

پژوهش در جراحی

حساسیت، ویژگی و ارزش پیش‌بینی

ترجمه و تلخیص: دکتر سیدعباس میرمالک* و گروه مترجمین**

اغلب اوقات در درک مفاهیم واژه‌های حساسیت (Sensitivity)، ویژگی (Specificity) و ارزش پیش‌بینی (Predictive Value) با اشکال روبرو هستیم و به طور غریزی نمره‌های بالاتر را در یک آزمایش بالینی با ارزش‌تر تلقی می‌کنیم! و خود درک این معنی است که مفاهیم توصیفی تا به شکل ریاضی تعریف نشوند، قابل فهم نبوده و تا به شکل اعداد و منحنی ارائه نشوند، کاربردی نیستند. این فصل توصیفی برای واژه‌های حساسیت و ویژگی و ارزش پیش‌بینی است.

مقدمه

کار پیشگامانه یاروشالمی (Yerushalmy) در مورد تغییرپذیری مشاهده‌گر در تفسیر تصاویر رادیوگرافی قفسه سینه، تمایل رو به گسترشی را در ارزیابی روند تشخیصی آغاز کرد. او، واژه حساسیت (Sensitivity) و تخصیص (Specificity) را به عنوان معیارهای اعتبار آزمون‌های تشخیصی معرفی کرد و مفهوم ارزش پیش‌بینی (Predictive Value) را با مفاهیم ریاضی برای اعتماد به آزمایش تشخیصی در طبابت بالینی، تعریف کرد.

اعتبار آزمایش تشخیصی

پزشکان بالینی از آزمون‌های تشخیصی برای تعیین اینکه آیا یک بیماری وجود دارد یا خیر؟ استفاده می‌کنند. برای مثال شما می‌توانید از اسکن گالیوم برای آزمون وجود یا عدم وجود آبسه‌های داخل شکمی استفاده کنید و یا ماموگرافی می‌تواند به شما در تعیین این که آیا یک توده قابل لمس، بدخیم است یا خیر؟ کمک کند.

موقعیت بیماری که آزمون تشخیصی قصد تعیین آن را داشته، گاهی اوقات به یک «بیماری مورد نظر» باز می‌گردد. اعتبار یک آزمایش تشخیصی به سادگی به قدرت آن در ثبت نتایج غیر طبیعی برای بیمارانی که در آنها بیماری مورد نظر وجود دارد و ثبت نتایج در بیمارانی که در آنها بیماری هدف وجود ندارد، باز می‌گردد. یک آزمایش تشخیصی مطلوب هر دو موقعیت را نشان می‌دهد، به این ترتیب که نتایج غیر طبیعی را تنها برای بیمارانی که بیماری مورد نظر را دارند ثبت می‌کنند و نتایج طبیعی تنها برای بیمارانی که فاقد بیماری مورد نظر هستند. چنین آزمایش تشخیصی از این نظر که نتایج آن در مورد وجود یا فقدان بیماری مورد نظر کاملاً متضمن پیش‌بینی است، آزمایش کاملی می‌باشد. متأسفانه بیشتر آزمایش‌های تشخیصی به این خوبی انجام نمی‌شوند، و تعدادی از

نویسنده پاسخگو: دکتر سیدعباس میرمالک

تلفن: ۸۸۷۸۷۵۶۱

Email: SAM@Mirmalek.net

* استادیار گروه جراحی عمومی، دانشگاه علوم پزشکی آزاد اسلامی، واحد تهران

** دکتر لیلا پرورش، دکتر شقایق تهرانی، دکتر پوریا حسینی، دکتر پانته آرضانیان، دکتر مریم سعیدیان،

دکتر مروا طهماسبی‌راد، دکتر علی غلامرضائزاد، دکتر مهدی کلانتری، دکتر الهام کنی، دکتر امیر تیمور مرعشی،

دکتر امید میرمطلبی، دکتر علی ناظمیان

تاریخ وصول: ۱۳۸۷/۰۴/۰۹

مفاهیم و تکنیک‌ها برای اندازه‌گیری دقیق اینکه چگونه یک آزمایش تشخیصی معین به شکلی رضایت‌بخش انجام می‌شوند، بسط داده شده‌اند.

برای قضاوت در مورد این که چگونه یک آزمایش برای پیش‌بینی بیماری مورد نظر در بیماران معین، انجام می‌گیرد، ما نیاز به بعضی روش‌ها برای تعیین درستی وجود و یا فقدان بیماری، در همان بیماران داریم. برای قضاوت در مورد قابلیت اسکن گالیوم در تشخیص آبه‌های داخل شکمی، نیاز داریم که بدانیم آیا چنین آبه‌هایی در یک گروه از بیماران وجود دارد یا خیر؟ تا اینکه بتوانیم این اطلاعات را با نتایج به دست آمده از اسکن گالیوم مقایسه کنیم. می‌توانیم از یافته‌های لاپاراتومی و تأیید نتایج آسیب‌شناختی آبه به عنوان عامل تعیین کننده موقعیت واقعی بیماری مورد نظر استفاده کنیم. همچنین می‌توانیم از نتایج لاپاراتومی به عنوان عاملی تعیین کننده برای وجود صدمات داخل صفاقی آن به منظور ارزیابی کاربرد لاواژ صفاقی به عنوان یک مانور تشخیصی استفاده کنیم. در ارزیابی ماموگرافی به عنوان وسیله شناسایی تومورهای بدخیم پستان، می‌توانیم از آزمون‌های پاتولوژیکی تومور به عنوان عامل تعیین کننده خود استفاده کنیم. روشی که برای تأیید وجود یا فقدان بیماری مورد نظر در چنین تشخیص‌هایی به کار می‌رود و به عنوان استانداردهای طلائی (Gold Standard) شناخته می‌شود، می‌تواند بهترین استاندارد بالینی رایج و در دسترس باشد.

هنگامیکه کاوش‌های جراحی و تأییدیه‌های آسیب‌شناختی، بخشی از کنترل مدیریت متداول بالینی باشند، انتخاب‌های آشکاری از استانداردهای طلائی هستند؛ در غیر این صورت سایر استانداردها باید مورد استفاده قرار گیرند. برای تعیین اعتبار تکنیک‌های اسکن جدیدتر به عنوان آزمایش‌های تشخیصی ترومبوز ورید عمقی (DVT) بهترین استاندارد بالینی در دسترس، ونوگرافی است. نتیجتاً از نتایج این روند رادیولوژیکی معمولاً به عنوان استاندارد طلائی برای وجود یا فقدان DVT استفاده می‌شود. همچنین برای بررسی اعتبار چنین آزمایش‌هایی مانند اسکن‌های رادیونوکلئیدی در تشخیص تنگی‌های شرایین کرونری، که تأییدیه‌های جراحی تنها در تعداد نسبتاً کمی از بیمارانی که برای جراحی بای پس مراجعه می‌کنند در دسترس است، آنژیوگرافی کرونری اغلب به عنوان استاندارد طلائی استفاده می‌شود.

تصویر ۱- ساختار ارزیابی یک آزمایش تشخیصی را در واژه‌های کلی تنظیم می‌کند. به شکل اصولی، ما صرفاً نتایج آزمایش‌های تشخیصی را (غیر طبیعی در مقابل طبیعی) با نتایج استاندارد طلائی برای بیماری مورد نظر (وجود بیماری در مقابل فقدان آن) مقایسه می‌کنیم. نتیجه آزمایش در ردیف‌های یک جدول ۲×۲ به شکل قراردادی تنظیم می‌شود، در حالی که نتیجه استاندارد طلائی در ستون‌ها نمایش داده شده است. انطباق با این قرارداد شما را در به یاد آوردن مجدد بعضی از تعاریف، که بعداً خواهیم آورد، کمک خواهند کرد.

