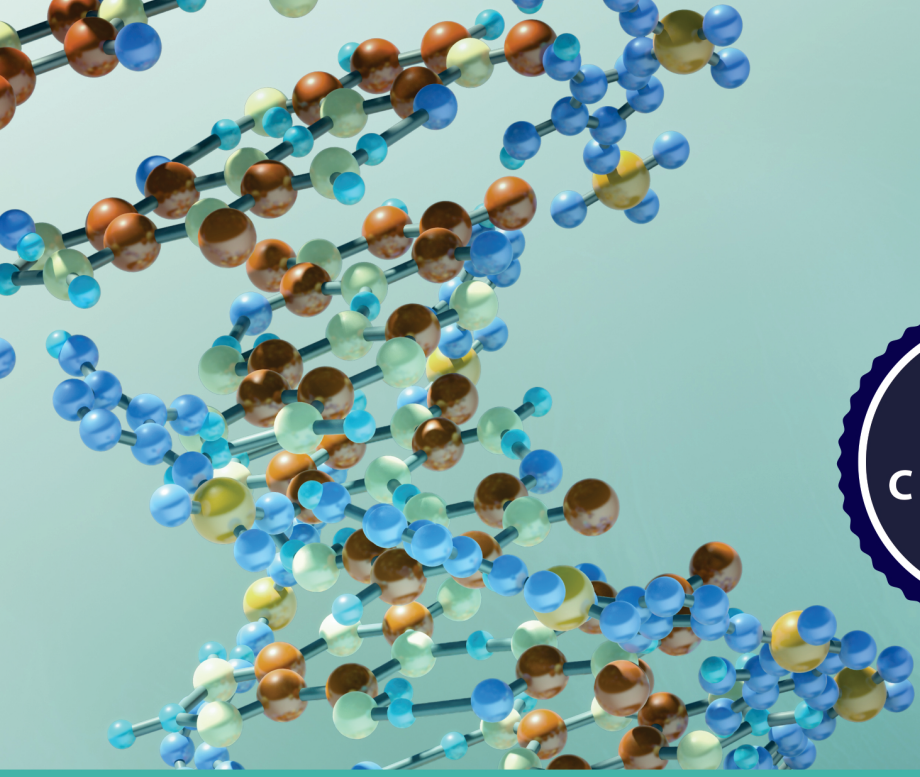
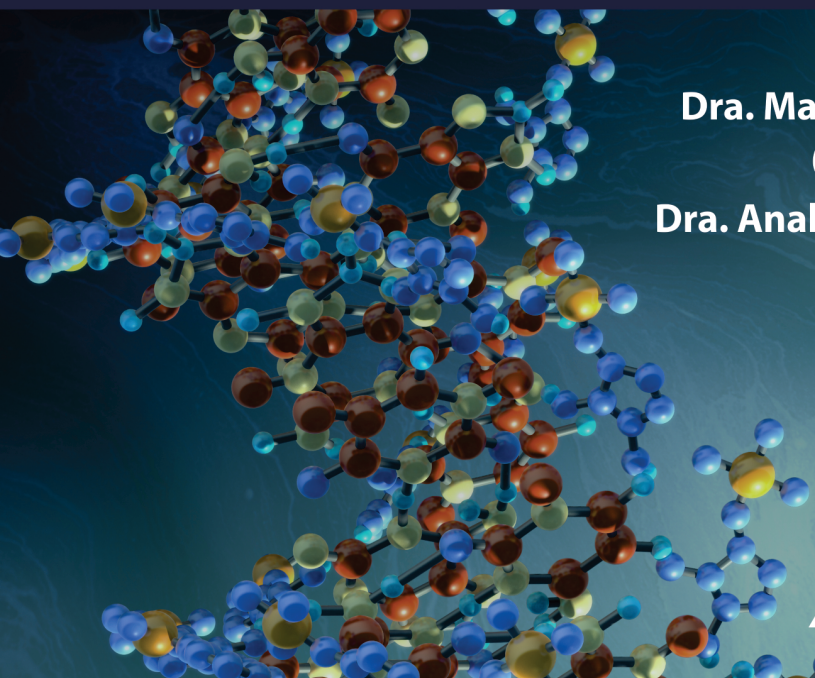


3ª EDIÇÃO
REVISTA E AMPLIADA

CONTÉM
CASOS
CLÍNICOS

BIOQUÍMICA CLÍNICA



Dr. Kanaan
Dra. Maria Alice Terra Garcia
(in memoriam)
Dra. Analúcia Rampazzo Xavier

Atheneu

Bioquímica Clínica

3ª EDIÇÃO



www.atheneu.com.br

(21) 99165-6798 Facebook.com/editoraatheneu Twitter.com/editoraatheneu Youtube.com/atheneueditora

Bioquímica Clínica

3ª EDIÇÃO

SALIM KANAAN

MARIA ALICE TERRA GARCIA
(*in memoriam*)

ANALÚCIA RAMPAZZO XAVIER

Atheneu
Rio de Janeiro • São Paulo
2023

EDITORA ATHENEU

São Paulo — *Rua Maria Paula, 123 – 18º andar*
Tel.: (11)2858-8750
E-mail: atheneu@atheneu.com.br

Rio de Janeiro — *Rua Bambina, 74*
Tel.: (21)3094-1295
E-mail: atheneu@atheneu.com.br

CAPA: Equipe Atheneu
PRODUÇÃO EDITORIAL: Efe Pê Ltda.

CIP-BRASIL. CATALOGAÇÃO NA PUBLICAÇÃO
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ

K24b
3. ed.

Kanaan, Salim
Bioquímica clínica / Salim Kanaan, Maria Alice Terra Garcia, Analúcia Rampazzo Xavier. - 3. ed. - Rio de Janeiro : Atheneu, 2022.
24 cm.

Inclui bibliografia e índice
ISBN 978-65-5586-580-6

1. Bioquímica clínica. I. Kanaan, Salim, II. Garcia, Maria Alice Terra. III. Xavier, Analúcia Rampazzo. IV. Título.

22-79798

CDD: 612.015
CDU: 612.015

Meri Gleice Rodrigues de Souza - Bibliotecária - CRB-7/6439

02/09/2022 02/09/2022

KANAAN, S.; TERRA GARCIA, M. A.; XAVIER, A. R.
Bioquímica Clínica – 3ª edição

© Direitos reservados à EDITORA ATHENEU – Rio de Janeiro, São Paulo, 2023

A Importância dos Exames Laboratoriais Confiáveis

1.1

Analúcia Rampazzo Xavier
Salim Kanaan

A importância dos exames laboratoriais confiáveis

A avaliação da efetividade do laboratório clínico e de sua contribuição para os desfechos tem sido matéria de crescentes discussões na literatura. A solicitação e a interpretação correta dos testes laboratoriais melhoram os diagnósticos clínicos e, por isso, também têm um grande impacto nos custos financeiros da assistência à saúde. A interpretação correta da informação, ou seja, o ato de discernir o significado e a importância do resultado de um determinado exame é a etapa final e mais crítica de uma série de eventos complexos que pode ser conhecida como o "ciclo de realização do exame" (Figura 1.1.1). Além do conhecimento em interpretar os resultados dos exames, o ato de reconhecer as etapas críticas de sua realização, bem como averiguar as medidas para diminuir a variabilidade desses diversos processos, permite uma avaliação da qualidade da informação obtida.

Na prática, todas as etapas desse ciclo de eventos podem sofrer a influência de fatores de variabilidade, com potencial impacto na validade da informação gerada. Vários pontos precisam ser averiguados como:

- **Desempenho técnico:** é a base de qualquer evidência. Além de precisão, exatidão, linearidade, os fatores de variabilidade pré-analíticos também devem ser considerados no desempenho técnico.
- **Desempenho diagnóstico:** averiguar a sensibilidade e especificidade dos métodos, analisando os valores preditivos positivos e negativos e a taxa de verossimilhança, importantes para uma correta análise de resultados.
- **Benefício clínico:** o impacto clínico dos exames deve ser baseado na estratégia diagnóstica, na estratégia terapêutica e no desfecho clínico (consequência dos itens anteriores).
- **Benefício operacional:** o benefício operacional pode, por exemplo, diminuir o tempo de internação de um paciente, reduzir a necessidade de recursos humanos, e ainda, a utilização de outros recursos da saúde.
- **Benefício econômico:** baseia-se na escolha de testes laboratoriais que possam gerar mais recursos ou minimizar gastos excessivos de alocação de recursos. Nem sempre o teste mais caro é necessário, bem como nem sempre o mais barato tem qualidade inferior.

