ارزیابی صحت آزمایش‌های تشخیصی در بیماران مشکوک به آپاندیست حاد با روش تحلیل منحنی راک

دکتر مهرداد برهانی: * استادیار، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی جندی شاپور اهواز
دکتر کاظم محمد: استاد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت و انستیتو تحقیقات بهداشتی، دانشگاه علوم پزشکی تهران
دکتر کورش هلاکویی نائینی: استاد، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت و انستیتو تحقیقات بهداشتی، دانشگاه علوم پزشکی تهران
دکتر سید رضا مجد زاده: دانشیار، گروه اپیدمیولوژی و آمار زیستی، دانشکده بهداشت و انستیتو تحقیقات بهداشتی، دانشگاه علوم پزشکی تهران
دکتر احمد کاویانی: استادیار، بخش جراحی، بیمارستان امام خمینی(ره)، دانشگاه علوم پزشکی تهران

فصلنامه پایش
سال پنجم شماره چهارم پاییز ۱۳۸۵ صص ۲۶۲-۲۵۵
تاریخ پذیرش مقاله:۱۳۸۴/۹/۲۸

چکیده

هفت درصد از جمعیت در طول عمر خود دچار بیماری آپاندیسیت حاد می‌شوند، ولی همچنان معیارهای نیاز به عمل اورژانس مورد بحث و پژوهش است. لذا تعیین ترکیب مناسب از آزمایش‌ها و متغیرهای تشخیصی با صحت بالا از اهداف مطالعات مربوط به تصمیم‌گیری برای عمل آپاندیسیت محسوب می‌شود. هدف این مطالعه هم تعیین چنین ترکیب مناسبی برای تعیین افراد نیازمند به عمل آپاندیسیت به کمک اندازه گیری صحت تشخیص آزمایش‌ها است. با مراجعه به پرونده پزشکی و دفاتر آسیب شناسی، اطلاعات تمامی ۵۷۷ مورد شکم حاد مشکوک به آپاندیسیت مراجعه کننده به اورژانس بیمارستان امام خمینی (ره) تهران طی سال‌های ۷۹ تا ۸۳ گردآوری گردید. داده‌ها با استفاده از منحنی راک (Receiver Operating Characteristic Curve-ROC) و سطح زیر منحنی راک AUC (Area Under the Curve) با استاندارد نتیجه آسیب شناسی پس از عمل، پردازش و تحلیل شد. شاخص صحت تشخیصی آزمایش‌ها و متغیرها، مقدار سطح زیر منحنی راک است. در این مطالعه، ترکیب متغیرهای جنسیت، سابقه استفراغ، درصد پلی مورفونوکلئر (Poly Morpho Nuclear-PMN) در آزمایش گلبول‌های سفید خون، از نظر سطح زیر منحنی راک، حساسیت آزمایش‌ها و حداقل بودن تعداد متغیرها بهترین ترکیب آزمایش‌ها و متغیرها محسوب شده و صحت بالایی را برای تشخیص آپاندیسیت در بیماران مشکوک به آپاندیسیت دارای حساسیت شکمی فراهم می‌سازد (AUC= ۷۷/۳، 95%CI= ٪۷۰-٪۸۴). نقطه برش مناسب (Cut off Point) برای تعداد گلبول‌های سفید خون۷/۵ (10^3 در میکرو لیتر)، برای درصد PMN، ۶۵ درصد و برای درجه حرارت دهانی، ^{oc} ۳۷/۷ تعیین گردید. در گروه آزمایش (Test set)، مقدار کاپا تطابق برای تشخیص گذاری با معیار حاصل از تحلیل راک و نتیجه آسیب شناسی پس از عمل بیماران، برابر ۶۱ درصد (SD=٪۱۵) بود. این مطالعه نشان داد حتی با تعیین بهترین ترکیب برای آزمایش‌های تشخیصی رایج فعلی آپاندیسیت، باز هم نتایج تشخیص اولیه با نتایج آسیب شناسی بعد از عمل بیماران انطباق خیلی خوبی ندارد. برای کاهش بیشتر موارد مثبت کاذب (آپاندکتومی نابجا) و منفی کاذب (بیماران نیازمند آپاندکتومی که تشخیص داده نشده‌اند) پیشنهاد می‌گردد حداقل در گروه‌های پر خطر برای تشخیص ناصحیح آپاندیسیت (کودکان، زنان جوان و سالمندان) از روش‌های تشخیصی بیشتری استفاده گردد.