قسمت بالایی سمت چپ نمودار تعداد بیمارانی را که بر طبق استاندارد طلائی دارای بیماری هستند و همچنین کسانی که نتایج آزمایش مثبت دارند، نشان می‌دهد. چنین است که آزمایش به درستی آنها را به عنوان دارندگان بیماری مورد نظر، شناسایی می‌کند. چنین بیمارانی مثبت حقیقی [True Positive (TP)] نامیده می‌شوند. قسمت پایینی سمت راست کسانی را که بیماری ندارند و نتایج آزمایش منفی دارند، نشان می‌دهد که منفی حقیقی [True Negative (TN)] هستند. این دو با هم، مثبت حقیقی و منفی حقیقی یعنی همه بیمارانی را که آزمایش‌های تشخیصی درست دارند، تشکیل می‌دهند. هرچه نسبت بیمارانی بیشتری در این دو گروه یافت شود، آزمایش تشخیصی دقیق‌تر است.

بیمارانی که در دو کادر باقی مانده قرار می‌گیرند، نشان‌دهنده موارد قابل تردید هستند، افراد قسمت بالایی سمت راست، بیمارانی را نشان می‌دهد که مبتلا به بیماری نبوده‌اند. اما نتایج آزمایش آنها به غلط، مثبت است. این بیماران مثبت کاذب (FP) نامیده می‌شوند. هزینه‌های این نوع اشتباه، بررسی‌ها و درمان‌های غیر ضروری است که باید انجام شوند و تأثیرات برچسب اشتباهی که بر بیماران زده می‌شود. قسمت پایینی سمت چپ بیمارانی را نشان می‌دهد که بیماری را دارند، اما اشتبهاً نتایج آزمایش منفی دارند و منفی کاذب (FN) می‌باشند، هزینه این نوع اشتباه مرگ و میر و بیماری همراه با فقدان درمان فوری می‌باشد. ما جدول را با اضافه کردن ردیف‌ها و ستون‌هایی کامل کرده‌ایم:

FN+TN, TP+FP, FP+TN, TP+FN

حساسیت به عنوان نسبت کسانی که بیمارند به کسانی که نتیجه آزمایش مثبت دارند، تعریف شده است این مورد برخی اوقات به صورت مثبت بودن در بیماری (Positivity in Disease) خلاصه شده است. حساسیت معیاری است در این مورد که یک آزمایش

تا چه حد در تشخیص بیماری در صورتی که بیماری وجود نداشته باشد خوب عمل کرده است. این مطلب می‌تواند از ستون اول جدول به وسیله این فرمول محاسبه شود.

$$\text{حساسیت} = \frac{TP}{TP+FN}$$

Target Disease According to Gold Standard

		Present	Absent	
Diagnostic Test Result	Positive (Abnormal)	TP	FP	TP+FP
	Negative (Normal)	FN	TN	FN+TN
		TP+FN	FP+TN	

تصویر ۱

تخصیص به عنوان نسبت کسانی که بیمار نیستند به کسانی که نتایج آزمایش منفی دارند تعریف شده است، بعضی اوقات به آن به عنوان "منفی بودن در سلامت" (Negativity in health) ارجاع می‌شود. تخصیص، معیاری است در این مورد که تا چه حد آزمایش در ثبت نتایج منفی در صورتیکه بیمار مورد نظر وجود نداشته باشد، خوب عمل کرده است. این مساله می‌تواند از ستون دوم جدول به وسیله این فرمول محاسبه شود:

$$\text{ویژگی} = \frac{TN}{FP+TN}$$

اگر واژه احتمالات شرطی را ترجیح می‌دهید، حساسیت احتمال شرطی آزمایش مثبت، به شرط وجود بیماری است و ویژگی احتمال شرطی آزمایش منفی به شرط عدم وجود بیماری است. تعاریف اخیر تنها از آن جهت آورده شده که ممکن است در مقالات با آنها مواجه شوید.

برای تشریح این مفاهیم، از مثالی که در تاریخ اخیر پزشکی به خوبی شناخته شده است یعنی ارزیابی اسید فسفاتاز پروستاتی (PAP) برای تشخیص سرطان پروستات که توسط فوتی (Foti) و همکارانش تنظیم شد، استفاده خواهیم کرد. اگرچه PAP توسط سایر آزمایشات، بخصوص آنتی‌ژن‌های ویژه پروستاتی (PSA) از دور خارج شده است، با این وجود به سبب رویدادهای مربوط به ارزیابی و اشاعه آن، به ویژه آزمایش اسید فسفاتاز پروستاتی (PAP) به عنوان مثالی عالی برای تشریح و مفاهیم این مقاله همچنان مورد استفاده قرار می‌گیرد. فوتی و همکارانش آن را در مورد ۱۱۳ بیمار مبتلا به کارسینومای پروستات، تایید شده توسط استاندارد طلایی بیوپسی جراحی و ۲۱۷ فرد فاقد سرطان پروستات، به کار بردند. گروه دوم مرکب از افراد طبیعی، بیماران مبتلا به هیپرپلازی خوش خیم پروستات، بیمارانی با سابقه پروستاتکتومی یا بیماری‌های گوارشی یا سرطان‌های غیر پروستاتی بودند. ۳۳۰ نفر برای اطلاع از مثبت بودن PAP بررسی شد و نمونه به عنوان مثبت (غیر طبیعی) یا منفی (طبیعی) مشخص شد. نتایج با وضعیت واقعی سرطان پروستات در ۳۳۰ نفر برای ایجاد تصویر ۲ مقایسه شد. از ۱۱۳ بیمار مبتلا به کارسینومای پروستات، آزمایش PAP به عمل آمد که نتیجه در ۷۹ نفر مثبت بود، با حساسیت ۷۹/۱۱۳ یا (۷۰٪) از یک حساسیت ۰/۷ یا ۷۰٪ این نتیجه بدست می‌آید که از ۱۰۰ بیمار مبتلا به کارسینومای پروستات، آزمایش تقریباً ۷۰ نفر از آنها را تشخیص می‌دهد و حدود ۳۰ نفر از آنها را از دست می‌دهد. در تصویر اخیر نتیجه توسط تفریق درصد حساسیت، از عدد ۱۰۰، به عنوان میزان منفی کاذب شناخته می‌شود.

از ۲۱۷ بیماری که کارسینومای پروستات نداشتند، آزمایش در ۲۰۴ نفر از آنها منفی بود این تخصیص ۲۰۴/۲۱۷ یا ۰/۹۴ می‌باشد. یک تخصیص ۰/۹۴ یا ۰/۹۴ یعنی از ۱۰۰ بیمار بدون کارسینومای پروستات، آزمایش در ۹۴ نفر از آنها منفی خواهد بود، و در ۶ نفر از آنها به صورت مثبت کاذب است. آخرین تصویر توسط تفریق درصد تخصیص، از عدد ۱۰۰ به دست آمد و به عنوان میزان مثبت کاذب شناخته شد.

حساسیت و تخصیص، معیارهایی از اعتبار یک آزمایش تشخیصی هستند. هرچه ارزش آنها بالاتر باشد، آزمایش، بهتر وجود یا عدم وجود بیماری را تشخیص می‌دهد. حساسیت بالاتر، میزان منفی کاذب پایین‌تر دارد و شانس از دست دادن موارد بیماری مورد نظر کمتر می‌شود، وقتی که بیماری وجود داشته باشد. ویژگی بالاتر، نتایج مثبت کاذب پائین دارد وقتی که بیماری مورد نظر موجود نباشد. آزمایش PAP، با حساسیت ۷۰٪ و ویژگی ۹۴٪، مثال ویژگی بالا با حساسیت متوسط است.