Do ponto de vista do laboratório, as fontes de possíveis variabilidades podem ser divididas entre as etapas pré-analíticas, analíticas e pós-analíticas.

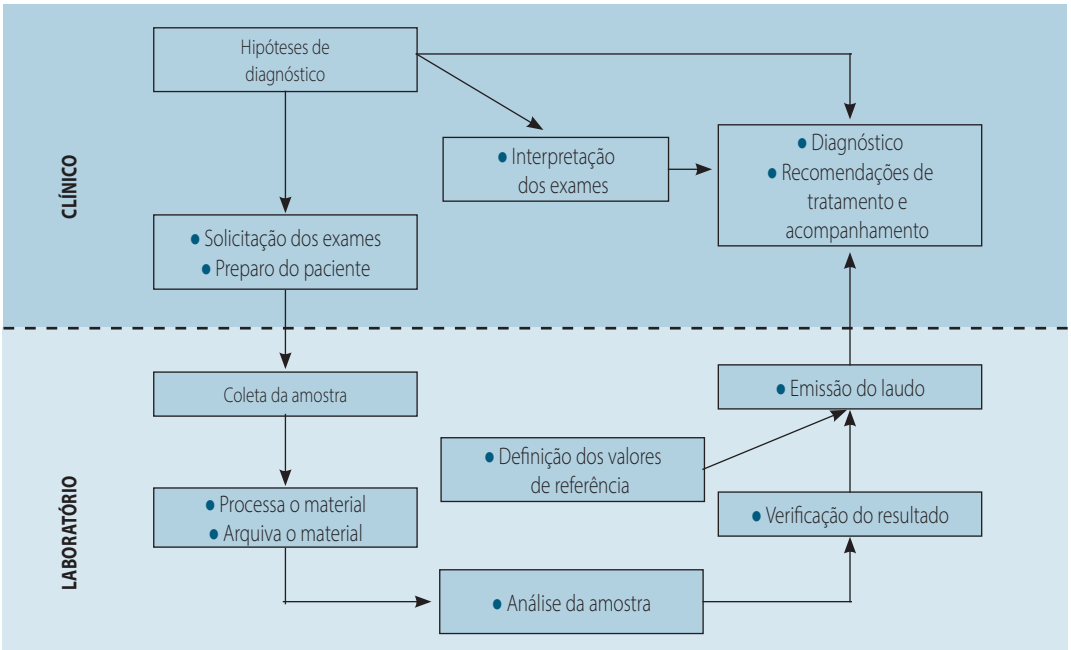


FIGURA 1.1.1 – Fluxograma de realização de exames.

Fonte de variabilidade nos resultados

A variabilidade dos resultados de um laboratório pode ser derivada do preparo inadequado do paciente (hora da coleta, jejum, certos alimentos, exercício físico e medicações), ou da coleta e manipulação das amostras (que podem gerar erros nos procedimentos de identificação de amostras e treinamento adequado da equipe técnica). Nesse caso, essas variáveis podem ser denominadas ***pré-analíticas***, pois ocorrem antes da análise laboratorial propriamente dita (Figura 1.1.2).

Quando falamos em ***variabilidade analítica*** devemos pensar no método analítico em si (reagentes, equipamentos, procedimentos e recursos humanos). Essa etapa, em especial, recebeu a atenção de especialistas em controle de qualidade nas últimas décadas, o que gerou a minimização dos erros analíticos, ocasionada pelo oferecimento de testes de proficiência, programas de credibilidade, auditorias externas e internas, onde a qualidade dos processos do laboratório é minuciosamente verificada. A automação dos métodos laboratoriais também pode ser responsabilizada pela diminuição de variabilidade nesta etapa do ciclo de realização de exames (Figura 1.1.2).

A ***variabilidade pós-analítica*** ocorre entre o término do método analítico e a assimilação dos resultados pelo clínico. A fonte principal de erro nesta etapa é a transcrição dos resultados, entretanto, os famosos "erros de digitação" estão se tornando cada vez menos frequentes, devido ao processo de interface entre os equipamentos e o sistema de informatização do laboratório. Cabe salientar a importância, nesta fase, da qualidade de laudos impressos, que muitas vezes não são claros nem adequados quanto ao formato, prejudicando a interpretação dos resultados (Figura 1.1.2).

Erros laboratoriais

Podem ser considerados erros a ocorrência e a verificação qualquer "defeito". Desde a requisição dos testes até o relatório dos resultados e

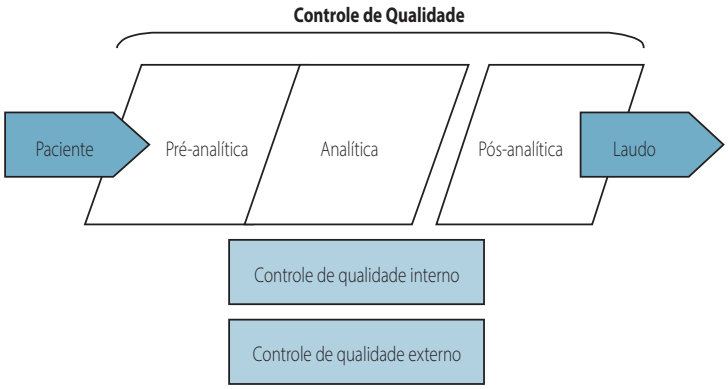


FIGURA 1.1.2 – Etapas do sistema de qualidade.

sua adequada interpretação. Apesar dos esforços dos laboratórios em eliminá-los, inevitavelmente nos deparamos com resultados de testes com erros. Uma revisão recente encontrou uma considerável atribuição da fase pré-analítica na ocorrência destes erros, sendo a troca de amostras na coleta a mais comum.

A possibilidade de erro deve ser levantada quando o resultado é impossível ou não fisiológico, quando for inconsistente com os resultados prévios do mesmo paciente ou incompatível com os resultados de outros testes realizados na mesma amostra, ou difere do esperado pelos achados clínicos.

Após a constatação da presença de erro, o laboratório deve identificar a causa e corrigi-la o mais rápido possível. Para isso o laboratório conta com ferramentas de controle de qualidade que facilitam e minimizam a ocorrência de resultados e análises errôneas.

Sistemas de controle de qualidade

Para realizarmos o controle de qualidade total, devemos dar ênfase a duas etapas importantes:

- A primeira é implantar o controle de qualidade intralaboratorial, também chamado de controle de qualidade interno, realizando análises diárias de amostras-controle para checar a ***precisão*** de cada metodologia.
- A segunda etapa é implantar o controle de qualidade interlaboratorial, também chamado de controle externo, cuja finalidade é realizar análises de amostras-controle recebidas periodicamente de fonte externa para comparar o nível de ***exatidão*** com diferentes laboratórios.

Os conceitos de precisão e exatidão são considerados qualitativos, embora possam refletir a repetitividade e reprodutibilidade de um método. A *precisão de medição* reflete o grau de concordância entre resultados de medição obtidos sob as mesmas condições (repetitividade). A *exatidão (acurácia) de medição* mede o grau de concordância entre o resultado de uma medição e um valor verdadeiro do mensurado (Figura 1.1.3).

Podemos confiar nas análises de um determinado método quando este estiver sob controle, tanto na sua exatidão como na precisão, e se as análises estiverem dentro dos limites aceitáveis de erro (LAE). Para a implantação do programa de controle de qualidade devemos inicialmente estabelecer certas normas a serem cumpridas, como:

- Selecionar o pessoal técnico, informando-o e treinando-o em cada fase da implantação do controle, através de reciclagens periódicas e palestras nas várias seções do laboratório. Realizar a manutenção e calibração de equipamentos e instrumentos de medições em locais de boa procedência, com mão de obra

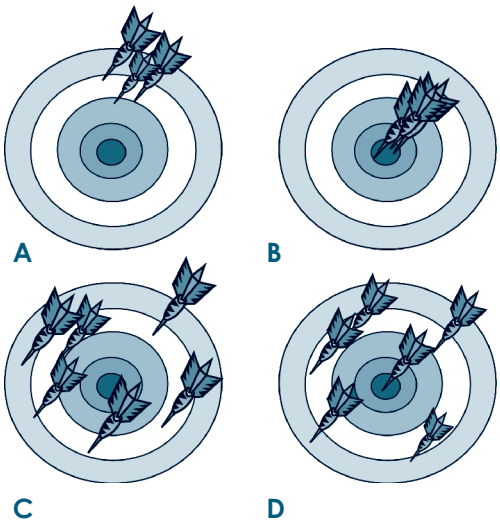


FIGURA 1.1.3 – Conceitos de exatidão e precisão. **A.** Resultados precisos e inexatos; **B.** Resultados precisos e exatos; **C.** Resultados imprecisos e inexatos; **D.** Resultados imprecisos e exatos.

especializada. Elaborar os reagentes com produtos químicos de qualidade e seguindo as recomendações das normas regulamentadoras específicas. Adequar e padronizar os métodos a serem utilizados. Proceder à coleta e conservação das amostras de acordo com a metodologia empregada. Realizar a lavagem e limpeza da vidraria com substâncias especiais, evitando contaminação química e biológica. Propiciar aos funcionários boas condições de trabalho.