کلیدواژه‌ها: روش منحنی راک، سطح زیر منحنی راک، آپاندیسیت حاد، آزمایش تشخیصی

* نویسنده پاسخگو: اهواز، شهرک دانشگاهی، گلستان، دانشکده پزشکی، گروه پزشکی اجتماعی
تلفن: ۳۳۸۸۳۳۴-۰۶۱۱ نمابر: ۳۳۶۶۹۹۰-۰۶۱۱
E-mail: mborhany@yahoo.com

Archive of SID
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]
www.SID.ir

مقدمه

آپاندیسیت حاد، شایع‌ترین علت جراحی اورژانس شکمی است [۱، ۲]. تقریباً ۷ درصد از جمعیت، در طول عمر خود دچار آپاندیسیت حاد می‌شوند [۳]، ولی همه بیماران در مراحل اولیه، دارای علایم کلاسیک برای تصمیم گیری صحیح جراحان نبوده و امکان درصدهای بالایی از خطای تشخیصی وجود دارد[۲].

به‌طور معمول تصمیم گیری برای آپاندکتومی (برداشتن آپاندیس) در مراجعات درد حاد شکمی، همزمان با رد سایر علل درد شکمی انجام می‌پذیرد. تشخیص آپاندیسیت بر اساس علایم ابراز شده توسط بیمار (بی‌اشتهایی، تهوع، استفراغ و مهاجرت درد)، معاینات و علایم بالینی (تب، تندرنس و ریباند تندرنس، گاردینگ، علامت کاف و روزینگ و....) و آزمایشات سرمی (شمارش گلبول‌های سفید خونی و درصد پلی‌مورفونوکلئرها) صورت می‌پذیرد[۴]. هیچکدام از این علایم و آزمایش‌ها به تنهایی مختص آپاندیسیت حاد نبوده و از نظر ویژگی تشخیص، ویژگی بالایی ندارند. هر چند اصولاً برای آن که بیماران مبتلا، شانس عمل به موقع را از دست ندهند، نیازمند آزمایش‌هایی هستیم که در درجه نخست حساسیت بالایی داشته باشند و ویژگی بالا از اولویت دوم برخوردار است. به دلیل همین مسایل در حال حاضر حتی در مراکز پیشرفته پزشکی، تشخیص اولیه آپاندیسیت در ۱۶ تا ۱۸ درصد موارد با انجام آسیب شناسی پس از عمل رد می‌شود[۵].به این ترتیب لازم است با ترکیب مناسب آزمایش‌های مذکور بالاترین میزان صحت (حاصل تقسیم موارد مثبت و منفی حقیقی بر کل موارد تشخیص‌گذاری) و حساسیت و ویژگی آزمایش‌ها را مشخص نماییم. برای این کار ضمن تعیین بهترین نقطه برش (Cut-off point) در آزمایش‌های کمی (مثل تعداد گلبول‌های سفید) می‌توان با استفاده از روش تحلیل راک (ROC) یا منحنی مشخصه عملی گیرنده (Receiver Operating Characteristic Curve)، میزان صحت هر آزمایش و ترکیب آزمایش‌ها را تعیین کرد. در تحلیل راک، صحت هر آزمایش با مقدار AUC (Area Under the Curve) یا سطح زیر منحنی راک (از ۰ تا۱۰۰ درصد) بیان می‌شود و امکان نمایش گرافیکی مقایسه صحت آزمایش‌ها وجود دارد[۶].

هدف ما از این مطالعه آن است که ضمن تعیین نقطه برش مناسب آزمایش‌های تشخیصی در مراجعات مشکوک به آپاندیسیت حاد، بهترین ترکیب این آزمایش‌ها را با بالاترین مقدار صحت، به کمک تحلیل منحنی راک مشخص نموده و با نتایج تحلیل رگرسیون لجستیک و نتایج آسیب شناسی پس از عمل، مقایسه نماییم.

مواد و روش کار

اطلاعات پرونده تمامی ۵۷۷ مراجعه کننده مشکوک به آپاندیسیت حاد که طی سال‌های ۷۹ تا ۸۳ در بخش‌های جراحی بیمارستان امام خمینی (ره) به دلیل وجود حساسیت شکمی و شک به آپاندیسیت حاد مورد آپاندکتومی قرار گرفته‌اند مورد بررسی قرار گرفت.

از ۵۷۷ بیمار مذکور گزارش آسیب شناسی (استاندارد طلایی تشخیص در این مطالعه) آپاندیسیت حاد را در ۱۲۵ مورد تأیید نکرده بود. اما آپاندیسیت حاد چرکی (در برخی موارد همراه فکالیت، گانگرن و یا پرفوراسیون) در ۴۵۲ مورد تأیید شده بود.

علاوه بر نتیجه آسیب شناسی آپاندیس عمل شده، اطلاعات متغیرهای تهوع، استفراغ، مهاجرت درد، بی‌اشتهایی، سن، جنس، درجه حرارت دهانی، شمارش گلبول‌های سفید خون و درصد پلی مورفونوکلئر در لکوسیت‌های خونی و علایم معاینه شکمی (ریباند تندرنس یا حساسیت برگشتی، گاردینگ، روزینگ و...) جمع‌آوری شد.

