

VALIDAÇÃO DE TESTES DE DIAGNÓSTICOS

"As aparências para a mente são de quatro tipos.

As coisas ou são o que parecem ser;

ou não são, e nem parecem ser;

ou são, e não parecem ser;

ou não são, mesmo assim parecem ser.

Identificar corretamente todos estes casos

é a tarefa do homem sábio."

Epictetus - Século II D.C.

I - A EPIDEMIOLOGIA E A CLÍNICA

A medicina clínica e a epidemiologia clínica começaram juntas. Os fundadores da epidemiologia foram, na maior parte, clínicos. Foi somente durante as últimas décadas que as duas se afastaram com escolas, treinamento, jornais e oportunidades de emprego separadas.

Nos últimos anos, clínicos e epidemiologistas vêm se conscientizando de que seus campos se interrelacionam, e a epidemiologia clínica (para alguns: epidemiologia aplicada à clínica) começou a se desenvolver em reconhecimento ao seguinte:

- na maioria das situações clínicas, o diagnóstico, prognóstico e resultado do tratamento são incertos para pacientes individuais e, portanto, devem ser expressos como uma probabilidade.
- a probabilidade para um paciente individual é melhor estimada recorrendo-se à experiência passada com grupos de pacientes similares.
- por serem as observações clínicas feitas em pessoas livres para fazer a sua própria vontade e por clínicos com habilidades e preconceitos variáveis, elas podem ser influenciadas por uma variedade de erros sistemáticos que podem distorcer a natureza real do evento e levar a conclusões erradas.
- para lidar com estes efeitos tendenciosos as observações clínicas devem ser baseadas em princípios científicos sólidos, incluindo modos de reduzir vícios e estimar o papel do acaso.
- estes princípios são importantes tanto para os médicos que desejam ser auto-suficientes no julgamento da informação clínica como para os pesquisadores que produzirão a pesquisa.

Sendo assim, o propósito básico da epidemiologia clínica é promover métodos de observação e interpretação clínica que levem a conclusões válidas. Esta atividade é fundamentada em princípios científicos básicos.

II - INTRODUÇÃO AOS TESTES DIAGNÓSTICOS

Os médicos dedicam grande parte de seu tempo diagnosticando a partir de queixas ou anormalidades de seus pacientes, chegando aos diagnósticos, depois de aplicar vários testes. No entanto, poucos desses médicos recebem um treinamento formal para a interpretação dos testes diagnósticos. A maioria dos bons clínicos usa juízo crítico, conhecimento amplo da literatura e uma abordagem rápida e informal de como organizar os dados. Entretanto, também existem alguns princípios básicos com os quais se precisa estar familiarizado ao interpretar os testes diagnósticos.

Um "teste diagnóstico" geralmente é concebido como um exame realizado em laboratório. Mas os princípios também se aplicam à informação clínica obtida da história, exame físico ou Raio X. Podem ser aplicados também onde um conjunto de achados serve como teste diagnóstico. Assim, pode-se falar do valor da artrite, cardite e coréia no diagnóstico da febre reumática, ou da hemoptise e da perda de peso em um fumante, como indicadores de câncer do pulmão.

É praxe geral em toda pesquisa que o investigador inicie o seu trabalho, desde a idéia inicial quando por ocasião da preparação do projeto, realização do estudo e por fim suas conclusões finais, que a precisão alcançada nos resultados seja sempre a melhor possível.

Quando um profissional de saúde deseja realizar um diagnóstico e precisa apresentar seu resultado considerando sentenças do tipo "normal-anormal", "positivo-negativo", "reator-não reator", "imune-não imune", deve ter em mente que a partir daí deverá se conscientizar de que existirão as mais variadas formas de sentenciar sobre isto, porém deve considerar que o seu objetivo principal terá que ser a precisão dessas sentenças, ou seja, o nível de acerto

deve ser o mais elevado, registrando positivo quando positivo e negativo quando negativo. Portanto um diagnóstico do tipo "positivo-negativo", quanto a classificação dos indivíduos, deve primar por apresentar uma precisão bastante acentuada.

O intuito neste tipo de estudo é mostrar que há interesse em identificar um teste ou meio de diagnóstico dentre tantos, que apresente um grau de precisão considerável para o diagnóstico de alguma doença.

Os clínicos têm consciência de que estabelecer um diagnóstico é um processo complexo, pois o interesse maior é precisar quando positivo ou negativo realmente.

Sabemos que há várias formas de diagnosticar as mais variadas doenças, porém devemos levar em conta que possuímos limitações para tal.