Características de um sistema de controle

O sistema de controle de qualidade tem como finalidade não só informar sobre a precisão e exatidão de cada método, mas também deve ser sensível o suficiente para detectar variações pequenas nas diversas fases; além disso, deve ser de fácil implantação, montagem e interpretação.

Um sistema bem montado deve ter a capacidade de revelar a presença de qualquer tipo de erro e, ao mesmo tempo, conseguir comparar o desempenho de método, técnicas e equipamentos, treinamento de recursos humanos, entre outras etapas da realização dos exames laboratoriais.

Dentre as fases necessárias para a implantação de um sistema de controle, podemos destacar a importância da doutrinação do pessoal do laboratório com a filosofia de controle. Quando os recursos humanos do laboratório estão envolvidos com a qualidade, as demais etapas ficam mais fáceis de serem implantadas. Entre as fases necessárias ao controle de qualidade, o laboratório deve sempre preparar amostras-controle, que sejam estáveis, com concentrações conhecidas normais e anormais (patológicas). Estas amostras também podem ser adquiridas comercialmente e devem ter os LAE estabelecidos para cada componente (no caso dos soros comerciais, estes valores são conhecidos e fornecidos pelos fabricantes).

Deve-se então preparar para cada método um cartão de controle baseado nos limites aceitáveis de erro e, diariamente, anotar neste cartão os resultados obtidos com a análise da amostra-controle; concomitantemente deve-se examinar e reconhecer os resultados que estão situados "dentro" e "fora" do considerado controle. Quando os resultados estiverem "fora" do controle, checar todos os reagentes, padrões, equipamentos, lavagem de material, cálculos e outras possíveis variáveis. Ao verificar os cartões diariamente, semanalmente e mensalmente, poderemos reconhecer as tendências de cada método.

Soluções de concentração conhecida

Se a análise for para determinar a *precisão* de um método, não necessariamente precisamos conhecer o valor real da amostra testada, ao contrário da determinação da *exatidão* de uma metodologia. No controle de qualidade do laboratório clínico, utilizamos dois tipos diferentes de soluções, de concentrações já conhecidas: padrões e soros-controle.

Solução-padrão

Serve para a calibração dos métodos ou cálculo das concentrações. Os valores são determinados por pesagem.

■ Padrões primários

São soluções que podem ser preparadas em meio aquoso ou em outros solventes, que apresentam unicamente a substância quimicamente pura a ser determinada. Possuem concentração conhecida, pois ela é preparada em balão volumétrico após pesagem quantitativa em balança analítica. Estas soluções-padrão podem ser elaboradas no próprio laboratório ou obtidas comercialmente. Esses padrões não têm as mesmas características de um soro, pois na solução-padrão a substância está isolada e pura, enquanto no soro humano encontramos outras substâncias e interferentes. Além disso, estas soluções-padrão não são submetidas a todas as manipulações que sofrem as amostras, já que a análise do padrão só entra no final, devido ao seu objetivo, que é o de determinar os erros de calibração dos aparelhos (fotômetros) ou calcular o fator do cálculo de concentração.

■ Padrões em meios aquosos adicionados com proteínas

São soluções preparadas em um meio aquoso, com a maioria das substâncias utilizadas em diagnóstico clínico, as quais são adicionadas de proteínas em concentrações muito similares às encontradas no soro humano normal. Estas soluções-padrão não têm as mesmas características físico-químicas encontradas no soro humano, como a viscosidade, que é importante até mesmo para uma simples medição de volume nas pipetas. As proteínas que são adicionadas nestas soluções e, também, os outros constituintes são exatos, pois são pesados em balança analítica e preparados em balão volumétrico.

■ Padrões em soro humano

Há uma procura na elaboração de um soro humano artificial sem perda das características físico-químicas, que tenha valores absolutos independentes de metodologia analítica. Estes soros devem ter valores conhecidos e estruturas proteicas intactas, que possam migrar eletroforeticamente de maneira similar às do soro humano.

Este tipo de padrão tem uma dupla finalidade:

- Padronização de técnicas como padrões de calibração.
- Controle de qualidade como soros de referência.

Soro-controle

São soluções cujos valores são obtidos por análise química, e assim para um mesmo soro e uma mesma substância obtemos resultados diferentes, dependendo do método utilizado.

- **Soro animal:** é o soro obtido de equinos ou bovinos. Não podemos esquecer que este soro tem características muito diferentes das do soro humano.
- **Soro humano:** pode ser obtido através de *pool* de soros diferentes ou pode ser encontrado na forma de liofilizados comerciais.
- ***Pool* de soros:** é preparado no próprio laboratório com sobras de amostras não hemolisadas, não lipêmicas e não ictéricas.
- **Liofilizados comerciais:** são reconstituídos adicionando um volume certo de água destilada no momento da utilização. As amostras-controle comerciais são acompanhadas de folhetos explicativos onde é descrita a metodologia empregada com os valores aceitáveis (geralmente média ± 2 desvios-padrões) para cada constituinte.

Os soros-controle podem ser caracterizados como "normais, anormais ou desconhecidos" em relação às concentrações das substâncias.

Cálculo estatístico

Utiliza-se o cálculo estatístico como ferramenta em controle de qualidade para analisar parâmetros e medir o grau de *precisão* ou *reprodutibilidade*. Através destes cálculos, podemos inferir a presença de erros, verificando o momento de sua ocorrência, possibilitando a identificação de suas causas. Esta maneira de abordar

o assunto exige um sistema com fundamental valorização quantitativa do soro-controle.

Material estatístico utilizado para cálculos no controle de qualidade:

- **valor médio:** é o valor que aponta para onde mais se concentram os dados de uma distribuição. Pode ser considerado o ponto de equilíbrio das frequências, num histograma;
- **desvio-padrão:** é o valor que quantifica a dispersão dos eventos sob distribuição normal, ou seja, a média das diferenças entre o valor de cada evento e a média central;
- **coeficiente de variação:** é uma medida de dispersão que se presta para a comparação de distribuições diferentes;
- **limite de atenção:** limites estabelecidos que denotem atenção à variação de valores em relação com à média;
- **limite de alarme:** limite estabelecido em que o valor isolado ultrapassa os limites aceitáveis de erro.

Exemplo de formulário para estatística

Dados:

Analito: Glicose
Método: GOD-PAP
*Soro-controle:*_____
*Nº Lote:*_____
Unidade de medida: mg/100 mL

Data	n	X_1	$X_1 - X$	$(X_1 - X)^2$
	1	100	0,5	0,25
	2	102	1,5	2,25
	3	98	2,5	6,25
	4	99	1,5	2,25
	5	101	0,5	0,25
	6	99	1,5	2,25
	7	102	1,5	2,25
	8	101	0,5	0,25
	9	104	3,5	12,25
	10	100	0,5	0,25
	11	99	1,5	2,25
	12	100	0,5	0,25
	13	100	0,5	0,25
	14	98	2,5	6,25
	15	98	2,5	6,25
	16	102	1,5	2,25
	17	102	1,5	2,25
	18	101	0,5	0,25
	19	101	0,5	0,25
	20	101	0,5	0,25
	21	99	1,5	0,25
	22	100	0,5	0,25
	23	102	1,5	2,25
	24	103	2,5	6,25
	25	104	3,5	12,25
	26	99	1,5	2,25
	27	101	0,5	0,25
	28	98	2,5	6,25
	29	102	1,5	2,25
	30	101	0,5	0,25
	31	99	1,5	2,25
	Σ	3116		85,85

1) Valor médio:

$$\bar{x} = \left[\frac{\sum x_1}{n}\right] = \left[\frac{3116}{31}\right] = 100,5$$

2) Desvio-padrão:

$$s = \sqrt{\frac{\sum (x_1 - \bar{x})^2}{n-1}} = \sqrt{\frac{85,75}{31-1}} = \sqrt{2,858} = 1,69$$

3) Coeficiente de variação:

$$C.V. = \frac{s}{\bar{x}} = \frac{1,69}{100,5} = 0,0168 = 1,68\%$$

4) Limites de atenção:
Superior: $\bar{x}$ + 2s = 103,4
Inferior: $\bar{x}$ – 2s = 97,1

5) Limites de alarme:
Superior: $\bar{x}$ + 3s = 105,6
Inferior: $\bar{x}$ – 3s = 95,6

Obs.:
Data:
Assinatura:

Legenda: x_1 = valor da medida diária. x = valor real; s = desvio-padrão da média.