این داده‌ها یک بار با تشکیل مدل رگرسیون لجستیک و استفاده از ضرایب این مدل به‌عنوان وزن هر متغیر (بر اساس روش‌های مرسوم) در تشخیص آپاندیسیت حاد مورد استفاده قرار گرفت [۷، ۸] و سپس بار دیگر داده‌های مذکور برای تعیین نقطه برش مناسب آزمایش‌های کمی (با تمام ۵۷۷ نمونه) و تعیین صحت ترکیب آزمایش‌ها و انتخاب بهترین آزمایش‌های تشخیصی (در نمونه‌ای ۱۴۸ تایی) براساس مراجع معتبر استفاده گردید[۱۱-۹].

در نهایت برای مقایسه نتایج تشخیص‌گذاری با دو معیار تشخیصی حاصل از دو روش (رگرسیون لجستیک و منحنی راک)، ابتدا یک گروه ۴۳ نفره به عنوان گروه آزمایش (Test set) انتخاب شد. سپس برای هر بیمار براساس معیار راک و معیار حاصل از مدل لجستیک، تشخیص‌گذاری شده و نهایتاً تشخیص‌های پیشنهادی با نتایج آسیب شناسی مقایسه گردید.

مقدار انطباق نتایج تشخیص‌گذاری دو روش با یکدیگر و با نتایج آسیب شناسی توسط شاخص توافق کاپا اندازه‌گیری گردید. برای محاسبات پیشرفته، از نرم‌افزارهای Stata 8 و Med calc 8 استفاده شد.

www.SID.ir
Archive of SID
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

يافته‌ها

در این بررسی سن بیماران از ۵ تا ۸۰ سال با میانگین ۲۶سال (انحراف معیار۱۲/۳) و نسبت جنسی مرد به زن تقریباً ۱/۵ بود.

در کسانی که ابتلای آنها به آپاندیسیت با آسیب شناسی تأیید شده، میانگین سنی ۲۶/۷ (خطای معیار ۰/۶) سال است.۶۰ درصد از موارد، تهوع و ۵۱ درصد استفراغ و بی‌اشتهایی داشته‌اند. مهاجرت درد و ریباند تندرنس (حساسیت برگشتی) در ۵۰ درصد از موارد و گاردینگ در لمس شکم در۱۰/۵ درصد از بیماران گزارش شده است (علامت کاف و روزینگ به ترتیب در۶ و ۳ درصد بیماران). در گروهی که آپاندیسیت حاد آنها در آسیب شناسی تأیید نشده است، ۶۵ درصد زن بوده و ۲۱/۵ درصد در سنین ۱۶ تا ۱۸ سال هستند (شایع‌ترین گروه سنی). سایر مشخصات توصیفی موارد آپاندکتومی در جداول شماره ۱ و ۲ آمده است. بهترین مدل رگرسیون لجستیکی که با داده‌های مذکور انطباق دارد در جدول شماره ۳ نشان داده شده است (متغیرهایی که در جدول وجود ندارند، حضورشان در مدل معنی‌دار نبوده است).