Alguns meios de diagnóstico, mais sofisticados não são acessíveis a todos os locais e profissionais. Portanto, devemos considerar que o objetivo maior será avaliar um meio que realmente contemple resultados satisfatórios.

Dependendo do que queremos investigar já haverá meios diagnósticos tradicionais conhecidos, que nos dão com bastante precisão o resultado "positivo" ou "negativo".

Tanto na clínica como em saúde comunitária, devemos nos conformar com métodos diagnósticos disponíveis e aceitáveis para o indivíduo. Estes métodos correm o perigo de serem menos eficazes que os testes de referência, tais como um diagnóstico apoiado em uma biópsia, uma citologia ou um achado anatomopatológico.

Tanto em indivíduos sadios como em enfermos, um método pode proporcionar resultados verdadeiros ou falsos, verificáveis a partir de um teste de referência.

Procura-se então obter informações sobre a conveniência ou vantagens de recomendar ou utilizar um meio de diagnóstico que seja tão bom quanto um já utilizado e considerado como teste ouro ou teste padrão.

Muitas vezes é difícil encontrar um teste ouro ou padrão, que sirva como ponto de referência para contrastar o novo método.

Devemos ressaltar que o objetivo maior será sempre utilizar um meio de diagnóstico (substituto), que proporcione resultados precisos, quando considerados positivos ou negativos.

No Manual Didático nº 9 - Risco Relativo e Testes Diagnósticos de autoria do Prof. Pedro Carvalho Rodrigues, editado em 1999 pelo INCa-MS encontramos as seguintes informações e exemplos envolvendo os Testes Diagnósticos:

CONCEITOS

Sensibilidade é a capacidade de um teste diagnóstico identificar os verdadeiros positivos nos indivíduos verdadeiramente

doentes. Quando um teste é sensível raramente deixa de encontrar pessoas com a doença.

Escolhemos um teste sensível quando:

- 1 - a doença é grave e não pode passar despercebida
- 2 - a doença é tratável
- 3 - os resultados errados (falsos) não determinam nenhum traumatismo psicológico, econômico ou social para o indivíduo.

Especificidade é a capacidade de um teste diagnóstico identificar os verdadeiros negativos nos indivíduos verdadeiramente sadios. Quando um teste é específico raramente cometerá o erro de dizer que pessoas sadias são doentes.

Utilizamos um teste mais específico possível quando:

- 1 - a doença é importante, mas difícil de tratar ou incurável
- 2 - o fato de saber que não se tem a doença tem importância sanitária e psicológica
- 3 - os resultados errados (falsos) positivos podem provocar traumas psicológicos, econômicos ou sociais.

Valor Preditivo (+) é a proporção de indivíduos verdadeiramente positivos em relação aos diagnosticados positivos apresentados pelo teste.

Valor Preditivo (-) é a proporção de indivíduos verdadeiramente negativos em relação aos diagnósticos negativos realizados pelo teste.

Os valores preditivos de um teste são variáveis. Dependem da prevalência da doença na população (a sensibilidade e a especificidade não se afetam). Se estudamos uma doença cuja prevalência é baixa, mesmo com um teste muito específico, obteremos muitos resultados falsos positivos, devido ao elevado número de indivíduos sadios na coletividade. Se a prevalência é alta, podemos esperar um maior número de resultados falsos negativos, na aplicação de teste com boa sensibilidade.

Portanto, quanto menor for a prevalência da doença, menor será o valor preditivo do resultado positivo e maior o resultado preditivo do resultado negativo. E, ao contrário, no caso da prevalência ser alta.

Acurácia é a proporção de acertos, ou seja, o total de verdadeiramente positivos e verdadeiramente negativos, em relação a amostra estudada.

Desejaremos uma elevada acurácia do teste quando:

- 1 - a doença é importante, mas curável
- 2 - há possibilidade de consequências graves na identificação de falsos positivos e falsos negativos

Na Tabela 1, apresentamos a relação entre o resultado de um teste diagnóstico e a ocorrência da doença.

Tabela 1 - Relação do teste diagnóstico e avaliação da doença (positivo-negativo)

TESTE	DOENÇA		
	PRESENTE	AUSENTE	TOTAL
POSITIVO	a (VP)	b (FP)	a + b
NEGATIVO	c (FN)	d (VN)	c + d
TOTAL	a + c	b + d	a + b + c + d = N

a- verdadeiros positivos

b- falsos positivos

c- falsos negativos

d- verdadeiros negativos

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	$VP / (VP+FN)$
Especificidade	=	$d / (b+d)$	=	$VN / (VN+FP)$
Valor Preditivo (+)	=	$a / (a+b)$	=	$VP / (VP+FP)$
Valor Preditivo (-)	=	$d / (c+d)$	=	$VN / (VN+FN)$
Acurácia	=	$(a+d) / N$	=	$(VP+VN) / N$

APLICAÇÕES

Na avaliação feita por exame clínico, 102 foram considerados positivos e 48 como negativos.