فایده آزمایش تشخیصی (سودمندی)

برای محاسبه حساسیت تخصیص، باید وجود یا فقدان حقیقی بیماری مورد نظر را بشناسید، این مطلب در ابتدا توسط ۱۱۳ بیماری که قبلاً بیوپسی، وجود کارسینومای پروستات را در آنها اثبات کرده بود، شرح داده شد. در بسیاری از موقعیت‌های بالینی نمی‌دانیم آیا بیماری مورد نظر وجود دارد یا خیر. اگر نتایج یک استاندارد طلایی در دسترس بودند، نیازی به آزمایش‌های تشخیصی دیگری نبود.

Prostatic Carcinoma According Biopsy Results

		Present	Absent	
Prostatic Acid Phosphatase Result	Positive (Abnormal)	79	13	92
	Negative (Normal)	TP	FP	TP+FP
		FN	TN	
		34	204	238
	TP+FN	FP+TN		
	113	217	330	

تصویر ۲

در طبابت بالینی، وقتی که وجود بیماری مورد نظر قطعی نیست، اغلب با مشکلات اضافی مواجه می‌شویم. ابتدا با فرض اینکه بیماری یک نتیجه مثبت داشته باشد، نیاز داریم بدانیم، که تا چه حد احتمال دارد که بیماری مورد نظر وجود داشته باشد و این ارزش پیش‌بینی مثبت [Negative Predictive Value (NPV)] آزمایش است. PPV منعکس می‌کند که چقدر می‌توانیم راجع به وجود بیماری مورد نظر در بیماران با نتیجه آزمایش مثبت مطمئن باشیم. همچنین در بیماران با نتایج منفی، نیاز داریم که احتمال این را که بیمار مورد نظر وجود ندارد بدانیم و این ارزش پیش‌بینی منفی (NPV) نامیده می‌شود. هر قدر PPV بالاتر باشد، آزمایش در تعیین وجود بیماری در صورتیکه نتایج مثبت باشد، بهتر است و هر قدر NPV بالاتر باشد آزمایش در تعیین عدم وجود بیماری در صورتیکه نتیجه منفی باشد بهتر است. درحالیکه حساسیت تخصیص به ترتیب، قابلیت ذاتی آزمایش را برای تشخیص وجود یا عدم وجود بیماری اندازه می‌گیرند، ارزش پیش‌بینی نیز فایده آزمایش را در طبابت بالینی اندازه می‌گیرد.

به بیماری که به صورت تصادفی از یک گروه ۳۳۰ نفری در مطالعه فوتی و همکارانش انتخاب شده بود توجه کنید (تصویر ۲) فرض کنید، نمی‌دانیم آیا فرد به کارسینومای پروستات مبتلا است یا خیر؟ همان گونه که در طبابت بالینی با آن مواجهیم و قصد داشته باشیم که به ارزیابی PAP برای پاسخ، اعتماد کنیم. از هنگامی که ۱۱۳ نفر از ۳۳۰ نفر در مطالعه، کارسینومای پروستات داشتند، می‌دانیم که یک احتمال ۱۱۳/۳۳۰ یا ۳۴/۲٪ وجود دارد که فردی به طور تصادفی انتخاب شده، بیمار باشد این اندازه نسبت بیماران مبتلا به محل جمعیت به عنوان شیوع شناخته می‌شود و با استفاده از این فرمول محاسبه می‌شود:

$$\text{شیوع} = \frac{TP + FN}{TP + FN + FP + TN}$$

صورت (TP + FN) تعداد کل افرادی است که مبتلا به بیماری هستند (یعنی مجموع قسمت‌ها در ستون دست چپ (تصویر ۲)؛ و مخرج تعداد کل افراد شرکت کننده است (یعنی مجموع همه چهار قسمت در تصویر ۲). نظر به این که یک شانس ۳۴/۲ درصدی وجود دارد که بیمار ما، که به شکل تصادفی از نمونه بیرون کشیده شده، قبل از اینکه تحت آزمایش PAP قرار گیرد سرطان پروستات داشته باشد، شیوع همچنین به احتمال قبل از آزمایش (Pretest Probability) یا امکان پیش از آزمایش (Pretest Likelihood) نیز باز می‌گردد.

اکنون ببینیم آزمایش PAP تا چه حد پیش‌بینی وجود یا عدم وجود بیماری را در بیمار، انتخاب شده به صورت تصادفی با فرض بر اینکه آزمایش او مثبت است، می‌تواند به خوبی انجام دهد. از آنجاییکه ۹۲ بیمار با نتایج مثبت آزمایش وجود دارند که ۷۹ نفر از آنها واقعاً سرطان پروستات دارند، یک شانس ۷۹/۹۲ یا ۸۵/۹ درصدی وجود دارد که بیمار ما که آزمایش مثبت دارد، کارسینومای پروستات داشته باشد. این، PPV برای این جمعیت است که عبارت است:

$$PPV = \frac{TP}{TP + FP}$$

از آنجائی که PPV، احتمال اینکه بیمار مبتلا به یک بیماری معین را وقتی که نتیجه آزمایش مثبت است، نشان می‌دهد، همچنین به احتمال بعدی یک آزمایش مثبت یا [Posttest Probability of a Positive Test (PTL)] نیز باز می‌گردد. در این مورد PPV یا (PTL +) به صورت قابل توجهی (۸۵/۹٪) از احتمال قبل از آزمایش ۳۴/۲٪ بالاتر است و آزمایش به خوبی در تعیین وجود کارسینومای پروستات توسط افزایش قابل توجه احتمال وجود بیماری در صورتی که آزمایش مثبت باشد، انجام گرفته است. اکنون فرض کنید نتیجه PAP در فردی که به صورت تصادفی شده است، منفی باشد. از آنجاییکه ۲۳۸ بیمار با نتیجه آزمایش منفی وجود دارند که ۲۰۴ نفر از آنها به بیماری مبتلا نیستند، یک شانس ۲۰۴ / ۲۳۸ یا ۸۵/۷ درصدی وجود دارد که وقتی آزمایش منفی است، بیمار فاقد کارسینومای پروستات باشد این یک ارزش پیش‌بینی منفی (NPV) برای این جمعیت است. (تصویر ۲)، که به این ترتیب محاسبه می‌شود.

$$NPV = \frac{TN}{FN + TN}$$

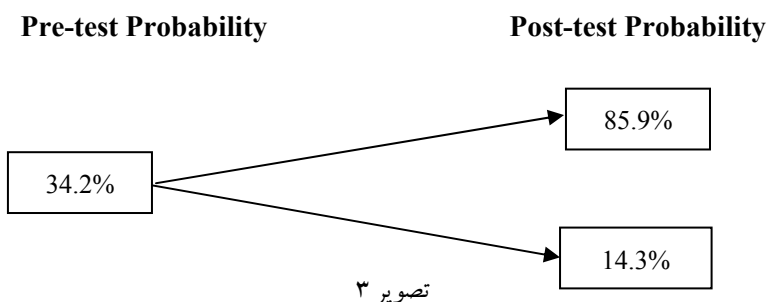
در این مثال، NPV ۸۵/۷٪ احتمال عدم وجود بیماری را زمانی که نتیجه PAP منفی است، نشان می‌دهد. اگر یک شانس ۸۵/۷ درصدی برای عدم وجود بیماری وجود داشته باشد، یک شانس ۱۴/۳ درصدی برای وجود بیماری، نیز وجود دارد، تفریق NPV از ۱۰۰ احتمال وجود بیماری را حتی وقتی که تست منفی است به دست می‌دهد که احتمال بعدی یک آزمایش منفی [Posttest Probability of negative Result (PTL-)] است. همانطور که ممکن است پیش‌بینی کرده باشید، احتمال بودن بیماری بعد از یک آزمایش منفی (۱۴/۳٪) این احتمال پایین‌تر از احتمال قبل از آزمایش بیماری (۳۴/۲٪) است. نتیجتاً، نتیجه منفی آزمایش در جهت تعیین وجود کارسینومای پروستات توسط کاهش احتمال بیماری از ۳۴/۲٪ به ۱۴/۳٪ دخیل بوده است. شما ممکن است به یک عدم تقارن توجه کرده باشید. ارزش پیش‌بینی مثبت (PPV) و احتمال بعدی یک آزمایش مثبت (PTL+) مترادف هستند، اما ارزش پیش‌بینی منفی (NPV) و احتمال بعدی یک آزمایش منفی (PTL-) مترادف نیستند. NPV به احتمال عدم وجود بیماری در کسانی که یک نتیجه منفی دارند، باز می‌گردد، در حالیکه (PTL-)، به احتمال وجود بیماری در کسانی که نتیجه منفی دارند. هرچند این دو کمیت یکسان نیستند، با این وجود ارتباط نزدیکی با یکدیگر دارند، چرا که جمع آنها می‌شود ۱۰۰٪ و یکی از آنها به سادگی می‌تواند نتیجه دیگری باشد.

