunoglobulin g	Nitrosourea compounds
Adenocarcinoma	Colitis, ulcerative	Immunoglobulin m	Osteoma
Adenocarcinoma, mucinous	Colon, sigmoid	Immunotherapy	Ovarian neoplasms

جدول ۲: موضوعات اساسی سرطان کولون در دوره زمانی دوم (سال‌های ۲۰۱۵-۱۹۹۰)

موضوعات اساسی			
1,2-dimethylhydrazine	Combined modality therapy	Immunity, innate	Oncogenes
Abdominal neoplasms	Common bile duct neoplasms	Immunization	Organ specificity
Abdominal pain	Constipation	Immunoblotting	Osteoma
Abscess	Crohn disease	Immunoenzyme techniques	Ovarian neoplasms
Acute disease	Crosses, genetic	Immunoglobulin g	Palliative care

کولون را فراهم می‌آورند، می‌تواند پولیپ آدنوماتو، سندرم‌های ارثی، عادات و رژیم غذایی، مصرف تنباکو و مشروبات الکلی و سایر عوامل پزشکی مانند بیماری التهابی روده (Inflammatory bowel disease) IBD، دیابت و بیماری کرون باشد.

از لحاظ روشی، پژوهش حاضر شباهت زیادی با مطالعه Noyons و van Raan (۳) داشت. آنان نیز دو گروه از مدارک را از دو دوره زمانی در نظر گرفته بودند و نتیجه پیش‌بینی از دوره اول را با آنچه که در دوره دوم رخ داده است، مقایسه کردند تا میزان صحت پیش‌بینی روش مبتنی بر زمان و بسامد موضوعات را به صورت درصدی بسنجند (۳). تفاوت اساسی پژوهش حاضر برای پیش‌بینی موضوعات آینده این است که تنها به موضوعات کلی توجه ندارد، بلکه تمامی موضوعاتی را که در مدارک مربوط به سرطان کولون به کار رفته است، مورد توجه قرار می‌دهد. پیش‌بینی از روی داده‌های مربوط به یک دوره ۲۵ ساله، برای یک دوره ۲۵ ساله (بعد از آن) صورت گرفته است. اگرچه طول زمان دوره اول زیاد می‌باشد، اما در همین اندازه، پیش‌بینی برای یک دوره طولانی صورت گرفته است. با این همه، گرچه این روش نتوانسته بود همه موضوعات اساسی نوپیدای دوره بعد از خود را تشخیص دهد، اما ۲۲ درصد از موضوعات مهم ۲۵ سال آینده به درستی پیش‌بینی شد.

علاوه بر آن، کامل نبودن پیش‌بینی موضوعات، از محدودیت‌هایی نشأت می‌گیرد که برای پژوهش حاضر و مطالعات مشابه وجود دارد. مانند آنچه که در تحقیق Tseng و همکاران اشاره شد (۱). بهره‌گیری از خبرگان باید به عنوان مکمل تحلیل‌های کتاب‌سنجی و علم‌سنجی برای پیش‌بینی آینده مورد توجه قرار گیرد که در عمل به علت محدودیت زمانی در بررسی یک به یک موضوعات، این امر به سادگی ممکن نیست. برای دقت کار در سیاست‌گذاری‌های علمی، بهتر است که هم‌زمان از روش‌های متفاوتی که در مقدمه پژوهش ذکر گردید، استفاده شود تا نتایج جامعیت و دقت بیشتری داشته باشند.

### نتیجه‌گیری

نتایج پژوهش حاضر نشان داد که سیاست‌گذاران حوزه پزشکی می‌توانند با استفاده از روش ساده متغیرهای زمان و فراوانی، آینده ظهور و افول موضوعات را در حوزه‌های مختلف تا اندازه‌ای پیش‌بینی نمایند. در هر حال، روشی وجود ندارد تا بتوان گرایش‌های موضوعی آینده را با قطعیت و دقت بالا پیش‌بینی کرد. به نظر می‌رسد که روش‌های مختلف علم‌سنجی و نظر خبرگان باید در کنار هم مورد توجه قرار گیرد.

### پیشنهادها

برای بالا بردن دقت این پیش‌بینی، به نظر می‌رسد که اگر پوشش زمانی

رابطه ۲

$$۲۲ = ۱۰۰ \cdot \frac{۲۳}{۳۳۸} = \text{موضوعات اساسی نوپیدای دوره اول و دوره دوم}$$

رابطه ۲ نشان می‌دهد که ۲۲ درصد از موضوعاتی که در دوره ۲۵ ساله دوم به عنوان موضوعات اساسی حوزه سبب‌شناسی سرطان کولون ظهور کرده‌اند، از طریق توجه به شاخص‌های قدمت استفاده و متوسط فراوانی استفاده سالانه و بسامد در موضوعات دوره اول امکان‌پذیر بوده است. توجه به این شاخص‌ها، ۱۲۸ موضوع را به عنوان موضوع اساسی برای آینده پیش‌بینی کرد که از آن تعداد، ۷۳ مورد (۵۷ درصد) به وقوع پیوست.