■ Considerações a serem desenvolvidas neste processo

1) Escolher o método a ser analisado, como por exemplo, a glicemia. Esta valorização é obtida através de experiências aleatórias, em número não menor do que 20.
2) O valor médio, que é a média aritmética, é uma medida chamada de posição; estas medidas de posição são valores centrais, e ao redor encontram-se as demais medidas de um conjunto de valores.
3) A medida de dispersão é o desvio-padrão que se relaciona com a dispersão das posições de todos os dados de um conjunto de valores em relação a um valor considerado central.
4) Escolha do intervalo de confiança e a distribuição de valores, sendo que o intervalo de confiança tem sua fundamental importância na avaliação da precisão dos métodos utilizados. Então a análise que fazemos do nosso soro-controle será considerada confiável se seus resultados se encontrarem dentro do intervalo de confiança; se os resultados obtidos forem inversos, estes deverão ser rejeitados, implicando este fato na busca dos fatores ou causas de erros que os afastaram do valor experimental daquele intervalo considerado como fidedigno.

A curva de distribuição de valores (Figura 1.1.4) é obtida colocando em abscissas os valores da medição, e em ordenadas a frequência relativa que esses valores oferecem numa série de medições. representa a média aritmética dos valores.

O desvio-padrão (DP ou s) é uma medida de dispersão derivada de curva de distribuição de valores que expressa uma série de dados nas proximidades do valor médio.

O coeficiente de variação é o desvio-padrão expresso em percentual e é também usado para expressar a precisão.

Mapas de uso diário no controle de qualidade

Estes mapas têm como finalidade permitir a rápida visualização das variações diárias, semanais e mensais existentes nos diferentes métodos da rotina laboratorial. Para que o la-

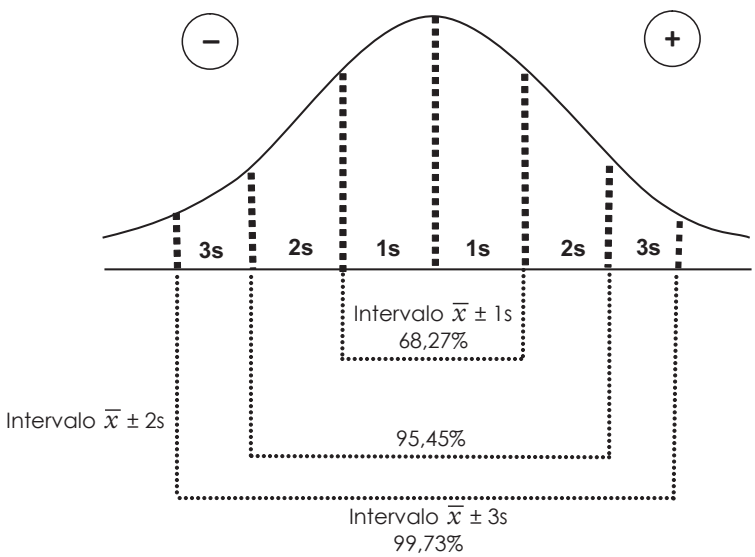


FIGURA 1.1.4 – Curva de distribuição de valores.

Legenda: $\overline{x}$ = valor da média diária; s = desvio padrão.

boratório tenha uma boa qualidade dos seus analitos, calculam-se os limites aceitáveis de erro (LAE) de cada método utilizando um *pool* de soros preparado no próprio laboratório ou, o que é mais utilizado, um soro comercial.

O soro comercial tem várias vantagens em relação ao soro preparado no laboratório, entre elas o fato de ser liofilizado e ter validade pré-estabelecida. A validade, geralmente superior a 1 ano, diminui a frequência do estabelecimento de novos limites de controle e também possui os valores médios, desvio-padrão e/ou coeficiente de variação já estabelecidos.

Como estabelecer estes mapas de controle

Para montar mapas de controle, devem-se ter em mãos os seguintes materiais e dados:

a) Utilizar papel milimetrado.

b) Dosar, no mínimo, 20 vezes cada analito contido no soro-controle (estabelecer estatisticamente os valores de $\bar{x}$, DP (s) e coeficiente de variância [CV]). A obtenção destes dados é feita com o padrão para cada analito como referência, dosando simultaneamente o soro-controle. Na rotina, o soro-controle e as amostras desconhecidas passam a ser determinadas contra o padrão de referência.

c) Estes valores determinados para cada analito são utilizados para o cálculo de média e desvio-padrão.

d) Com isso, estabelecemos os limites aceitáveis de erro (LAE), que podem ser calculados de dois modos:

- Utilizando-se o parâmetro de $\bar{x} \pm 2$ DP, que permitirá a colocação dos valores obtidos, em unidade de concentração, diretamente sobre os mapas de controle.
- Utilizando-se a fórmula de *Tonks* em termos percentuais (ver Figura 1.1.5). A utilização desta fórmula deve ser ajustada, pois o sistema apresenta deficiências, já que alguns valores calculados LAE em % podem necessitar de grandes ajustes.

$$\text{LAE \%} = \frac{\text{¼ da variação normal}}{\text{média da normalidade}}$$

FIGURA 1.1..5 – Fórmula de *Tonks*.

e) Elaborar o mapa de seguinte forma:

- Marcando na abscissa (eixo horizontal) os dias do mês e na ordenada (eixo vertical) os valores de média e desvio-padrão calculados; nos soros-controle, esses valores já são preestabelecidos.

O soro-controle deve ser dosado diariamente, em concomitância com as amostras desconhecidas, e os resultados obtidos para cada analito deverão ser plotados no mapa sob a forma de pontos (Figura 1.1.6).

Como interpretar os mapas de controle

São realizadas, diariamente, análises das amostras-controle de cada analito, juntamente com o desconhecido dos pacientes, utilizando os mesmos métodos nos padrões. Após estas determinações, podemos obter as seguintes possibilidades:

- **Dentro dos limites aceitáveis:** quando o método tem boa precisão e exatidão, em amostras-controle de concentrações corretas, quase todos os pontos ficam entre os limites $\bar{x} \pm 2$ DP, sendo distribuídos aproximadamente metade em cada lado da média, sendo que $^2/_3$ ficam entre $\bar{x} \pm 1$ DP, onde se dispõem cerca de 68% dos valores na curva de distribuição.
- **Quando um em cada 20 resultados ficar fora dos limites** (pois os limites de confiança são de 0,5%): entretanto, quando mais de dois pontos estão fora dos limites, o método deveria ser checado imediatamente, e o nosso controle analisado.
- **Resultados fora de controle:** se um resultado estiver "fora dos LAE", deverá ser repetida a análise da amostra-controle. Quando a

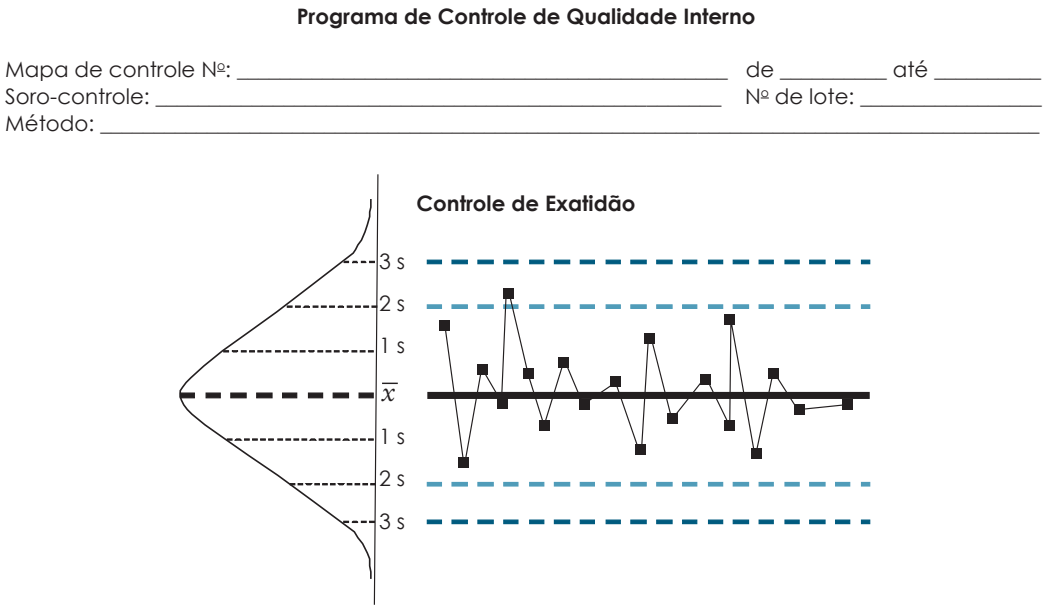


FIGURA 1.1.6 – Curva de *Gauss*. Mapa de Controle segundo Levey-Jennings.