بهترین نقطه برش آزمایش‌های کمی بر اساس ماکزیموم بودن مثبت حقیقی و منفی حقیقی و با استفاده از شاخص بودن (حساسیت + ویژگی -۱=Youden index) [۱۲] به شرط تأمین حساسیت حدود ۹۰ درصد محاسبه گردید (جدول شماره ۴). در مورد تب، استثنائاً نقطه برش معمول (۳۷/۷ درجه سانتیگراد) شرایطی بهتر از نقطه به‌دست آمده با ضوابط مذکور داشته و در تحلیل‌های بعدی، ملاک تب داشتن، حرارت دهانی بالاتر از ۳۷/۷ درجه است. برای هر متغیر و ترکیب متغیرها مقدار سطح زیر منحنی راک (AUC) به عنوان شاخص صحت آزمایش‌ها محاسبه شد که ترکیب‌های با صحت تشخیصی بالا در جدول شماره ۵ مشاهده می‌گردند. با توجه به مندرجات این جدول، بهترین ترکیب متغیرها برای تشخیص آپاندیسیت بر اساس مقدار AUC و تأمین حساسیت بالا و نیز تعداد متغیر کمتر، چنانچه مشاهده می‌شود ترکیب: جنسیت+ درصد PMN + استفراغ است. لذا اگر بر این اساس بخواهیم گروه در معرض خطر (Risk Set) برای آپاندیسیت را تعریف کنیم، می‌توانیم مردان مشکوک به آپاندیسیتی که حداقل سابقه استفراغ را دارند و یا درصد PMN آزمایش آنها ۶۵ درصد یا بالاتر است و زنانی که حداقل دارای هر دو ویژگی (استفراغ و درصد بالایPMN) هستند، به‌عنوان گروه در معرض خطر مورد عمل آپاندکتومی قرار دهیم. اگر ضرایب مدل لجستیک با روش امتیاز دهی، مشابه آنچه در مرجع [۷] برای داده‌های مطالعه فرامینگهام انجام پذیرفته، ملاک امتیاز دهی هر متغیر باشد (ضمیمه الف)، مهاجرت درد، استفراغ، بالا بودن گلبول های سفید (مساوی یا بیش از۷/۵ هزار) هرکدام ۱ امتیاز، بالا بودن درصد PMN (۶۵ درصد یا بالاتر) ۲ امتیاز، داشتن تب (دمای دهانی بیش از ۳۷/۷ درجه سانتیگراد) ۴ امتیاز و جنسیت مرد ۳ امتیاز خواهند داشت و مردان با داشتن حداقل یکی از این علایم (حداقل ۴ امتیاز) و زنان اگر حداقل تب داشته باشند و یا با داشتن حداقل ۳ علامت (به‌شرط آن که هر سه علامت یک امتیازی نباشند) می‌توانند حداقل ۴ امتیاز کسب نموده و به عنوان گروه در معرض خطر، شناخته شوند. اگر معیار های تشخیصی به‌دست آمده با دو روش رگرسیون لجستیک و تحلیل راک را در یک گروه آزمایش ۴۳ نفره (انتخاب تصادفی از بین بیمارانی که اطلاعات آنها کامل بوده) به‌کار بندیم و با نتایج آسیب شناسی مقایسه نماییم، منحنی‌های راک (شکل ارایه شده) به‌دست می‌آید. چنانچه مشاهده می‌گردد هر دو معیار سطح زیر منحنی بالایی دارند و به عنوان روش‌های تشخیصی می‌توانند مطرح باشند. از طرفی برتری گرافیکی، معیار برگرفته از تحلیل راک (بالاتر بودن این منحنی از منحنی روش لجستیک) در آزمون معنی‌داری اختلاف صحت (مقدار درصدAUC) تأیید نمی‌گردد (P=۰/۱۸۵). نتایج تشخیص‌گذاری با دو معیار تشخیصی (راک و لجستیک) با هم انطباق و توافق خیلی خوبی دارند (کاپا ۹۳ درصد با انحراف معیار ۱۵ درصد) اما توافق هیچکدام از دو روش با نتایج آسیب شناسی خیلی خوب و ایده‌آل نیست (کاپای هر دو روش با توجه به جداول شماره ۶ ، ۷ و ۸ کمتر از ۸۰ درصد است).

جدول شماره ۱- توزیع متغیرهای کمی پیوسته در ۵۷۷ بیمار تحت مطالعه

	حداکثر	حداقل	میانگین	انحراف معیار
سن (سال)	۸۰	۵	۲۶	۱۲/۳
حرارت بدن (درجه سانتیگراد)	۴۰	۳۶	۳۷/۴	۰/۵
شمارش گلبول سفید (۱۰۰۰ در میکرولیتر)	۳۴/۳	۴/۱	۱۲/۴	۴/۵
درصد پلی مورفو نوکلئر	۹۶	۲	۷۶	۱۱/۷

Archive of SID
www.SID.ir
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

جدول شماره ۲- فراوانی بر حسب درصد در متغیرهای کیفی

	آپاندیسیت حاد چرکی -	آپاندیسیت حاد چرکی +
نتیجه آسیب شناسی	۲۱/۷	۷۸/۳
استفراغ	ندارد ۵۱/۸	دارد ۴۸/۲
تهوع	ندارد ۴۰	دارد ۶۰
مهاجرت درد	ندارد ۵۴/۱	دارد ۴۵/۹
بی‌اشتهایی	ندارد ۵۰/۳	دارد ۴۹/۶
حساسیت برگشتی	ندارد ۵۰/۳	دارد ۴۷/۵ (مشکوک ۲/۳)
سایر علایم (گاردینگ، کاف، ...)	ندارد ۶۹/۳	دارد ۳۰/۷
جنسیت	زن ۴۰/۶	مرد ۵۹/۴

جدول شماره ۳- نسبت شانس (Odds Ratio) با معنی استخراجی از مدل رگرسیون لجستیک در ۵۷۷ نمونه برای پیامد نتیجه آسیب شناسی

(پس از حذف متغیرهای سن، بی‌اشتهایی، تهوع، حساسیت برگشتی و سایر علایم که در مدل کلی اولیه بی‌معنی بوده‌اند)