### بحث

پیش‌بینی موضوعاتی که در آینده به عنوان موضوعات اساسی حوزه سبب‌شناسی سرطان کولون به وقوع خواهد پیوست، به عنوان بخشی از روند تحولات موضوعی این حوزه، مورد توجه پژوهش حاضر بود و رویکرد آن مانند مطالعه Noyons و van Raan، بهره‌گیری از داده‌های سری زمانی بود (۳). نتایج حاصل به صورت جامع نتوانست تمامی موضوعات مورد ظهور آینده را نشان دهد. همه موضوعات داغ و در حال ظهور را نمی‌توان با متغیر فراوانی مدرک و متغیرهای سری زمانی به دست آورد. به عبارت دیگر، نمی‌توان ادعا کرد که روشی وجود دارد تا بتوان با قطعیت و دقت بالا، گرایش‌های موضوعی آینده را پیش‌بینی نمود. این نکته در تحقیق Price (۲) نیز مشخص گردید. وی که به بررسی مفهوم جبهه‌های تحقیق می‌پردازد، نتیجه گرفت که یک جبهه تحقیق، از ۴۰ تا ۵۰ درصد مقالات تشکیل شده است. به عبارت دیگر، موضوعاتی در باقی‌مانده ۵۰ درصد مقالات می‌تواند باشد که شاید به عنوان موضوعات داغ آینده محسوب گردد، اما از این روش قابل تشخیص نیست. Ji و Zhi نیز در مطالعه خود با استفاده از روش آماری، به پیش‌بینی موضوعات آینده تالاب پرداختند و تنها نتوانستند موضوعات کلی را پیش‌بینی نمایند (۴).

نکته مورد توجه دیگر در پژوهش حاضر، محتوای سرعنوان‌های اساسی و روابط بین آن‌ها بود. از آن‌جا که این پژوهش به تحلیل محتوای مدارک مربوط به سبب‌شناسی سرطان کولون پرداخت، انتظار می‌رود که موضوعات اساسی آن مدارک، مرتبط با عوامل خطر ساز در ایجاد سرطان کولون باشد، اما با بررسی سرعنوان‌های موضوعی مورد استفاده در مدارک و خوشه‌بندی آن‌ها، این نتیجه به دست آمد که سرعنوان‌های موضوعی به کار رفته در مدارک، با تشخیص و درمان و علائم سرطان کولون نیز در ارتباط می‌باشند؛ چرا که هیچ کس دلایل دقیق ابتلا به سرطان کولون را نمی‌داند و پزشکان نیز اغلب نمی‌توانند توضیح دهند چرا شخصی به سرطان کولون مبتلا می‌گردد و شخص دیگری به این بیماری مبتلا نمی‌شود. با این حال، برخی دلایل که امکان ابتلا به سرطان

### تشکر و قدردانی

بدین وسیله از پشتیبانی و مساعدت همکاران گروه علم اطلاعات و دانش‌شناسی دانشگاه یزد به جهت ارایه مشاوره‌های علمی در طول اجرای این پژوهش، تشکر و قدردانی به عمل می‌آید.

مدارکی که موضوعات اساسی نوپیدای آینده بر اساس آن‌ها پیش‌بینی می‌شود، طولانی‌تر باشد و در مقابل، پیش‌بینی برای دوره کوتاه‌تری صورت گیرد، نتیجه دقیق‌تر خواهد بود. برای رسیدن به نتیجه مطلوب‌تر نیز بهتر است که چندین روش مختلف برای پیش‌بینی هم‌زمان مورد استفاده قرار گیرد.