نتایج آزمایش PAP در این جمعیت ویژه می‌تواند توسط یک نمودار درختی تشخیصی رسم شود (تصویر ۳). یک چنین بازنمایی‌هایی در تحلیل تصمیمات بالینی، و ساختار علمی تصمیمات بالینی، بسیار مفید است. بیمار با احتمال پیش از آزمایش کارسینومای پروستات که ۳۴/۲٪ است، از سمت چپ نمودار وارد می‌شود. آزمایش اگر آزمایش مثبت باشد، این احتمال به احتمال بعدی یک آزمایش مثبت (۸۵/۹٪) افزایش می‌یابد، اگر آزمایش منفی باشد، این احتمال به احتمال بعدی یک آزمایش منفی

(۱۴/۳٪) کاهش می یابد. بنابراین آزمایش مذکور امکان دستیابی به برخی اطلاعات بالقوه مفید در این جمعیت را نشان می دهد. در مورد بیماران با نتایج مثبت، احتمال بیماری به قدر کافی برای این که بیویسی و کاوش ممکن جراحی را توجیه کند، بالا (۸۵/۹٪) است، اما اگر نتایج منفی باشد احتمال بیماری به قدر کافی پائین (۱۴/۳٪) است که چنین بیمارانی با سایر آزمایشات غیر تهاجمی در صورتیکه در دسترس باشد؛ پیگیری شوند.

بر پایه نتایج حساسیت و تخصص، فوتی و همکارانش نتیجه گرفتند که آزمایش رادیوایمونواسی PAP می تواند به عنوان یک آزمایش غربالگری مناسب برای تشخیص اولیه سرطان پروستات، به خدمت گرفته شود. فوتی و گیتز در یک گزارش ضمیمه سرمقاله اظهار داشتند: «یافته ناگوار این بود که ۹۰٪ موارد زمانی تشخیص داده شدند که مناسبتاز داده بودند. مفهوم روشن گزارش ضمیمه این است که غربالگری گروه عمده ای بر مبنای آزمایش خون، به تنهایی می تواند این تجربه اندوهبار را معکوس سازد». جرائد ملی گزارش کردند که یک آزمایش خون جدید برای سرطان پروستات وعده داده می شود، همچنان که پاپ اسمیر برای سرطان گردن رحم انجام می شود. فایده آزمایش به سودمندی و قابلیت آن در تأثیر مثبت بر مراقبت بیماران در یک موقعیت ویژه بالینی باز می گردد. این امر می تواند تنها در رابطه با یک موقعیت ویژه بالینی مورد قضاوت قرار گیرد و تخصص و حساسیت نسبتاً خوب برای تأیید مزیت بالینی یک آزمایش معین، در همه شرایط کفایت نمی کند. اگرچه حساسیت و تخصص کمیت های ذاتی اعتبار یک آزمایش را اندازه می گیرند و می توان فرض کرد که این کمیت ها در موقعیت های مختلف بالینی پایدار بمانند؛ لیکن ارزش های پیش بینی و احتمالات بعدی یک آزمایش می توانند به شدت تغییر کنند. از آنجایی که ارزش های پیش بینی و احتمالات بعدی یک آزمایش به شدت به احتمال پیش از آزمایش (شیوع) وابسته است، تغییرات در شیوع می تواند منجر به تغییرات قابل توجه در ارزش پیش بینی و مزیت بالینی یک آزمایش شود. به عبارت دیگر، به منظور ارزیابی مناسب یک آزمایش تشخیصی برای یک هدف ویژه بالینی باید آن را در رابطه با همان هدف ویژه تحلیل کنید.

تحلیل اصلی ارزیابی PAP در جمعیتی از بیماران که شیوع سرطان پروستات در بین آنها ۳۴/۲٪ بود، به سختی نمایانگر موقعیت غربالگری متداول بود. سه سال بعد از گزارش فوتی و همکارانش، واتسون و تانگ تحلیل خود را از آزمایش PAP به عنوان یک آزمایش غربالگری برای سرطان پروستات ارائه کردند. بر پایه داده های ملی در سال ۱۹۶۴ در ایالات متحده این مؤلفان فرض کردند که شیوع کارسینوما پروستات در بین مردان سفید پوست آمریکایی ۳۵ مورد به ازاء ۱۰۰/۰۰۰ نفر است. آنها سپس از حساسیت و تخصص تثبیت شده آزمایش PAP به ترتیب ۷۰٪ و ۹۴٪ برای محاسبه ارزش پیش بینی در موقعیت غربالگری استفاده کردند. برای انجام این محاسبات، شما با قرار دادن جمعیت فرضی ۱۰۰/۰۰۰ نفری به عنوان مجموع در گوشه سمت راست و چپ پائینی جدول ۲×۲ (TP+FP+FN+TN) شروع می کنید (تصویر ۴A) از آنجایی که ۳۵ مورد تخمین زده شده در بین این گروه فرضی وجود دارد، مجموع داخل ستون پائین دست چپ (TP+FN) باید ۳۵ خوانده شود. سپس به این صورت ادامه می یابد که مجموع داخل ستون راست (FP+TN) باید ۹۹۹۶۵=۱۰۰۰۰-۳۵ باشد (تصویر ۴A). از ۳۵ مردی که سرطان پروستات داشتند، آزمایش در حدود ۷۰٪ موارد مثبت خواهد بود (حساسیت)، این تقریباً ۲۵ مورد است و این کمیت می تواند در کادر TP وارد شود (تصویر ۴B). باقی مانده ۱۰=۲۵-۳۵ می تواند در کادر FN وارد شود. به همچنین از ۹۹۹۶۵ مرد بدون کارسینوما پروستات، آزمایش در حدود ۹۴٪ از مردان منفی خواهد بود (ویژگی). این حدود تقریباً ۹۳۹۶۷ می باشد و این کمیت می تواند در TN وارد می شود (تصویر ۴C). باقی مانده ۵۹۹۸=۹۳۹۶۷-۹۹۹۶۵ در قسمت FP وارد می شود (تصویر ۴C). جدول می تواند توسط اضافه کردن میزان کل به سادگی تکمیل شود (تصویر ۵). اکنون می توانیم ارزش های پیش بینی و احتمالات بعدی یک آزمایش را برای موقعیت غربالگری محاسبه کنیم. PPV یا (PTL+)، TP/(TP+FP) یا ۶.۲۳/۲۵ است که ۴۲٪ می باشد.