segunda avaliação estiver dentro de controle, uma das hipóteses é que a reconstituição não foi completa (amostras congeladas e liofilizadas) ou houve algum problema na manipulação do soro-controle. Os resultados obtidos dos pacientes poderão ser liberados e o ponto fora deve permanecer no cartão, apesar de não ser incluído no cálculo do coeficiente de variação. Mas se a segunda análise estiver fora do controle, os resultados não são liberados.

Como analisar a perda da precisão

Se muitos resultados estiverem próximo de $\overline{x} \pm 2$ DP, pode-se observar uma perda de precisão do método. Esta situação é estatisticamente apresentada quando mais de cinco pontos ficam próximos de 2 DP ou 1/3 dos pontos fica "fora" de $\overline{x} \pm 1$ DP. A perda de precisão pode ser causada por vários motivos, entre eles:

- Pipetas de má qualidade ou descalibradas, provavelmente pelo uso excessivo na sua secagem (acima de 60ºC).
- Material de vidraria quimicamente sujo devido ao uso de detergentes inadequados ou enxágue insuficiente.
- Banho de água com temperatura oscilante devido a defeito no termostato ou localização inadequada (próximo de aparelhos de ar condicionado, aquecedores e fornos etc.).
- Condições emocionais do analista, ocasionadas por fadiga ou acúmulo de serviço.
- Tubos de leitura fotométrica inadequados apresentando paredes sem espessura uniforme.
- Defeito eletrônico no fotômetro originando variações nas leituras.
- Oscilações de rede elétrica.
- Célula fotoelétrica com "cegueira" parcial devido ao recebimento de luz muito intensa (nunca se deve usar o fotômetro enquanto um dos filtros não estiver no lugar, bem como se deve desligar o aparelho para a troca dos mesmos).

Como analisar a perda da exatidão

Configura-se como perda de exatidão quando mais de cinco pontos se aproximam dos limites de $\bar{x} \pm 2$ DP. A perda de exatidão é provocada geralmente por erro sistemático. Estes erros podem ser ocasionados pela troca dos controles ou valor incorreto nas dosagens, causados pelo preparo inadequado ou instabilidade dos reagentes do método, variação de temperatura de banho-maria durante a realização dos exames, alteração do tempo das fases de preparo ou leituras de comprimento de onda diferentes do recomendado.

Quando não há exatidão, as análises deverão ser suspensas e nenhum resultado pode ser liberado.

Como analisar a perda de sensibilidade

Quando as variações e os erros inerentes da metodologia e manipulação laboratoriais levam a uma reprodução sucessiva dos resultados, de forma repetitiva (se mais de cinco pontos seguidos, que deveriam oscilar em torno da média, apresentam o mesmo valor estatisticamente), podemos caracterizar uma perda da sensibilidade. A elevação dos valores de absorbância do branco da reação (ideal se abaixo de 0,100) e a presença de valores baixos de controle podem levar a uma perda de sensibilidade. Os arredondamentos nas leituras e/ou cálculo dos resultados e o desgaste da célula fotoelétrica, de modo que pequenas variações não sejam mais detectadas, podem também ser causas de uma sensibilidade reduzida do exame.

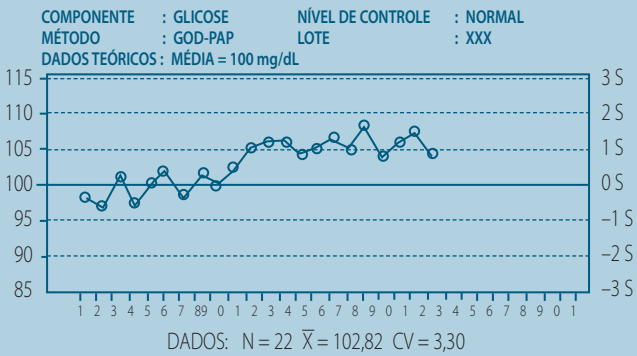
Resultados tendenciosos

Quando mais de 6 (seis) pontos estão situados do mesmo lado da média, esses resultados podem estar tendenciosos. A deterioração gradativa dos reativos (geralmente substratos enzimáticos) e defeitos nos aparelhos são exemplos que podem causar tendências nos resultados.

CASOS CLÍNICOS

■ Caso 1

Um analista clínico, responsável pela gestão da qualidade de um laboratório privado, ao verificar semanalmente os controles do equipamento automatizado de bioquímica, observou o seguinte resultado para o controle de nível normal para o analito glicose:



DADOS: N = 22 $\overline{X}$ = 102,82 CV = 3,30

Seguindo as regras de Westgard, ele observou que houve uma quebra de controle 10_{Xm}. Baseado no gráfico acima, qual a providência a ser tomada para controlar o sistema?

■ Caso 2

Em um laboratório público, o técnico de laboratório de plantão, após manutenção e passagem do controle interno diário, e liberou o uso do aparelho para a análise dos materiais dos pacientes. O responsável técnico, ao verificar o gráfico de Levey-Jennings, verificou que alguns analitos tiveram as regras de Westgard quebradas. Sendo observado para o analito magnésio – controle normal: 1_{3s}; fosfato – controle normal e patológico: R_{4s}.

Que tipo de erro ocorreu? O que deve ser feito com os analitos já dosados nestas condições?

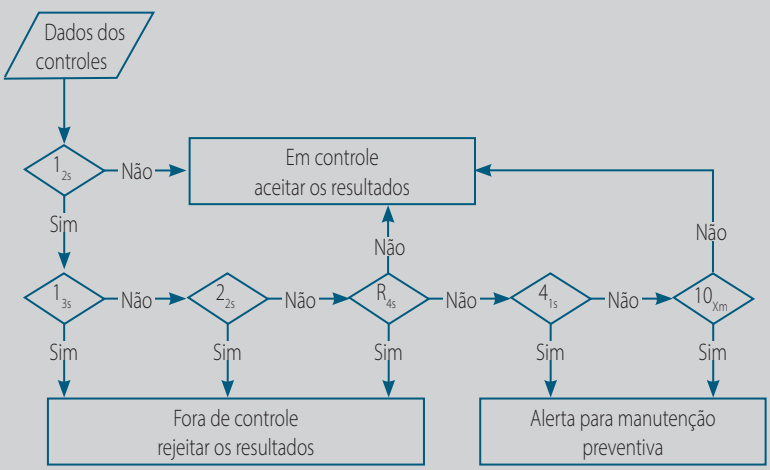
Bibliografia consultada

1. Barros E, Xavier RM, Albuquerque GC. Laboratório na Prática Clínica: Consulta Rápida. 1a ed. São Paulo: Artmed Editora, 2005.
2. Burtis CA, Ashwood RE, Bruns DE. TietzFundamentals of Clinical Chemistry. 6thed. St. Louis, Missouri: Saunders Elsevier, 2008.
3. Burtis CA, Ashwood RE, Bruns DE. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics. 4th ed. St. Louis, Missouri: Saunders Elsevier, 2006.
4. Henry JB. Clinical Diagnosis and Management by Laboratory Methods. 20th ed. Pennsylvania, USA: W.B. Saunders Company, 2001.
5. McPherson R, Pincus M. Henry's Clinical Diagnosis and Management by Laboratory Methods. 21th ed. St. Louis, Missouri: SaundersElsevier, 2006.
6. Neto GV, et al. Critérios para a Habilitação de Laboratórios Segundo os Princípios das Boas Práticas de Laboratório (BPL). Disponível em: <http://apostilas.cena.usp.br/Valdemar/BPL-ANVISA.pdf>. Acessado em 29/08/2012.