Exemplo 1: Uma amostra de 150 indivíduos foi avaliada quanto ao diagnóstico de faringite através da cultura, sendo que 38 foram considerados positivos e 112 negativos (Tabela 2).

Ao considerar a utilização de um meio de cultura como teste padrão ouro, qual a sua avaliação quanto ao exame clínico? Ele pode substituir o meio de cultura?

Tabela 2 - Diagnóstico de Faringite por Meio de Cultura e Exame Clínico

		MEIO DE CULTURA		
		Positivo	Negativo	Total
Exame	Positivo	8 (a)	94 (b)	102
Clínico	Negativo	30 (c)	18 (d)	48
Total		38	112	150 ^N

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	$8 / 38$	=	0,21 (baixa)
Especificidade	=	$d / (b+d)$	=	$18 / 112$	=	0,16 (baixa)
Valor Preditivo (+)	=	$a / (a+b)$	=	$8 / 102$	=	0,078 (baixo)
Valor Preditivo (-)	=	$d / (c+d)$	=	$18 / 48$	=	0,38 (baixo)
Acurácia	=	$(a+d) / N$	=	$26 / 150$	=	0,17 (baixa)

Pelos resultados encontrados, verifica-se que se considerarmos o meio de cultura como o teste padrão para identificar os indivíduos doentes e sadios, o exame clínico não é recomendável para proceder ao diagnóstico da faringite, pois os vários indicadores utilizados apresentaram baixos valores.

Exemplo 2: Avaliando-se uma amostra de 200 indivíduos, que diagnosticados quanto a pneumonia através raios X simples, tiveram outra avaliação por meio de auscultação.

Sendo o raio X, o método padrão, faça suas conclusões sobre os dados da Tabela 3.

Tabela 3 - Diagnóstico de pneumonia por raios X e auscultação.

Diagnóstico por auscultação	RAIOS X		
	SIM	NÃO	TOTAL
Positivo	195 (a)	5 (b)	200
Negativo	10 (c)	190 (d)	200
Total	205	195	400

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	195 / 205	=	0,95 (alta)
Especificidade	=	$d / (b+d)$	=	190 / 195	=	0,97 (alta)
Valor Preditivo (+)	=	$a / (a+b)$	=	195 / 200	=	0,98 (alto)
Valor Preditivo (-)	=	$d / (c+d)$	=	190 / 200	=	0,95 (alto)
Acurácia	=	$(a+d) / N$	=	385 / 400	=	0,96 (alta)

Face aos resultados obtidos, concluímos que o método de auscultação apresenta resultados apreciáveis quanto ao diagnóstico de pneumonia considerando o método de raios X, como o método padrão de diagnóstico.

Exemplo 3: Para avaliação de 120 pacientes, quanto a incidência de câncer de próstata, foram realizados os métodos de toque retal (padrão) e PSA, obtendo-se os dados da Tabela 4. Faça as suas conclusões.

Tabela 4 - Diagnóstico de câncer de próstata por toque retal e PSA.

PSA	TOQUE RETAL		
	POSITIVO	NEGATIVO	TOTAL
POSITIVO	06 (a)	4 (b)	10
NEGATIVO	34 (c)	76 (d)	110
TOTAL	40	80	120 ^N

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	6 / 40	=	0,15 (baixa)
Especificidade	=	$d / (b+d)$	=	76 / 80	=	0,95 (alta)
Valor Preditivo (+)	=	$a / (a+b)$	=	6 / 10	=	0,60 (médio)
Valor Preditivo (-)	=	$d / (c+d)$	=	76 / 110	=	0,69 (médio)
Acurácia	=	$(a+d) / N$	=	82 / 120	=	0,68 (média)

Os resultados obtidos indicam que o diagnóstico realizado por PSA, para diagnóstico de câncer de próstata, apresentam baixo valor de sensibilidade, ou seja, grande quantidade de falsos negativos.

Quanto aos valores preditivos (+) e (-), assim como a precisão, tiveram resultados avaliados como regulares.

Em relação a especificidade o seu valor é bastante elevado, por conseguinte, quando o PSA for julgado negativo é bastante confiável o seu diagnóstico.