	نسبت برتری (OR) و ضریب [B]	احتمال (P)	فاصله اطمینان ۹۵٪ برای OR
مهاجرت درد (دارد، ندارد)	(۱/۹) [۰/۶۲]	۰/۰۴۹	۳/۴-۱
استفراغ (دارد، ندارد)	(۲/۳) [۰/۸۳]	۰/۰۰۸	۴/۳-۱/۲
تب (دارد، ندارد)	(۱۳) [۲/۶]	۰/۰۰۳	۵۹-۳
گلبول سفید (۱۰۰۰ در میکرولیتر)	(۱/۱) [۰/۰۸]	۰/۰۴	۱/۲-۱
پلی‌مورفونوکلئر (درصد)	(۱/۰۵) [۰/۰۵]	۰/۰۰۰۱	۱/۰۸-۱/۰۲
جنسیت (زن، مرد)	(۴/۸) [۱/۶]	۰/۰۰۰۱	۸/۴-۲/۶

مقدار ثابت مدل۵- (B_0)= (CONSTANT)

جدول شماره ۴- خصوصیات نقاط برش انتخابی برای آزمایش‌های کمی پیوسته

	نقطه برش	حساسیت	ویژگی	LR^{+*}	LR^{-*}
شمارش گلبول سفید خون (۱۰۰۰ در میکرولیتر)	۷/۵	۹۰/۲	۲۶/۴	۱/۲	۰/۳۷
پلی‌مورفونوکلئر (درصد)	۶۵	۹۱/۴	۳۹/۲	۱/۵	۰/۲۲
تب (درجه سانتیگراد)	۳۷	۹۵/۵	۴	۱	۱/۱
سن (سال)	۱۴	۹۳/۶	۹/۶	۱	۰/۶۷

* LR: نسبت درست‌نمایی + و -

Archive of SID
www.SID.ir
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

جدول شماره ۵- مقدار سطح زیر منحنی راک (AUC) برای ترکیب دو یا چند آزمایش و خصوصیت

	مقدار AUC بر حسب 95%CI
پلی‌مورفونوکلئر و گلبول سفید (درصد)	۶۳/۳ (۷۱-۵۵)
پلی‌مورفونوکلئر (درصد) و تب	۷۰/۴ (۷۷/۶-۶۲/۳)
پلی‌مورفونوکلئر (درصد)، حساسیت برگشتی، مهاجرت درد و استفراغ	۷۱/۴ (۷۸/۵-۶۳/۴)
پلی‌مورفونوکلئر (درصد)، مهاجرت درد و استفراغ	۷۳/۴ (۸۰/۳-۶۵/۵)
استفراغ و مهاجرت درد	۶۵/۶ (۶۸-۵۱/۴)
مهاجرت درد و حساسیت برگشتی	۶۰ (۶۸-۵۱/۴)
استفراغ و جنسیت	۷۴/۲ (۸۱-۶۶/۳)
استفراغ و پلی‌مورفونوکلئر (درصد)	۶۹/۵ (۷۶/۸-۶۱/۴)
پلی‌مورفونوکلئر (درصد) و جنسیت	۷۴/۳ (۸۱-۶۶/۵)
*پلی‌مورفونوکلئر (درصد)، استفراغ و جنسیت	۷۷/۳ (۸۴-۷۰)
پلی‌مورفونوکلئر (درصد)، استفراغ، جنسیت و مهاجرت درد	۷۹/۶ (۸۵/۸-۷۲/۲)
استفراغ، جنسیت، تب، مهاجرت درد و پلی‌مورفونوکلئر (درصد)	۸۲/۳ (۸۸-۷۵)

* آزمون معنی‌داری اختلاف بین سطح زیر منحنی راک برای متغیرهای استفراغ، جنسیت و درصد پلی‌مورفونوکلئر با AUCهای مدل‌های ۴ و ۵ متغیری (دو سطر آخر جدول) بیانگر معنی‌دار نبودن اختلاف با مدل ۴ متغیری (P=۰/۳۳) و اختلاف معنی‌دار با مدل ۵ متغیری است (P=۰/۰۷ با سطح قضاوت P ≤۰/۳۳).

جدول شماره ۶- نتایج تطابق روش‌های مختلف با نتیجه آسیب شناسی و با یکدیگر در گروه آزمایش (Test Set)

	کاپا بر حسب درصد	انحراف معیار بر حسب درصد	احتمال P
روش لجستیک با AUC راک	۹۳	۱۵	۰/۰۰۰۰۱
روش لجستیک با آسیب شناسی	۵۳	۱۵	۰/۰۰۰۳
AUC راک با آسیب شناسی	۶۱	۱۵	۰/۰۰۳

جدول شماره ۷- نتایج تشخیص‌گذاری با معیارهای حاصل از تحلیل راک در مقایسه با نتایج آسیب شناسی در گروه آزمایش