DISCUSSÃO DOS CASOS CLÍNICOS

Caso 1

Uma regra de Westgard 10_{Xm} quebrada significa que os valores do controle estão do mesmo lado da média em 10 ou mais dias consecutivos. Essas observações podem ocorrer para o valor de um dos controles ou dois níveis, significando a observação de 10 ou 5 dias respectivamente (5 observações para o nível normal e 5 para o patológico). No caso, houve quebra de um dos controles (normal) por mais de 10 dias consecutivos caracterizando um erro sistemático, em que há a necessidade de verificação do sistema, pois há perda de exatidão nas suas análises, uma vez que ao final do período mensal de análise pode haver desvio da média, e o aparelho ficará com erro de exatidão inserido. O fluxograma abaixo mostra quais providências devemos realizar para sanar o erro sistemático:



Fonte: Controles de Qualidade em Laboratório Clínico | Eng Produção e Mecatrônica – Academia.edu

Portanto, o erro sistemático visto é evitado com a manutenção preventiva do equipamento e análise correta do controle interno diário.

Caso 2

As análises deveriam ter sido rejeitadas, pois ambos as regras violadas se trata possivelmente de erros aleatórios, os quais denotam falta de precisão de suas análises e resultados dos pacientes com valores errados. O erro no controle nível normal do analitos magnésio (1_{3s}) mostra que o referido controle está 3 desvios padrões acima ou abaixo da média esperada, significando que os resultados das análises sob estas condições deveriam ter sido rejeitadas, ou melhor, nem analisadas antes de resolver o problema. A violação desta regra indica um aumento do erro aleatório, mas também pode significar eventualmente um erro sistemático de grandes dimensões. Como ação corretiva, deve ser verificado todo o sistema, da água às estabilidades dos reagentes, e eventualmente realizar a manutenção corretiva com calibração.

O erro encontrado nos controles normal e patológico do analito fosfato (R_{4s}) nos mostra que a diferença entre dois controles é maior que 4 desvios. Assim, quando o valor de um controle excede + 2 desvios e o valor do outro controle ultrapassa o – 2 desvios, cada observação ultrapassa 2 desvios, mas em direções opostas, fazendo uma diferença maior que 4 desvios padrões. Deve ser considerado que a diferença deve ser maior que 4 desvios, mesmo que um resultado não ultrapasse os 2 desvios aceitáveis. É indicativa da ocorrência de erros aleatórios, e como ação corretiva, deve-se verificar todo o sistema, como no caso anterior, inclusive a necessidade de ações corretivas com a calibração do equipamento.

1.2 Acurácia dos Testes Diagnósticos

Analucia Rampazzo Xavier
Gisele Caldas Alexandre
Hélia Kawa
Maria Luiza Garcia Rosa
Thiago Pavoni Gomes Chagas

Acurácia dos testes diagnósticos

Os testes diagnósticos são requisitados para a detecção, o monitoramento ou a predisposição à uma condição. Uma vez que o resultado deste teste é gerado, ele deve ser interpretado de modo a possibilitar a tomada de decisão médica. No entanto, a interpretação de testes diagnósticos pode trazer alguns desafios ao profissional de saúde. Um teste de diagnóstico é geralmente entendido como um teste realizado em laboratório, mas os princípios discutidos aqui aplicam-se igualmente bem às informações obtidas a partir da história clínica, do exame físico e dos procedimentos de imagem. Em muitos casos, a compreensão desses princípios ajudará o profissional de saúde a reduzir a incerteza diagnóstica.

Diante de qualquer resultado de um teste diagnóstico, o profissional de saúde deverá avaliar se este resultado é plausível. Um resultado, seja ele positivo ou negativo, pode ser verdadeiro ou falso, sendo esse o maior desafio na interpretação. A confiança do profissional de saúde no resultado é de grande importância para a tomada da decisão, e para orientar os passos seguintes de sua conduta. E esse próximo passo poderá significar a realização de um procedimento de risco (invasivo) ou de alto custo.

O presente capítulo pretende discutir as diferentes formas de um profissional de saúde avaliar a acurácia de um teste diagnóstico, ou seja, a sua capacidade de discriminar entre presença ou ausência de uma condição, que pode ou não ser considerada uma doença. Para facilitar a compreensão, indivíduos com a presença de qualquer condição serão chamados de doentes e os demais, sadios.

Serão utilizados os dados do ESTUDO DIGITALIS (Garcia Rosa, et al. – 2015), descrito no Box 1.2.1, no capítulo, para apresentar e discutir os conceitos abordados.

BOX 1.2.1. – Estudo Digitalis

O ESTUDO DIGITALIS, uma iniciativa de pesquisadores de diferentes unidades da UFF e da Fundação Municipal de Saúde de Niterói, examinou entre 2011 e 2012, 633 indivíduos de 45 a 99 anos, selecionados aleatoriamente, assistidos pelo Programa Médico de Família de Niterói-RJ. Todos os participantes foram examinados por médicos, realizaram uma Ecocardiografia por Doppler Tecidual (EDT) e dosaram o peptídeo natriurético cerebral (BNP) no plasma. Os examinadores não tiveram acesso a outras informações dos participantes. Indivíduos com fração de ejeção inferior a 50% (determinada pelo EDT), na presença de um sinal ou sintoma de insuficiência cardíaca foram classificados como portadores de insuficiência cardíaca com fração de ejeção reduzida (ICFER) (Jorge *et al.*, 2016). O BNP plasmático em pg/mL foi avaliado como teste para o diagnóstico de ICFER.

Qualidade do teste diagnóstico: acurácia e confiabilidade

Para que o resultado de um teste diagnóstico represente uma informação fidedigna e, assim, seja útil para a detecção, o diagnóstico, o monitoramento de uma determinada condição, ele deve ser avaliado sob vários aspectos. A confiabilidade metodológica, sua praticidade e viabilidade na rotina clínica, sua robustez diagnóstica e a acurácia do teste compreendem algumas características indispensáveis. A confiabilidade do teste diagnóstico refere-se a capacidade deste em fornecer resultados repetidos e consistentes. Particularmente com relação a acurácia de um teste, esta refere-se à quanto, em termos quantitativos ou qualitativos, um teste é útil para diagnosticar uma condição ou para predizê-la.

A acurácia de um teste diagnóstico: avaliando testes qualitativos

Estabelecer um diagnóstico é um processo imperfeito, resultando em uma probabilidade e não em uma certeza. Diferentes situações clínicas implicarão em diferentes probabilidades de um resultado estar correto. O Quadro 1.2.1 apresenta um quadro com as quatro possibilidades diante de um teste qualitativo: doença presente ou doença ausente. O resultado do teste será verdadeiro quando ele for positivo para os doentes (a) e negativo para os sadios (d) e será falso quando for positivo para os sadios (b) ou negativo para os doentes (c).

Como saber se a doença está presente ou ausente? O padrão ouro

Um teste diagnóstico é considerado útil quando ele é capaz de identificar corretamente a presença de determinada condição. Antes de ser adotado para uma rotina diagnóstica, o teste deve ser avaliado para verificar essa sua capacidade de acerto. Para essa validação é necessário comparar o teste em questão com um teste padrão ouro ou padrão de referência com alto poder discriminante. Para isso, opta-se pelo método com maior precisão disponível. Muitas vezes, esse padrão ouro pode ter um maior custo, ou ser mais invasivo, ou ser mais demorado, ou ainda, implicar em algum risco para o paciente. Tais características justificam

Quadro 1.2.1. Relação entre o resultado de um teste diagnóstico e a ocorrência da doença

		Doença	
		Presença	Ausência
Teste	Positivo	(a) Verdadeiro positivo	(b) Falso positivo
	Negativo	(c) Falso negativo	(d) Verdadeiro negativo

a busca de um novo teste, com uma acurácia semelhante e que seja menos invasivo, ou mais rápido, ou de menor custo.

Nas doenças que não são autolimitadas, cujo diagnóstico definitivo não é possível senão após alguns meses depois da suspeita diagnóstica, sua evolução pode ser selecionada como padrão ouro. Nessa situação, o período de acompanhamento deve ser longo o suficiente para que a enfermidade possa ter um diagnóstico definitivo.