### References

1. Tseng YH, Lin YI, Lee YY, Hung WC, Lee CH. A comparison of methods for detecting hot topics. *Scientometrics* 2009; 81(1): 73.
2. Price DJ. Networks of scientific papers. *Science* 1965; 149(3683): 510-5.
3. Noyons ECM, van Raan AFJ. Monitoring scientific developments from a dynamic perspective: Self-organized structuring to map neural network research. *J Am Soc Inf Sci* 1998; 49(1): 68-81.
4. National Institute of Science and Technology Policy (NISTEP). The 8<sup>th</sup> Science and Technology Foresight Survey- Study on Rapidly-developing Research Area -Interim Report. Tokyo, Japan: NISTEP; 2005.
5. Krampen G, von Eye A, Schui G. Forecasting trends of development of psychology from a bibliometric perspective. *Scientometrics* 2011; 87(3): 687-94.
6. Zhi W, Ji G. Constructed wetlands, 1991-2011: A review of research development, current trends, and future directions. *Sci Total Environ* 2012; 441: 19-27.
7. Chen C, Hu Z, Liu S, Tseng H. Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. *Expert Opin Biol Ther* 2012; 12(5): 593-608.
8. Glanzel W. Bibliometric methods for detecting and analysing emerging research topics. *El profesional de la informacion* 2012; 21(1): 194-201.
9. Reitz JR. Foundations of electromagnetic theory. 3<sup>rd</sup> ed. Reading, MA: Addison-Wesley; 1979.
10. Noukarizi M, Narmenji M. Introduction to information and communication. Tehran, Iran: Samt; 2013. [In Persian].
11. Zali MR. Rectum and colon cancer: Whatever patients should know about rectum and colon cancer. Tehran, Iran: Research Center for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences; 2005. [In Persian].
12. Salari A, Dehghan H. Evaluation and treatment of colorectal cancer in Shahid Rahnemoon and Afshar hospitals, Yazd-Iran. *J Shahid Sadoughi Univ Med Sci* 2007; 15(3): 20-2. [In Persian].
13. Qin J. Semantic similarities between a keyword database and a controlled vocabulary database: An investigation in the antibiotic resistance literature. *J Am Soc Inf Sci* 2000; 51(2): 166-80.

## Predicting Basic Concepts of a Field, Based on the Factors of Oldness and Frequency Use of Subject Terms: A Case Study on Colon Cancer

Mojtaba Mokhtari-Shamsi<sup>1</sup>, Mohammad Tavakolizadeh-Ravari<sup>2</sup>, Ebrahim Zalzadeh<sup>2</sup>, Mahmoud Baghbanian<sup>3</sup>

### Original Article

#### Abstract

**Introduction:** Prediction of future subjects is a necessity for policy making in science field, which is done by different variables. Accordingly, present research was conducted to predict basic concepts in the field of colon cancer based on the factor of oldness and frequency use of subject terms.

**Methods:** The present study was a descriptive-applied research. PubMed database was used to extract the data. Two separate searches were done on two 25-year-period studies, one from 1965 to 1989, and one from 1990 to 2015. In order to determine and predict important subjects of colon cancer domain (Colonic Neoplasms), two factors of oldness and annual frequency use of the terms were used. To do this, frequency and duration of the use of each of these terms were noticed in a period of 25 years. Then, climax or deterioration of the subject terms in future documents was predicted based on their reducing and increasing frequency in use during a special period.

**Results:** Studying the degree of conformity between predicted subjects and newly found basic subjects in colon cancer etiology during 1990 to 2015, 128 subjects were predicted as the key subjects for future out of which, 73 cases (58%) occurred in the second period. In addition, studying the degree of conformity between the mentioned 73 cases and the second period basic subjects indicated that predicted subject issues constitute 22% of the key subjects of the second period.

**Conclusion:** The results of the present research and the studies on predicting future trends showed that there was no method to predict the future subject issues with certainty and high accuracy. It seems that various methods of scientometrics and experts' views should be considered together in order to have a precise prediction.

**Keywords:** Medical Subject Headings; Bibliometrics; Colon Cancer

Received: 7 Feb, 2016

Accepted: 27 Nov, 2016

**Citation:** Mokhtari-Shamsi M, Tavakolizadeh-Ravari M, Zalzadeh E, Baghbanian M. **Predicting Basic Concepts of a Field, Based on the Factors of Oldness and Frequency Use of Subject Terms: A Case Study on Colon Cancer.** Health Inf Manage 2016; 13(5): 354-9.

Article resulted from MSc thesis No. 1157993 with the financial support of the University of Yazd.

1- MSc, Scientometrics, Department of Knowledge and Information Science, School of Social Sciences, University of Yazd, Yazd, Iran

2- Assistant Professor, Knowledge and Information Science, Department of Knowledge and Information Science, School of Social Sciences, University of Yazd, Yazd, Iran (Corresponding Author) Email: tavakoli@yazd.ac.ir

3- Assistant Professor, Hepatology and Gastroenterology, Department of Internal Sciences, School of Medicine, Shahid Sadooghi University of Medical Sciences, Yazd, Iran