	آسیب شناسی -	+	جمع
روش راک			
تشخیص -	۷	۳	۱۰
تشخیص +	۳	۳۰	۳۳
جمع	۱۰	۳۳	۴۳

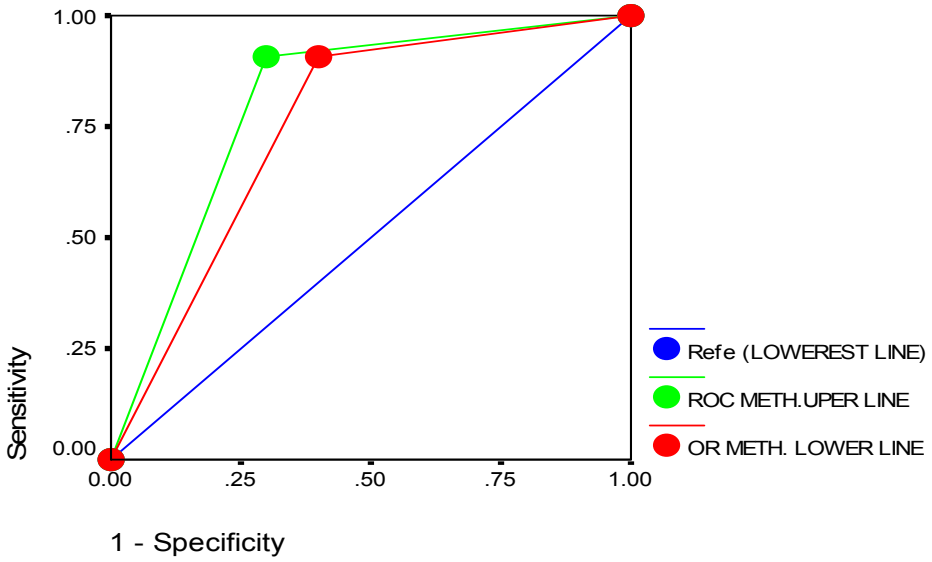
جدول شماره ۸- نتایج تشخیص‌گذاری با معیارهای حاصل از OR مدل رگرسیون لجستیک در مقایسه با نتایج آسیب شناسی در گروه آزمایش

	آسیب شناسی -	+	جمع
روش OR			
تشخیص -	۶	۳	۹
تشخیص +	۴	۳۰	۳۴
جمع	۱۰	۳۳	۴۳

Archive of SID
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

www.SID.ir

منحنی راک و سطح زیر آن برای دو روش تشخیصی*



*منحنی بالاتر مربوط به معیار حاصل از تحلیل راک (ROC METH.) و منحنی پایین‌تر مربوط بـه معیـار رگرسـیون لجستیک (OR METH.) است. تفاوت سطح زیر منحنی راک برای این دو روش معنی‌دار نیست و با فاصـله اطمینـان ۹۵٪، سطح زیر منحنی برای روش راک برابر ۶۲ تا ۹۸ درصد می‌باشد.

بحث و نتیجه‌گیری

نتایج توصیفی این پژوهش به یـک مطالعـه قبلـی در بیمارسـتان سینا تهران نزدیـک اسـت [۱۳]، ولـی یافتـه اصـلی مـا معیارهـای تشخیصی جدیدی برای تعیین گـروه در معـرض خطـر آپاندیـسیت حاد است.

اصولاً تحلیل راک برای متغیرهـای تشخیـصی در آپاندیـسیت بـه تعداد محدودی انجام شـده اسـت [۱۶-۱۴] ولـی همزمـان در ایـن مطالعات نقطه برش مناسب، تعیین سـطح زیـر منحنـی راک بـرای تمام متغیرها و ترکیب آنهـا (بـا محاسـبه فاصـله اطمینـان) انجـام نپذیرفته است [۱] و از امتیازات این مطالعه انجام تمام ایـن امـور و حصول راهکار برای تعیین گروه نیازمند بـه آپانـدکتومی مـی‌باشـد. دستورالعمل‌های حاصل در این مطالعه با آنچه سیستم امتیاز دهـی تعـدیل شـده آلـوارادو (Modified Alvarado Score) نامیـده می‌شود [۲] همخوانی ندارد. امتیاز دهی آلـوارادو یـک سیـستم ۱۰ امتیازی است که کسب ۷ امتیاز و بیشتر، بیمـار را کاندیـدای عمـل آپاندکتومی می‌کند. همه علایم (مهاجرت درد، تهوع و استفراغ، تب، بـی‌اشـتهایی، حـساسیت برگـشتی و افـزایش PMN) در ایـن امتیازگذاری، یک امتیازی هستند ولی حساسیت ربع راست تحتانی شکم و لکوسیتوز دو امتیاز می‌گیرند.

نقطه برش برای تب با درجه حرارت دهانی بالای ۳۷/۲ سانتیگراد و برای درصد پلی مورفو نوکلوئرها، بالای ۷۴ درصد است[۲]. چنانچه دیده می‌شود امتیاز متغیرها در امتیـازدهی آلـوارادو بـا مطالعـه مـا متفاوت است. علاوه بر جمعیت مـورد مطالعـه متفـاوت، تفـاوت در نقاط برش آزمایش‌های تشخیصی و توجه مطالعـه مـا بـه جنـسیت بیماران و مشکل بالاتر بودن خطای تشخیص‌ها در زنان (به واسـطه تعدد تشخیص‌های افتراقی) و لزوم تفاوت ملاک تشخیصی در مرد و زن، مواردی است که می‌تواند توجیه کننده اختلاف یافته‌هـای ایـن مطالعه با توصیه‌های تشخیصی امتیازدهی آلوارادو باشد.