No exemplo descrito no Box 1.2.1, o uso do Ecocardiograma por Doppler Tecidual (EDT) para diagnosticar a presença de disfunção de ventrículo esquerdo, caracterizado por uma fração de ejeção inferior a 50%, aliado aos sinais e sintomas de insuficiência cardíaca, é o padrão ouro para testar a acurácia do peptídeo natriurético cerebral (BNP) para o diagnóstico de insuficiência cardíaca com fração de ejeção reduzida (ICFER) (Lagoeiro *et al.,* 2016*)*.

Um outro exemplo sobre o uso do padrão ouro é o descrito por Piroozmand e colaboradores (2017), que testaram a acurácia dos anticorpos imunoglobulina G (IgG) anti *Helicobacter pylori* (*H. pylori*) na saliva e soro utilizando ELISA como método alternativo ao diagnóstico definitivo do *H. pylori* feito por histopatologia (padrão ouro). A sensibilidade e a especificidade foram de 75% e 79%, respectivamente. Os autores concluíram que os valores obtidos feitos pelo método ELISA salivar foram comparáveis aos do ELISA sérico e que aquele pode ser usado como alternativa para o diagnóstico da infecção por *H. pilory*.

Os testes utilizados como padrão ouro não são perfeitos e isso se estabelece como um desafio que precisa ser enfrentado. Assim, é importante considerar o teste mais acurado disponível e essa escolha, no entanto, não exclui a possibilidade de um novo teste ser mais acurado que o padrão ouro escolhido. Caso isso ocorra, a acurácia do teste apreciado será erroneamente mal avaliada.

■ Testes diagnósticos e os pontos de corte: conceituando o intervalo de referência

Com o objetivo de auxiliar a interpretação de resultados, o intervalo de referência é uma tentativa de definir as populações de indivíduos normais (sadios) e anormais (doentes). Para que os resultados dos testes diagnósticos possam ser assumidos como informações úteis e, portanto, possam cumprir com as suas finalidades já expostas anteriormente, é fundamental que sejam interpretados a partir de intervalos de referência válidos para a população na qual são aplicados. Para algumas análises, os intervalos de referência oferecem uma resposta simples de sim ou não, sendo utilizado um único ponto de corte para separar indivíduos normais (sadios) e aqueles anormais (doentes). Em outras situações, o intervalo de referência pode apresentar um ponto de corte no limite superior e outro ponto de corte no limite inferior. Nesse caso, o intervalo de referência define três grupos de pacientes: pacientes normais (sadios), pacientes anormais (doentes) com valores baixos, e pacientes anormais (doentes) com valore altos (Figura 1.2.1). Hipoteticamente, qualquer ponto de um intervalo de referência entre duas distribuições poderia diferenciar perfeitamente entre ambas a populações normais e anormais.

No entanto, infelizmente, como será melhor discutido posteriormente, não existem testes diagnósticos perfeitos, ou seja, testes que geram populações distintas de pacientes normais ou anormais. Entre os resultados do teste diagnóstico, forma-se uma área com certa sobreposição de valores, uma zona cinzenta, sobre a qual o teste é incapaz de discriminar a existência ou ausência da condição que está sendo investigada (Figura 1.2.1). Os resultados falsos positivos ou falsos negativos ocorrem devido a essa sobreposição das duas populações.

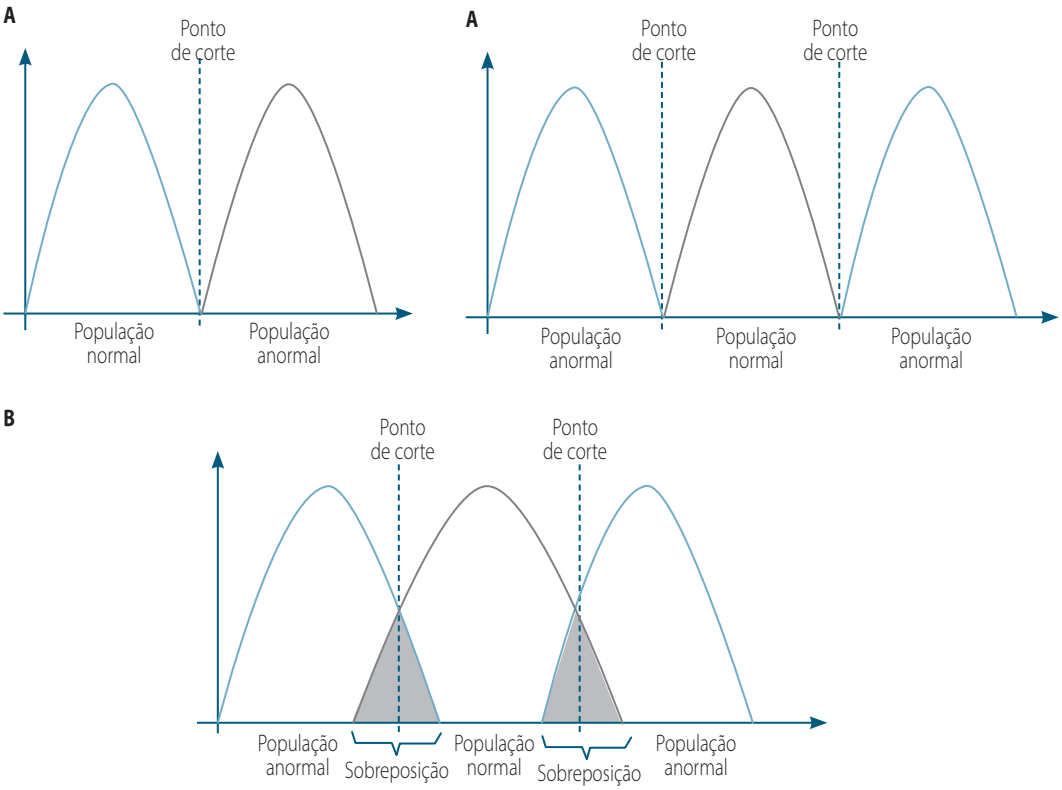


FIGURA 1.2.1 – Distribuição dos grupos de pacientes de acordo com o ponto de corte. A. Distribuição hipotética. B. Distribuição real.

Propriedades do teste que independem da condição do paciente: a sensibilidade e a especificidade

A condição do paciente avaliado interferirá na interpretação do resultado de um teste, como descrito a seguir. Contudo, há propriedades específicas do teste que independem da população em que é aplicado: a sensibilidade e a especificidade.

Para apresentar esses dois conceitos, relacionados à acurácia do teste diagnóstico será utilizado o exemplo descrito no Box 1.2.1. O BNP é dosado no plasma e seu resultado é quantitativo. Aqui, será empregado o ponto de corte de 41,5 pg/mL e, mais adiante, ficará claro como se chegou a esse ponto de corte. Os indivíduos com níveis de BNP < 41,5 pg/mL foram considerados negativos para ICFER e aqueles com níveis ≥ 41,5 pg/mL foram considerados positivos (Quadro 1.2.2).

Dos 633 indivíduos estudados (realizaram EDT, consulta médica e dosaram o BNP), 29 foram classificados como doentes (com ICFER) e 604 como sadios (sem ICFER). Entre os 29 indivíduos classificados como doentes, 23 (79,31%) apresentaram BNP ≥ 41,5 pg/mL (positivos), sendo classificados como verdadeiros positivos, isto é, doentes que o BNP, no ponto de corte escolhido, detectou. Os outros 6 doentes (20,69%) apresentaram BNP < 41,5 pg/mL (negativo), sendo, portanto, falsos negativos, isto é, doentes que o BNP no ponto de corte escolhido não detectou. Assim, a sensibilidade do BNP para detectar portadores de ICFER foi de 79,31%.

Quadro 1.2.2. Acurácia do BNP plasmático para o diagnóstico da ICFER.

BNP	ICFER		
	Sim	Não	Total
Teste positivo ≥ 41,5 pg/mL	23 Verdadeiros positivos	84 Falsos positivos	107
Teste negativo < 41,5 pg/mL	6 Falsos negativos	520 Verdadeiros negativos	526
Total	**29**	**604**	**633**

Fonte: Estudo Digitalis.

Pode-se definir a sensibilidade como a capacidade do teste em detectar os verdadeiramente doentes entre todos os doentes.

Sensibilidade = número de verdadeiros positivos/total de doentes

Sensibilidade = 23/ 29 = 79,31%

Já entre os 604 identificados como sadios, houve 520 (86,09%) que apresentaram BNP < 41,5 pg/mL (negativo), sendo classificados como verdadeiros negativos, isto é, indivíduos sadios que o BNP, no ponto de corte escolhido, detectou. Os outros 84 sadios (13,91%) apresentaram BNP ≥ 41,5 pg/mL (positivo), sendo classificados como falsos positivos, isto é, sadios que o BNP classificou erradamente como positivos. Desta forma, a especificidade do BNP foi de 86,09%.