Exemplo 4: Avaliando-se 80 indivíduos, para diagnóstico de câncer de laringe, foram utilizados para diagnóstico os métodos de biópsia (padrão) e punção de linfonodo, sendo os dados constantes da Tabela 5.

Tabela 5 - Diagnóstico de câncer de laringe por biópsia e punção de linfonodo

PUNÇÃO DE LINFONODO	BIÓPSIA		
	POSITIVO	NEGATIVO	TOTAL
POSITIVO	45 (a)	5 (b)	50
NEGATIVO	15 (c)	15 (d)	30
TOTAL	60	20	80 ^N

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	45 / 60	=	0,75 (alta)
Especificidade	=	$d / (b+d)$	=	15 / 20	=	0,75 (alta)
Valor Preditivo (+)	=	$a / (a+b)$	=	45 / 50	=	0,90 (alto)
Valor Preditivo (-)	=	$d / (c+d)$	=	15 / 30	=	0,50 (médio)
Acurácia	=	$(a+d) / N$	=	60 / 80	=	0,75 (média)

Quanto aos resultados verificamos que o valor preditivo (+) foi o destaque, apresentando-se igual a 0,90. Os demais, como sensibilidade e especificidade foram considerados regulares.

Exemplo 5: Numa amostra de 90 indivíduos, foram empregados testes diagnósticos para avaliação de presença de enfisema pulmonar. Foram utilizados raios X (teste padrão) e exame clínico, segundo dados da Tabela 6.

Tabela 6 - Diagnóstico de enfisema pulmonar por raios X e exame clínico.

EXAME CLÍNICO	RAIOS X		
	POSITIVO	NEGATIVO	TOTAL
Positivo	31 (a)	19 (b)	50
Negativo	31 (c)	9 (d)	40
TOTAL	62	28	90

Cálculos:

Sensibilidade	=	$a / (a+c)$	=	$31 / 62$	=	0,50 (média)
Especificidade	=	$d / (b+d)$	=	$9 / 28$	=	0,32 (baixa)
Valor Preditivo (+)	=	$a / (a+b)$	=	$31 / 50$	=	0,62 (médio)
Valor Preditivo (-)	=	$d / (c+d)$	=	$9 / 40$	=	0,23 (baixo)
Acurácia	=	$(a+d) / N$	=	$40 / 90$	=	0,44 (baixa)

Conclui-se que sensibilidade, especificidade, valores preditivo (+) e (-), e precisão apresentam baixos valores, indicando que o exame clínico não é um bom teste diagnóstico para identificação ou não de enfisema pulmonar.

Razões de verossimilhança ("likelihood ratios")

As razões de verossimilhança são uma forma de descrever o desempenho de um teste diagnóstico. Elas resumem o mesmo tipo de informação que a sensibilidade/especificidade e podem ser usadas para calcular a probabilidade de doença depois de um teste positivo ou negativo.

Uma vez que as razões de verossimilhança são expressões em chances, para entendê-las é necessário, primeiro, distinguir chances de probabilidade. Probabilidade, usada para expressar sensibilidade, especificidade e valor preditivo, é a proporção de pessoas nas quais uma determinada característica (tal como o teste positivo) está presente. Chances, por outro lado, é a razão de duas probabilidades. Chances e probabilidade contêm a mesma informação, mas a expressam de formas diferentes. As duas podem ser interconvertidas usando fórmulas simples:

$\text{Chances} = (\text{Probabilidade do evento}) / (1 - \text{probabilidade do evento})$

$\text{Probabilidade} = (\text{Chances}) / (1 + \text{Chances})$

Esses termos são familiares à maioria, porque são usados no cotidiano. Por exemplo, podemos dizer que as chances são de 4:1 que um determinado time ganhe, ou que ele tem 80% de probabilidade de vencer.

A razão de verossimilhança (também interpretável como razão de probabilidades) para um determinado valor de um teste diagnóstico é dada em chances, e é definida como a probabilidade de um resultado do teste em pessoas com a doença, dividida pela probabilidade do resultado do teste em pessoas sem a doença.