یک مطالعه ۵۴۴ موردی آینده‌نگر در نـروژ [۱] در خـصوص ارزش تشخیصی تب و شاخص‌های التهابی با استفاده از تحلیل راک، مقدار AUC را برای ترکیب این آزمایش‌ها محاسبه نمـوده و مقـدار آن را کمتر از ۷۰ درصد دانسته است.

مطالعه دیگری هم [۱۶] با روش تحلیل راک نتایج مشابهی به‌دست آورده است. در مطالعه ما با لحاظ کردن شرط بالا بـودن حـساسیت آزمایش، نقطه برش پیشنهادی برای شمارش گلبول‌هـای سـفید بـا

[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]
Archive of SID
www.SID.ir

مطالعه نروژ متفاوت است، اما باز هم بهترین ترکیب تشخیصی برای آزمایش‌های التهابی نتوانسته صحتی بیش از ۷۰/۴ درصد ایجاد کند (جدول شماره ۵) و اختلاف واضحی بین نتایج این تحقیقات وجود ندارد. از طرفی معیارهای حاصل از تحلیل راک حتی با معیارهای حاصله از مدل رگرسیون لجستیک ساخته شده در مطالعه ما در گروه آزمایش، تفاوت معنی‌داری از نظر صحت تشخیصی و انطباق با نتایج آسیب شناسی نشان نداده است (به شکل منحنی‌های راک حاصله نگاه کنید). علاوه بر تورش‌های (Bias) رایج در مطالعات تحلیلی، در مطالعات آزمایش‌های تشخیصی و از جمله در این تحقیق به‌طور خاص، سه نوع تورش باید مورد نظر باشد. اول ورود درصد بالایی از بیمارانی که تابلو واضح بالینی دارند به مطالعه: این وضعیت باعث می‌شود بیمار آگاه از شرایط خود علایم کلاسیک بیماری را ابراز کند و معالجه‌گران بیش از معمول، علایم رایج و کلاسیک را جستجو و اعلام کنند[۱۷]. برای کاهش این مشکل، بیمارانی که قبل از مراجعه به بیمارستان محل تحقیق ما در مراکز دیگر بررسی شده و به‌عنوان آپاندیسیت محتمل جهت انجام عمل معرفی شده بودند و یا بیمارانی که با تابلوی آپاندیسیت حاد پرفوره مراجعه نموده بودند (جمعاً ۲۳ نفر) از مطالعه کنار گذاشته شده و نمونه‌ها به افراد مشکوک به آپاندیسیت (و نه بیماران به تشخیص رسیده در سایر مراکز) منحصر گردید. تورش خاص دوم، تورش تعیین بیماری است (Verification Bias). در صورتی که آزمایش تشخیص نهایی (استاندارد طلایی) گران و یا تهاجمی بوده و انجام آن برای شروع درمان الزامی نباشد، تعدادی از بیماران ممکن است از انجام آزمون استاندارد طلایی خودداری نموده و از مطالعه کنار گذاشته شوند[۱۱].

خوشبختانه در مورد تشخیص قطعی آپاندیسیت، نمونه‌گیری آسیب شناسی حین عمل درمانی برای همه بیماران انجام می‌گیرد و درمان جراحی آپاندیسیت حاد به‌دلیل ارزانی نسبی در مراکز دولتی و کم خطر و کم عارضه بودن عمل، معمولاً با مخالفت بیماران روبرو نمی‌شود. تورش خاص سوم در مطالعات اپیدمیولوژی آزمایش‌های تشخیصی، تورش استاندارد ناکامل است (Imperfect standard bias): این تورش هنگامی رخ می‌دهد که محقق به جای استفاده از استاندارد طلایی، از آزمایشی ارزان‌تر یا کمتر تهاجمی و یا سریع‌تر استفاده نماید تا مشکل تورش تعیین بیماری را حل کند (مثل استفاده از کیت‌های آنتی ژنی تشخیص میکروبی به جای استفاده از کشت میکروبی ترشحات حلق). به دلیل الزام انجام بررسی آسیب شناسی در تمامی موارد آپاندکتومی، این مطالعه در معرض تورش استاندارد ناکامل قرار ندارد. با وجود مزایای این تحقیق و هرچند که ثبت اطلاعات بیماران نامزد جراحی معمولاً با دقت انجام می‌گیرد، ولی مثل سایر مطالعات گذشته نگر این تحقیق هم در معرض خطای ثبت اشتباه و ناکامل اطلاعات قرار داشته و لازم است نتایج آن در یک کارآزمایی مداخله‌ای تصادفی شده به‌کار گرفته شده و مورد تحلیل قرار گیرد. از طرفی چنان که دیدیم حتی با کاربرد دو روش تحلیل متفاوت و انتخاب ترکیباتی با بالاترین سطح ممکن برای تشخیص صحیح، باز هم ۲۲/۷ درصد از تشخیص‌ها می‌تواند ناصحیح (مثبت یا منفی کاذب) باشد و لازم است با افزایش آزمون‌های با صحت تشخیصی بالا، در جستجوی ترکیبات مناسب‌تری از آزمایش‌های رایج فعلی باشیم. برای حل مشکلات جاری در تشخیص آپاندیسیت (و کاهش موارد منفی کاذب و مثبت کاذب)، تعیین ترکیب بهینه آزمون‌ها و متغیرهای تشخیصی مرسوم، با روش تحلیل راک هم کارآمد نبوده و پیشنهاد می‌گردد در گروه‌های کودکان، سالمندان و زنان جوان از آزمایش‌های سرمی دیگر و یا روش‌های تصویری بیشتری استفاده نمود و افزایش صحت تشخیص در آن شرایط را اندازه گیری کرد.