		Prostatic Carcinoma						
		Present	Absent					
Prostatic Acid Phosphatase Result	Positive (Abnormal)	<table border="1"> <tr> <td>TP</td> <td>FP</td> </tr> <tr> <td>FN</td> <td>TN</td> </tr> </table>		TP	FP	FN	TN	TP+FP
	TP			FP				
FN	TN							
Negative (Normal)			FN+TN					
		TP+FN	FP+TN					
		35	99.965	100.000				

تصویر ۴A

		Prostatic Carcinoma							
		Present	Absent						
Prostatic Acid Phosphatase Result	Positive (Abnormal)	25	<table border="1"> <tr> <td>TP</td> <td>FP</td> </tr> <tr> <td>FN</td> <td>TN</td> </tr> </table>		TP	FP	FN	TN	TP+FP
	TP	FP							
FN	TN								
Negative (Normal)	10			FN+TN					
		TP+FN	FP+TN						
		35	99.965	100.000					

تصویر ۴B

		Prostatic Carcinoma								
		Present	Absent							
Prostatic Acid Phosphatase Result	Positive (Abnormal)	25	5.998	<table border="1"> <tr> <td>TP</td> <td>FP</td> </tr> <tr> <td>FN</td> <td>TN</td> </tr> </table>		TP	FP	FN	TN	TP+FP
	TP	FP								
FN	TN									
Negative (Normal)	10	93.967		FN+TN						
		TP+FN	FP+TN							
		35	99.965	100.000						

تصویر ۴C

$TN \times NPV / (TN + FN)$ یا $93967/93977$ است، که $99/99\%$ می‌باشد. نتیجتاً، احتمال بعدی یک آزمایش منفی، توسط تفریق NPV از 100% به دست آمد که $0/01\%$ است. احتمال سرطان پروستات قبل از آزمایش (احتمال پیش از آزمایش با شیوع) $1/100000$ یا $0/035\%$ تعیین شد.

در تصویر ۶ یک نمودار درختی تشخیصی، این نتایج را برای موقعیت غربالگری خلاصه می‌کند. یک مرد معمولی و بدون علامت، که توسط چنین آزمایشی غربالگری شده باشد، به آزمایش سمت چپ نمودار با یک احتمال پیش از آزمایش سرطان پروستات $0/035\%$ نزدیک می‌شود (۳۵ شانس در $100/000$ نفر) اگر آزمایش او مثبت باشد احتمال تنها $0/42\%$ افزایش می‌یابد (۱ شانس در 240). حتی با یک نتیجه آزمایش مثبت، شانس داشتن سرطان پروستات هنوز بی‌نهایت اندک است و این امر به سختی می‌تواند آزمایش تهاجمی تری را توجیه کند. وقتی که در هنگام انجام غربالگری نتیجه آزمایشی مثبت باشد. از نظر بالینی اندک خواهد بود. به علاوه از آن جایی که تقریباً 6% ($6023/100,000$) از همه مردان آمریکائی سفید پوست آزمایش غربالگری PAP مثبت داشتند، هر خط مشی بررسی بیشتر موارد مثبت، تعداد زیادی آزمایش غیر ضروری از مردان سالم را ایجاد می‌کند، اگر آزمایش منفی باشد، احتمال بیماری به $0/01\%$ ($10/100000$) کاهش می‌یابد. اگر چه ممکن است کسی بگوید وقتی که آزمایش منفی باشد، در موقعیت غربالگری مفید است، چرا که تقریباً کارسینومای پروستات را غیر متحمل می‌شمارد این کاهش در احتمال از فوائد کمی برخوردار است، زیرا این بیماری در این جمعیت معین بسیار نادر است ($35/100000$).

مثال ما رابطه بین احتمال بعد از آزمایش و احتمال پیش از آزمایش را شرح می‌دهد. در تحلیل تشخیصی که توسط فوتی و همکارانش انجام شد، احتمال بعدی یک آزمایش مثبت و منفی نسبتاً بالا بود (به ترتیب $85/9\%$ و $14/3\%$) زیرا احتمال پیش از آزمایش، قبل از انجام آزمایش، بالا بود ($34/2\%$). در تحلیل غربالگری که توسط واتسون و تانگ انجام شد، احتمالات بعدی یک آزمایش مثبت و منفی بی‌نهایت بودند (به ترتیب $42/0\%$ و $0/01\%$). زیرا احتمال پیش از آزمایش بسیار پائین بود ($0/035\%$). احتمالات پیش و پس از آزمایش می‌تواند توسط یک بیان ریاضی به نام قضیه Baye به هم متصل شوند. در ادامه دو عدد از عبارتهای متساوی قضیه Baye آمده است:

$$PPV = P \times SENS / (P \times SENS) + (1 - P) \times (1 - SPEC)$$

$$NPV = (1 - P) \times SENC / (1 - P) \times SPEC + P \times (1 - SENS)$$

Prostatic Carcinoma

		Present	Absent	
Prostatic Acid Phosphatase Result	Positive (Abnormal)	25	5,998	6,023
	Negative (Normal)	10	93,967	93,977
		TP+FN	FP+TN	
		35	99,965	100,000

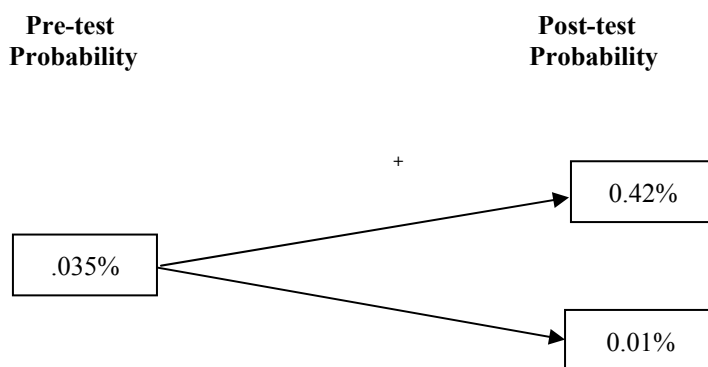
$$PPV = TP / (TP + FP) = 25 / 6,023 = 0.0042 \text{ or } 0.42\%$$

$$NPV = TN / (TN + FN) = 93,967 / 93,977 = 0.9999 \text{ or } 99.99\%$$

در حالی که P, SENS, SPEC به ترتیب احتمال قبل از آزمایش، حساسیت و ویژگی در نسبت اعشاری (مثلاً ۰/۸۰) به جای درصد (مثلاً ۸۰٪) را نشان می‌دهد. از بیان‌های متفاوتی برای قضیه Baye استفاده می‌شود که شکل متداول توصیف احتمالات پس از آزمایش یا ارزش پیش‌بینی با واژه‌هایی مانند احتمال پیش از آزمایش است. همه نشان می‌دهند که چگونه ارزش‌های قبلی وابسته به ارزش بعدی هستند و یک روش برای محاسبه احتمالات بعد از آزمایش و ارزش‌های پیش‌بینی در یک آزمایش تشخیصی معین و یک احتمال پیش از آزمایش معین را فراهم می‌کنند. شما به محاسبه ارزش‌های پیش‌بینی برای آزمایش PAP در موقعیت غربالگری توسط قرار دادن ۰/۰۰۰۳۵ به جای P، ۰/۷۰ به جای SENS و ۰/۹۴ به جای SPEC در فرمول بالا، ترغیب می‌شوید. باید ارزش‌های پیش‌بینی را که قبلاً به دست آمده، استنتاج کنید مگر در صورتی که تفاوت‌های احتمالی، مختصری ناشی از اشتباه باشد. این فرمول‌ها یک جایگزین جالب توجه به جای مجموعه‌های محاسبات و جداول را که برای دستیابی به ارزش‌های پیش‌بینی و احتمالات بعد از آزمایش به کار بردیم، فراهم می‌کنند (تصاویر ۴C، ۴A، و ۵). قضیه Baye در هر کدام از این اشکال، روش نسبتاً ساده‌ای برای استخراج ارزش‌های پیش‌بینی و احتمالات بعد از آزمایش، که مرکز بررسی مزیت بالینی هستند، می‌باشند. از این نوع از تحلیل گاهی اوقات به عنوان «تحلیل Bayesian» یاد می‌شود.