As razões de verossimilhança expressam quantas vezes mais provável (ou menos) se encontra um resultado de um teste em pessoas doentes comparadas com as não doentes. Se um teste é dicotomizado (positivo/negativo), dois tipos de razões de verossimilhança descrevem sua capacidade para separar pessoas doentes de não doentes: um associado a um teste positivo e o outro a um teste negativo, como apresentado a seguir:

Razão de Verossimilhança de um resultado de Teste Positivo (RVP)

$\text{RVP} = (\text{sensibilidade}) / (1 - \text{especificidade})$

ou,

$\text{RVP} = (a / a+c) / (b / b+d)$

- quanto maior a RVP, melhor o teste

- para ser um bom teste a RVP deve ser muito maior que 1

- um resultado positivo é mais provável de ser verdadeiro positivo (sensibilidade) do que falso-positivo (1 - especificidade)

Razão de Verossimilhança de um resultado de Teste Negativo (RVN)

$\text{RVN} = (1 - \text{sensibilidade}) / (\text{especificidade})$

ou,

$\text{RVN} = (c / a+c) / (d / d+b)$

- quanto menor a RVN, melhor o teste
- quanto mais perto de 1, melhor será o teste
- um resultado negativo é mais provável de ser um verdadeiro resultado negativo (especificidade) do que um falso-negativo (1 - sensibilidade)

Uma das vantagens da razão de verossimilhança é que ela permite descrever a acurácia de um teste de resultados numéricos (exames hormonais e bioquímicos, por exemplo) em diversos pontos de corte de resultado. (Gomes, 2001). Outra vantagem é ser menos suscetível a mudanças em função da prevalência da doença.

Gomes (2001), apresenta o seguinte exemplo:

Tabela 7 - Exemplo do cálculo da RV (teste rápido para faringite estreptocócica)

Teste Rápido	Cultura Positiva	Cultura Negativa	TOTAL
Teste Rápido Positivo	21 (a)	6 (b)	27
Teste Rápido Negativo	1 (c)	154 (d)	155
TOTAL	22	160	182

Cálculos:

Sensibilidade	=	95,45%	
Especificidade	=	96,25%	
Valor Preditivo (+)	=	77,78%	
Valor Preditivo (-)	=	99,35%	
RV positiva	=	$(a / a+c) / (b / b+d)$	= 25,3
RV negativa	=	$(c / a+c) / (d / d+b)$	= 0,05

BIBLIOGRAFIA

- 1- Almeida Filho, N. & Rouquayrol, M.Z. - *Introdução à Epidemiologia Moderna*. Belo Horizonte/Salvador/Rio de Janeiro:COOPMED/APCE/ABRASCO, 1992.
- 2- Drummond, J.P. & Silva, E. - *Medicina Baseada em Evidências - Novo Paradigma Assistencial e Pedagógico*. São Paulo: Atheneu, 1998.
- 3- Fletcher, R.H.; Fletcher, S.W.; Wagner, E.H. - *Epidemiologia Clínica*. Porto Alegre:Artes Médicas, 1989.
- 4- Forattini, O.P. - *Epidemiologia Geral*. São Paulo:Artes Médicas, 1986.
- 5- Gomes, M.M. - *Medicina Baseada em Evidências - Princípios e Práticas*. Rio de Janeiro:Reichmann&Affonso, 2001.
- 6- Jekel, J.F.; Elmore, J.G. & Katz, D.L. - *Epidemiologia, Bioestatística e Medicina Preventiva*. Porto Alegre: ARTEMED, 1999.
- 7- Klotzel, K. - *Medicina Ambulatorial - Princípios Básicos*. São Paulo:EPU, 1999.
- 8- López, M. - *O Processo Diagnóstico nas Decisões Clínicas - Ciência, Arte, Ética*. Rio de Janeiro:Revinter, 2001.
- 9- Medronho, R.A. (org) - *Epidemiologia*. São Paulo: Atheneu, 2002.
- 10- Pereira, M.G. - *Epidemiologia - Teoria e Prática*. Rio de Janeiro:Guanabara Koogan, 1995.
- 11- Rouquayrol, M.Z. - *Epidemiologia & Saúde*. Rio de Janeiro:MEDSI, 1994.

"Ao avaliar a acurácia diagnóstica do teste rápido para detecção de antígeno estreptocócico, Joslyn et al. estudaram 182 pacientes com faringite aguda. A prevalência de faringite estreptocócica, diagnosticada pelo teste padrão, a cultura da orofaringe, era de 12% (22 pacientes). O teste rápido foi capaz de diagnosticar 21 desses casos e afastar a probabilidade de ocorrência da doença em 154 dos 160 com cultura negativa (tabela 7). O teste rápido foi capaz de detectar quase todos os casos de faringite estreptocócica (95,45% - sensibilidade). Diante de um resultado positivo, mais de 3/4 (77,78% - VP+) dos pacientes são verdadeiros positivos. O resultado positivo aumenta em 25 vezes a chance da doença (RVpositiva). A capacidade de refutar o diagnóstico também foi boa. Por esse estudo, o uso do teste rápido permite um diagnóstico preciso, rápido e uma conduta terapêutica imediata.