تشکر و قدردانی

دانشکده بهداشت و انستیتو تحقیقات بهداشتی دانشگاه علوم پزشکی تهران هزینه این پژوهش را از اعتبارات پژوهشی تأمین نموده است. نگارندگان مقاله از بخش‌های جراحی، پاتولوژی مرکزی و واحد آمار و مدارک پزشکی بیمارستان امام خمینی (ره) تهران که زمینه عملی تحقیق را فراهم نمودند، مراتب قدردانی خود را اعلام می‌نمایند.

Archive of SID
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

www.SID.ir

منابع

1- Korner H, Soreide J. Diagnostic accuracy of inflammatory markers in patients suspected acute appendicitis. The European Journal of Surgery 1999; 165: 679-85

2- Arzu Denizbasi, Erol Erden Unluer. The role of the emergency medicine resident using the Alvarado score in the diagnosis of acute appendicitis. European Journal of Emergency Medicine 2003, 10: 296-301

3- Way L, Current of surgery. LANGE medical book, 2003

4- Elangovan S. Clinical and laboratory finding in acute appendicitis in the elderly. The Journal of American Board of Family Practice 1996; 9: 75-8

۵- امامی میبدی مصطفی، ترجمه اصول جراحی شوارتز ۱۹۹۹، چراغ دانش، چاپ دوم، تهران، ۱۳۸۲

6- Massimo M. Comparison of 3 clinical models for predicting the probability of pulmonary embolism. Medicine 2005; 84: 107-14

7- Sullivan L, Massaro J. Presentation of multivariate data for clinical use: the Framingham study risk score functions. Statistics in Medicine 2004; 23: 1631-60

8- Afina S, Lijmer J. The diagnostic odds ratio: a single indicator of test performance. Journal of Clinical Epidemiology 2003; 56: 1129-35

9- Greiner M, Pfeiffer D, Smith R. Principles and practical application of the ROC analysis for diagnostic tests. Prevention Veterinary Medicine 2000; 45: 23-41

10- Pepe S, Thompson M. Combining diagnostic test results to increase accuracy. Biostatistics 2000; 1: 123-40

11- Zhou X, Obuchowski N. Statistical methods in diagnostic medicine. John Wiley& Sons Inc, 2002

12- Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. Epidemiology 2005; 16: 73-81

۱۳- ظفرقندی محمد رضا، قراگوزلو پرهام، نوع پرست مریم، معزاردلان کوروش، .بررسی آینده نگر نتایج۲۱۴ جراحی با تشخیص آپاندیسیت از میان ۶۱۱ مورد شکم حاد بیمارستان سینا، مجله علمی نظام پزشکی، ۸۲، ۲۱، ۳۳-۲۷

14- Van Dieijen, Brombacher PJ. The value of lab tests in patient suspected of acute appendicitis. Clinical Chemistry 1991; 29: 749-52

15- Gronroos JM, Forsstrom J. Phosphatase A2, CRP, and white blood count in diagnosis of acute appendicitis. Clinical Chemistry 1994; 40: 1775-60

16- Paajanen H, Mansika A. Are serum inflammatory markers age dependent in acute appendicitis? Journal of the American College of Surgens 1996; 184: 303-8

17- Andersson REB. Meta analysis of clinical & lab: diagnosis of appendicitis. The British Journal of Surgery 2004; 91: 28-37

Archive of SID
[Downloaded from payeshjournal.ir on 2026-11-01]
[DOR: 20.1001.1.16807626.1385.5.4.3.9]

www.SID.ir