رادبویایمونواسی PAP چندین مورد بنیادی را درباره تست‌های تشخیصی نشان می‌دهد:

۱. حساسیت و ویژگی، اعتبار (دقت) یک آزمایش تشخیصی را اندازه می‌گیرند و ارتباط مستقیمی با مزیت بالینی آزمایش ندارند.
۲. مزیت بالینی یک آزمایش بهتر است با ارزیابی ارزش‌های پیش‌بینی و احتمالات بعد از آزمایش در یک موقعیت بالینی ویژه مورد بررسی قرار گیرد.
۳. ارزش‌های پیش‌بینی و احتمالات بعد از آزمایشی وابستگی بسیار شدیدی به احتمال پیش از آزمایش (شیوع) بیماری دارند. و به عنوان نتیجه، موقعیت بالینی و جمعیت بیمار در چنین آزمایشی محسوب می‌شوند. این به میزان زیادی از تصور غلط مردم مبنی بر این که نتایج آزمایش قطعی هستند، (به عبارت دیگر آزمایشات مثبت دلالت بر این دارند که افراد بیمار هستند و آزمایشات منفی بر این که آنها بیمار نیستند) متفاوت است و این که استنتاجاتی که از نتایج آزمایش ترسیم شده‌اند، مستقل از بیمارانی هستند که آزمایش شده‌اند.



تصویر ۶

ارزیابی آزمایش‌های تشخیصی

آزمایش‌های تشخیصی و فن‌آوری جدید دائماً در پیشرفتی رو به افزایش هستند. برای مثال، آزمایش آنتی‌ژن ویژه پروستاتی (PSA)، توموگرافی خروج پوزیترون (PET)، و تصویربرداری مقاومت مغناطیسی (MRI) را داریم. این آزمایش‌ها و تکنیک‌های

جدید باید قبل از این که به چرخه مصرف بالینی وارد شوند، مورد ارزیابی قرار گیرند و بایگانی متون پزشکی باید دربرگیرنده مقالات بیشتری در مورد ارزیابی آنها باشد.

رهنمودهایی برای ارزیابی آزمایش‌ها و تکنیک‌های جدید

هدف از آزمایش تشخیصی

هر ارزیابی مناسبی از یک آزمایش تشخیصی با عبارتی روشن از هدف بالینی پیشنهاد شده آن، آغاز می‌شود. آزمایش رادیوایمونواسی اسید فسفاتاز پروستاتی، نشان می‌دهد که چگونه یک آزمایش بالینی ممکن است به خوبی در یک موقعیت (تشخیص) انجام پذیرد، اما در موقعیت دیگری (غربالگری) با شکست مواجه شود. عملکرد بالینی پیشنهاد شده، چگونگی انجام ارزیابی را دیکته می‌کند.

آزمایشات بالینی در پنج عملکرد بالینی متفاوت به خدمت گرفته می‌شوند:

تشخیص، غربالگری، مرحله‌بندی، مراقبت و طبقه‌بندی. تشخیص (Diagnosis) حکم به ورود یا خروج از یک بیماری، در بیمار است که بیماری در او به صورت بالینی مشکوک است (مانند آنژیوگرافی برای تشخیص بیماری شراین کرونری در یک بیمار مبتلا به آنژین). غربالگری (Screening) یعنی کشف احتمالی یک بیماری معین، در یک گروه از افرادی که بدون علامتند و مشکوک به داشتن بیماری نیستند، (مثلاً ماموگرافی برای کشف احتمالی سرطان پستان در خانم‌های میانسال). مرحله‌بندی (Staging) استفاده از یک آزمایش بالینی، به عنوان راهنمای درمان برای اندازه‌گیری میزان پیشرفت بیماری است (مانند دیدن مדיاستن "Mediastinoscopy" برای تعیین قابلیت جراحی از سرطان ریه). نظارت (Monitoring) استفاده از یک آزمایش، برای ارزیابی و تعدیل درمان مداوم است (مانند زمان‌های پروترومبین برای نظارت تأثیر درمان با ضد انعقادها). تقدم به استفاده از یک آزمایش برای تعیین این که چه بیمارانی باید تحت آزمایش‌های تهاجمی تری قرار گیرند، برمی‌گردد (مانند مطالعات داپلر برای تعیین این که کدام یک از بیماران باید آنژیوگرافی مغزی "Cerebral" شوند). بر این اساس شما باید نوع بیمار، جایگاه بالینی خواسته شده و هدف از یک آزمایش را قبل از آن که هرگونه ارزیابی از آن را شروع نمایید، بیان کنید. همچنان که ادامه می‌دهیم اهمیت این نیازها بیشتر بر ما آشکار خواهد شد.

طیف و تعداد بیماران

هنگامی که بیمارانی را که در ارزیابی یک آزمایش تشخیصی شرکت می‌کنند، انتخاب می‌کنیم، قانون زیر را مشاهده می‌کنیم: نوع بیمار مورد بررسی در ارزیابی یک آزمایش، باید منطبق شود با نوع بیماری که آزمایش برای او در طبابت بالینی درخواست شده است. متأسفانه این قانون همیشه دنبال نمی‌شود. برای انجام یک ارزیابی، شما به یک جمعیت با بیماری مورد نظر برای این که حساسیت آن را تخمین بزنید و به یک جمعیت عاری از بیماری مورد نظر برای این که ویژگی آن را تخمین بزنید، نیاز دارید. ممکن است وسوسه شوید که گروهی با بیماری تثبیت شده که قبلاً شناسایی شده‌اند را به عنوان گروه بیمار و گروهی که از نظر سلامتی کنترل شده‌اند را به عنوان گروه غیر بیمار مورد مطالعه قرار دهید. چند لحظه تفکر باید شما را متقاعد کند که چنین گروه‌های غیرمتجانسی در یک آزمایش بالینی با طیف گسترده بیمارانی که شما و سایرین در تجربیات بالینی روزمره با آنها مواجه هستید، قابل رقابت نیستند. یک آزمایش پیشنهاد شده برای تشخیص باید در میدان گسترده‌ای از بیماران مشکوک به داشتن بیماری مورد نظر و یا سایر بیماری‌هایی که اغلب با بیماری مورد نظر اشتباه می‌شوند، ارزیابی شود. یک آزمایش غربالگری باید با افراد بدون علامت شبیه کسانی که در تجربه بالینی با آنها مواجه هستید، مورد رقابت قرار گیرند. استفاده از افرادی که به طور مشخص بیمار و یا سلامتند به عنوان، موردها و شاهد‌ها، به صورت جعلی موجب افزایش ارزش‌های حساسیت و ویژگی می‌شود؛ زیرا این افراد احتمالاً، به ترتیب آزمایش‌های مثبت و منفی خواهند داشت. اگر با بیان عملکرد پیشنهاد شده آزمایش را آغاز کنید، نوع بیمار پیشنهاد شده‌ای که قرار است آزمایش در مورد او انجام شود و جایگاه پیشنهاد برای استفاده از آن، صرفاً نیاز به جمع‌آوری همه بیمارانی دارید که متوالیاً با شرح حال ویژه منطبق باشند. بیمارانی که در جایگاه‌های مناسب در طی یک دوره زمانی کافی دیده شوند. استفاده از بیماران متوالی، احتمال پیشداوری در انتخاب بیماران را به حداقل می‌رساند و طیف بیمارانی که که آزمایش در مورد آنها انجام خواهد شد، برای استفاده بالینی مشروط آن تکرار می‌کنند. اگر می‌خواهید

یک آزمایش تشخیصی جدید برای DVT جهت استفاده در مورد بیمارانی که برای اولین بار خودشان را به سطح مراقبت‌های اولیه معرفی می‌کنند ارزیابی کنید، به هر بیماری که به طور منظم با پزشک خانوادگی مشورت می‌کند و یا با ساق پای متورم و مشکوک به DVT در یک دوره زمانی به یک اورژانس محلی ویژه وارد می‌شود، توجه کنید. چنین گروهی شامل بیمارانی خواهد بود که به درجات متنوعی از شدت به DVT مبتلا هستند و همچنین برخی افرادی که اختلالات مبهمی مانند کیست بیکر پاره شده یا ترومبوفلیت‌های سطحی دارند. انجام آزمایش در چنین گروهی از بیماران دقیقاً انجام آن را در گروه‌های مشابهی در جایگاه مراقبت اولیه مشابه، منعکس می‌کند. اگر قصد دارید که یک آزمایش غربالگری، برای کشف زودرس کارسینومای پروستات را ارزیابی کنید، باید بیان کنید که آزمایش برای استفاده در مردان مسن که هم بدون علامت هستند و هم از نظر بالینی طبیعی، پیشنهاد می‌شود و سپس آن را در یک جمعیت بزرگ از چنین افرادی، ارزیابی کنید.

غالباً تعداد غیرکافی از بیماران مورد مطالعه قرار می‌گیرند، یک محقق ممکن است یک آزمون را در بین ۲۵ فرد بیمار ارزیابی کند و دریابد که ۲۰ نفر از آنها نتایج مثبت دارند و استنتاج کند که حساسیت ۸۰٪ است. اگر چه این از لحاظ تکنیکی بهترین تخمین برای حساسیت بر مبنای داده‌ها است، اطمینان ۹۵٪ محدود به پیرامون آن می‌باشد که بسیار گسترده است (۶۴٪ تا ۹۵٪). محقق باید یک رویکرد محافظه‌کارانه داشته باشد و از حد پائین (۶۴٪) به عنوان تخمین حساسیت استفاده کند. اما در بیشتر موارد این گونه نیست، حساسیت به عنوان این که ۸۰٪ بوده است، بیان می‌شود. نتیجه یک تخمین بالاتر از اعتبار آزمایش است، تنها استثناء قابل قبول استفاده از تعداد کافی بیماران برای اطمینان از تخمین دقیق زمانی رخ می‌دهد که، یک آزمایش برای یک بیماری مورد نظر نادر در حال ارزیابی بوده و تعداد بیماران در دسترس به صورت غیرقابل اجتنابی کم باشند.

استاندارد طلایی

استاندارد طلایی مجموعه‌ای از معیارهاست، که برای تعیین این که کدام فرد واقعاً بیمار و کدام یکی بیمار نیست توسط محققین بکار می‌رود. این معیارها تأثیری قاطع بر جدول ۲×۲ و بر تعیین حساسیت و ویژگی یک آزمایش دارند. استاندارد طلایی ممکن است، قطعی باشد (مانند نتایج هیستوپاتولوژیکی، بیوپسی، جراحی یا اتوپسی). ممکن است به سادگی نتایج سایر آزمون‌های تشخیصی پذیرفته شده رایج به عنوان استانداردهایی برای تشخیص بیماری مورد نظر را مورد سؤال قرار دهد. در یک موقعیت معین، استاندارد طلایی ممکن است مجموعه‌ای از علائم و نشانه‌ها و نتایج آزمایش باشد (مانند سیستم‌ها دسته‌بندی برای تشخیص آرتريت روماتوئید، لوپوس اریتماتوسیسستمیک و تب روماتیسمی). در ارزیابی آزمایش تشخیصی، این مطلب اهمیت حیاتی دارد که استاندارد طلایی به خوبی تعریف شده، قابل تکرار بوده و عنوان یک استاندارد بالینی رایج برای تشخیص بیماری مورد نظر، پذیرفته شده باشد.

هر فردی که گزارشی از ارزیابی یک آزمایش تشخیصی را بخواند، می‌خواهد بداند که آزمایش چگونه برای کشف وجود یا عدم وجود بیماری مورد نظر در رابطه با استاندارد طلایی، عمل خواهد کرد. اگر استاندارد طلایی به خوبی تعریف نشده باشد یا استاندارد رایج را برای تشخیص بیماری هدف، ارائه نکرده باشد، ارزیابی آن از استفاده کمی برخوردار است زیرا در عملکرد بالینی مستقیماً قابل اجرا نخواهد بود. برای مثال، از آن جایی که نتایج اتوپسی یا یافته‌های جراحی برای بیماران مبتلا به بیماری شراین کرونر (CAD) معمولاً قابل دسترسی نیستند، نتایج آنژیوگرافی در سطح گسترده‌ای به عنوان استاندارد بالینی برای تشخیص، مورد پذیرش قرار گرفته‌اند. نتیجتاً اگر بخواهید که انجام یک آزمایش جدید در تشخیص CAD را ارزیابی کنید، ممکن است به خوبی از آنژیوگرافی کرونر به عنوان استاندارد طلایی خارجی خود برای بررسی وجود یا عدم وجود CAD در هر بیمار، استفاده کنید؛ معیارهای آنژیوگرافی که برای تأیید وجود CAD به کار می‌برید باید به وضوح در گزارش ارزیابی شما بیان شوند. چنین بیانات دقیقی باید شامل چنین معیارهای روشن و قابل تکراری باشند. مثل «یک تنگی بیشتر از ۷۵٪ که توسط بازنگری مستقل دو متخصص قلب مشاهده شده است» و از ابهامی مانند هر موقعیت غیرعادی که توسط آنژیوگرافی مشاهده شده «باید اجتناب کرد. روش‌های استفاده شده برای انجام استاندارد طلایی را مو به مو شرح دهید، آنگونه که خواننده بتواند تعیین کند که استاندارد طلایی، به آن صورت که در مطالعه شما استفاده شده، با استفاده آن در جایگاه بالینی آنها مشابه می‌باشد یا خیر؟

آزمایش تشخیصی

شرح دقیق جزئیات، همان طور که برای استاندارد پلائی تعیین شد، به صورت مشابهی برای آزمایش تشخیصی که در حال ارزیابی آن هستید، لازم است. به این ترتیب که باید به وضوح روش شناسی آزمایش، شرایطی که آزمایش تحت آن انجام شد و این که چگونه بیماران مهیا شدند را شرح دهید، جزئیات کافی را ارائه کنید تا خوانندگان شما بتوانند آزمایش تشخیصی را عیناً همان طور که شما آن را در ارزیابی خود انجام داده‌اید، به انجام برسانند.

هنگامی که آزمایش یک نتیجه کمی را به بار می‌آورد (مانند تمرکز یک جزء اصلی سرم) باید دقت آن را ارزیابی کنید. تغییرپذیری تولید شده توسط تکنولوژی آزمایش می‌تواند با مقایسه نتایج برای چندین نمونه گرفته شده از بیماران مشابه در زمان مشابه ارزیابی شود. یا در همان بیمار در زمان‌های متفاوت ارزیابی شود. حساسیت، ویژگی و ارزش پیش‌بینی نباید در چندین ارزش آستانه محاسبه شود (مثلاً ارزشی که نتایج طبیعی و غیرطبیعی را از هم جدا می‌کند). هنگامی که یک ارزش آستانه منفرد انتخاب می‌کنید، انتخاب شما باید تأیید شود.

اگر آزمایش یک نتیجه کیفی را به بار آورد که نیاز به تفسیر داشته باشد (مانند سی تی اسکن‌هایی که باید توسط رادیولوژیست تفسیر شوند)، معیارهای تفسیر را به صورتی قابل تکرار و به روشنی بیان کنید. به علاوه، تغییرپذیری درون و بین مشاهده‌گر در اجزاء تفسیری آزمایش را با ارائه همان فهرست متوالی از نتایج آزمایش به دو یا چند مشاهده‌گر برای بازنگری مستقل و به یک مشاهده‌کننده منفرد برای بازنگری مستقل در دو یا چندین موقعیت جداگانه، ارزیابی کنید.

استقلال

زمانی که بیماری در نمونه شما تحت انجام آزمایشات تشخیصی و استاندارد پلائی قرار می‌گیرد نتایج برای تعیین اعتبار و مزیت آزمایش در جدول ۲×۲ (تصویر ۲) مقایسه می‌شود. آزمایش تشخیصی و آزمایش استاندارد پلائی باید از یکدیگر مستقل باشند. هر دو آزمایش باید در مورد هر بیمار انجام شود. نتیجه آزمایش تشخیصی نباید در مورد انتخاب بیمار برای انجام آزمایش استاندارد پلائی، تأثیر بگذارد. این روند باید در صورت امکان کوری سه جانبه (Triple-Blinded) باشد. به این ترتیب که، کسانی که آزمایش را انجام می‌دهند، کسانی که آن را تفسیر می‌کنند و کسانی که مورد آزمایش قرار می‌گیرند (بیماران) تا انجام و تفسیر هر دو آزمایش باید کور (بی‌اطلاع) نگهداشته شوند. همچنین افراد گوناگونی که آزمایش تشخیصی را انجام می‌دهند، تفسیر می‌کنند و مورد آزمایش قرار می‌گیرند، باید از نتایج آزمایش استاندارد پلائی در صورت امکان، بی‌اطلاع باشند. عکس آن نیز صادق است، افراد گوناگونی که آزمایش استاندارد پلائی را انجام می‌دهند، تفسیر می‌کنند، و مورد آزمایش قرار می‌گیرند در صورت امکان باید از نتایج آزمایش تشخیصی بی‌اطلاع باشند. اگر این استقلال رعایت نشود، تورش مورد انتظار می‌تواند رخ دهد و منجر به افزایش نادرست حساسیت، ویژگی و ارزش‌های پیش‌بینی شود. این امر به ویژه زمانی صادق است که یک بیمار بسیار ذهن‌گرا و تفسیر کننده، به عنوان یک طرف تشکیل‌دهنده در آزمایش تشخیصی وجود داشته باشد.

ارزیابی و اعتبار و مزیت

هنگامی که یک جدول ۲×۲ تهیه کرده، حساسیت و ویژگی و ارزش‌های پیش‌بینی را استنتاج کردید و دریافتید که آزمایش مفید است، اگر آزمایش‌ها به صورت معقولی بالا باشند، ممکن است تصور کنید کار شما پایان یافته است. یک ارزیابی مناسب، به هر ترتیب، با بررسی پیش‌بینی، احتمالات بعد از آزمایش و اهمیت ویژگی آنها در موقعیت بالینی پیشنهاد شده برای آزمایش پیش می‌رود. از آن جایی که این ارزش‌ها بستگی نزدیکی به شیوع بیماری مورد نظر، در نمونه بیمار شما، براساس تعیین استاندارد پلائی دارند، باید ارزیابی کنید که آیا این شیوع به صورت معقولی به شیوع واقعی همان بیماری در جمعیت مورد نظر پیشنهاد شده، نزدیک است یا خیر؟ اگر از اصولی پیروی کنید که براساس آن، استفاده از کلیه بیماران متوالی که شبیه به آنهایی که مورد آزمایش قرار می‌گیرند و آنهایی که به یک یا چندین جایگاه مشابه با محل انجام آزمایش، وارد می‌شوند، هستند، دقیقاً وضع بدین منوال خواهد بود. اگر شیوع در نمونه بیمار شما یک تخمین واقع‌گرایانه برای جمعیت مورد نظر نباشد، ارزش‌های پیش‌بینی و احتمالات بعد از آزمایش را بر پایه شیوع واقع‌بینانه‌تری، مجدداً محاسبه کنید. می‌توانید این

کار را به سادگی با استفاده از قضیه Baye با مقدار شیوع جدید و حساسیت به دست آمده و ویژگی، انجام دهید. ارزیابی شما از مزیت بالینی باید بر این اساس که آیا آزمایش اطلاعات کافی و مبتنی بر ارزش‌های پیش‌بینی و احتمالات بعد از آزمایش تولید می‌کند، تا موجب ایجاد تغییراتی در کنترل مدیریت شود، متمرکز گردد. آزمایش رادیوایمونواسی برای PAP در موقعیت غربالگری شکست خورد، زیرا احتمالات بعدی در آزمایش مثبت و منفی به قدر کافی با احتمال پیش از آزمایش (شیوع) متفاوت نبودند تا بر هر یک از تصمیم‌های مدیریت تأثیر بگذارند. یک ارزیابی مناسب، بر پایه شیوع واقع‌گرایانه، می‌تواند این مطلب را در تحویل اولیه قبل از استفاده از PAP به عنوان پیش‌بینی آزمایش غربالگری پیشنهاد شده، تشریح کند.

زمانی که مزیت بالینی و تأثیر ارزش‌های پیش‌بینی و احتمال بعد از آزمایش را بر تصمیمات بالینی ارزیابی می‌کنید، باید نتایج را برای مثبت کاذب و منفی کاذب در نظر بگیرید. اگر بیماری مورد نظر غالباً مهلک باشد و درمانی وجود داشته باشد که به صورت قابل توجهی پیامد را تغییر دهد (مانند خونریزی داخل شکمی، هماتوم ساب‌دورال و مننژیت باکتریال) منفی کاذب بی‌نهایت ناخوشایند خواهد بود، در این صورت حساسیت بالا مورد نیاز است. مقدار حساسیت بالا بر کاهش قابل توجه احتمال بعدی در یک آزمایش منفی تأثیر دارد (یعنی ارزش پیش‌بینی منفی بالا رفته است) و این اطمینان منطقی را ایجاد می‌کند که زمانی که آزمایش منفی است، بیماری مورد نظر وجود ندارد. اگر در حال غربالگری برای یک بیماری که به فوریت کشنده نیست باشید، در صورتی که خطا کنید به جهت در دسترس نبودن درمان مؤثر و یا بالا بودن هزینه‌های طبقه‌بندی و تحقیقات بعدی در موارد مثبت (مانند سیستمیک فیبروزیس در نوزادان) مایلید که تعداد موارد مثبت کاذب را در حداقل نگه دارید و در این حالت ویژگی بالائی مورد نیاز خواهد بود. ویژگی بالا بر افزایش قابل توجه احتمال بعدی یک آزمایش مثبت، تأثیر دارد (یعنی ارزش پیش‌بینی مثبت بالا رفته است). و این اطمینان منطقی را ایجاد می‌کند که هنگامی که آزمایش مثبت است، بیماری مورد نظر وجود دارد. عموماً، تأثیرات احتمالات بعد از آزمایش و ارزش‌های پیش‌بینی بر تصمیمات بالینی را بدون مراجعه به مدیریت بعدی نتایج موارد مثبت و منفی، نمی‌توان ارزیابی کرد.

تفسیر

واژه‌های حساسیت و ویژگی غالباً در متون پزشکی مشاهده می‌شوند، و نیاز به درکی قوی برای سهولت قضاوت هوشمندانه نسبت به متون دارند. مفاهیم حساسیت و ویژگی به دقت تعریف شده و سپس به خوبی با مثال جراحی نشان داده شده‌اند، پس از آن سپس به سمت استفاده از ارزش‌های پیش‌بینی منفی و مثبت و این که چگونه می‌توانیم این اطلاعات را در ارزیابی اعتبار آزمون‌های آزمایشگاهی و تصویربرداری مورد استفاده قرار دهیم، هدایت می‌شویم. به دقت به این امر توجه شده است که در حالی که حساسیت و ویژگی ممکن است اعتبار یک آزمایش تشخیصی را اندازه‌گیری کنند، ممکن است مستقیماً به مزیت بالینی آن آزمایش مربوط نباشد. بهترین ارزیابی مزیت بالینی آزمودن ارزش‌های پیش‌بینی است.