A especificidade pode ser definida como a capacidade do teste em detectar os verdadeiramente sadios entre os sadios.

Especificidade = número de verdadeiros negativos/ total de sadios

Especificidade = 520/ 604 = 86,09%

A acurácia é definida como a capacidade do teste de discriminar doentes de sadios: do total de exames realizados qual o percentual de acertos.

Acurácia = (verdadeiros positivos + verdadeiros negativos)/ total de exames

Acurácia = (23+520)/ 633 = 90,52%

A sensibilidade e a especificidade são importantes para interpretar o resultado de um teste, medindo até que ponto o teste identifica corretamente os indivíduos normais (sadios) e anormais (doentes).

A Figura 1.2.2 apresenta três situações relacionando níveis diferentes de sensibilidade e especificidade para ficar claro a relação desses com a capacidade do teste de discriminar doentes e sadios. No ítem A da Figura 1.2.2, há a mesma proporção de verdadeiro positivo (VP), falso negativo (FN), falso positivo (FP) e verdadeiro negativo (VN). A conclusão é que o teste não discrimina. Enquanto, no B, a sensibilidade é maior que a especificidade, e o percentual de falsos negativos é menor que o de verdadeiros negativos. Quanto menor for o percentual de falsos negativos em relação aos verdadeiros negativos, maior é a probabilidade de um resultado negativo ser de um sadio e maior a confiança no teste negativo. Já em C, a especificidade é maior que a sensibilidade, havendo um menor percentual de falsos positivos em relação aos verdadeiros positivos. Quanto maior for essa diferença, ou seja, quanto menor o percentual de falsos positivos, maior é a probabilidade de um resultado positivo ser de um doente, maior a confiança que se pode depositar em um teste positivo.

Quanto maior a sensibilidade, maior a probabilidade de o teste negativo ser de um sadio, portanto, maior a confiança em se descartar o diagnóstico. Quanto maior a especificidade, maior a probabilidade de um teste positivo ser

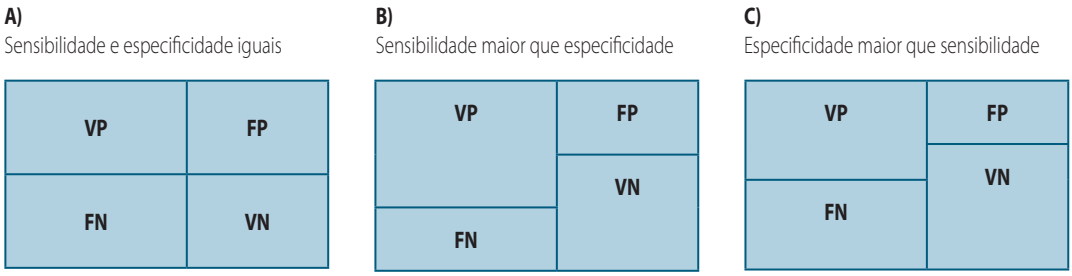


FIGURA 1.2.2 – Relação entre sensibilidade e especificidade e interpretação do resultado do teste.

de um doente, ou seja, maior a confiança em se confirmar um diagnóstico.

No exemplo ilustrado no Quadro 1.2.2, a especificidade foi um pouco maior que a sensibilidade (86,09 versus 79,31 %), induzindo uma maior confiança num teste positivo para confirmar a ICFER do que em um negativo para excluí-la.

Um teste sensível (aquele que é positivo na presença da doença) deve ser escolhido quando a penalidade de se considerar um doente como sadio e liberá-lo, é muito alta. Seria o caso de doenças consideradas potencialmente graves, como o linfoma de Hodgkin, ou ainda de doenças transmissíveis como, por exemplo no caso da tuberculose e sífilis. Com a sensibilidade alta, há menor proporção de falsos negativos, aumentando a segurança que um teste negativo seja de um sadio. A opção por um teste de maior sensibilidade também é importante para afastar hipóteses em fase inicial do diagnóstico, uma vez que existirão poucos falsos negativos.

Testes específicos são úteis para confirmar um diagnóstico. Isso ocorre porque um teste altamente específico raramente é positivo na ausência de doença, o que resulta em poucos falso-positivos. Logo, se o resultado do teste for positivo, há grande probabilidade de ser um verdadeiro positivo. Testes altamente específicos são particularmente necessários quando os resultados falso-positivos podem prejudicar o paciente fisicamente, emocionalmente ou financeiramente. Por exemplo, antes que os pacientes sejam submetidos a tratamentos com altos efeitos colaterais, exames altamente específicos, como a biópsia, devem ser realizados.

O *trade-off* entre a sensibilidade e a especificidade

O ideal seria termos acesso a um teste que fosse simultaneamente altamente sensível e específico, o que raramente é possível. Na maioria dos casos, ao se aumentar a sensibilidade, perde-se em especificidade. Nessas situações, a localização de um ponto de corte no *continuum* entre normal (sadio) e anormal (doente) se baseia em evidências clínicas, e num equilíbrio entre a sensibilidade e a especificidade.

Ao contrário do que se pensa, sensibilidade e especificidade não são fixas dos testes: a escolha do ponto de corte pode alterar a especificidade e sensibilidade. Ao se estabelecer um ponto de corte baixo (quando o aumento significa piora da condição clínica), o teste terá maior sensibilidade, mas baixa especificidade, com muitos falsos positivos. Por outro lado, com o ponto de corte alto, o teste será mais específico, mas com menor sensibilidade, tendo mais falsos negativos. O limite intermediário para o ponto de corte, implicará em um menor número de falsos positivos e negativos com um menor de erro de classificação (Figura 1.2.3).

Seguindo no exemplo do Box 1.2.1, a Figura 1.2.3 apresenta o ponto de corte do BNP com menor erro de classificação de doentes e sadios. Pode-se observar, na curva normal da direita, a distribuição hipotética da dosagem do BNP dos indivíduos com diagnóstico de ICFER e na cur-

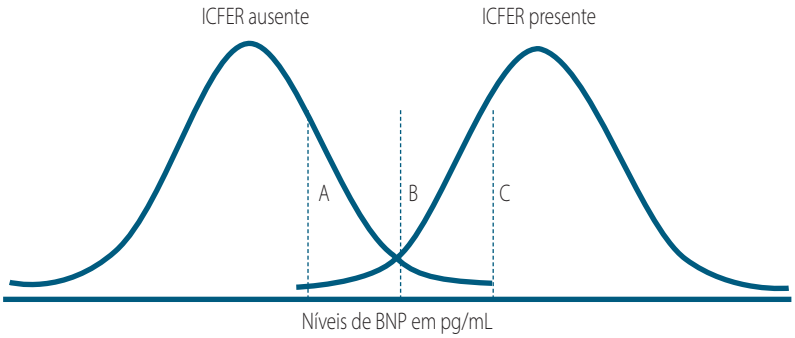


FIGURA 1.2.3 – Ponto de corte do BNP para o diagnóstico de ICFER, com menor erro de classificação de indivíduos doentes e sadios.

va normal da esquerda, sem o diagnóstico. Se for adotado o menor ponto de corte de BNP que um indivíduo doente apresentou, 10,5 pg/mL (linha A pontilhada), será alcançado uma alta sensibilidade (dados apresentados à frente na Tabela 1.2.2), mas haveria um grande percentual de indivíduos sadios que seriam considerados doentes (toda a área à direita da linha pontilhada A, abaixo da curva representando ICFER ausente). Por outro lado, se for adotado o ponto de corte mais alto que um indivíduo sadio apresentou, 282,00 pg/mL (linha pontilhada B), haveria um alto percentual de doentes considerados sadios, com BNP abaixo desse valor (toda a área à esquerda da linha pontilhada C, abaixo da curva representando ICFER presente). A opção é sempre por escolher um ponto de corte que minimize tanto o percentual de falsos positivos, quanto o de falsos negativos.

Evidentemente, em situações específicas, o profissional de saúde pode buscar minimizar o percentual de falsos negativos, ou seja, não deseja liberar indivíduos doentes. Nesse momento, poderá adotar um ponto de corte menor, aumentando a sensibilidade e diminuindo os falsos negativos. Em outro momento, o interesse pode ser não incluir indivíduos sadios, ou seja, o profissional de saúde deseja excluir ao máximo os falsos positivos. Nesse caso, o ponto de corte adotado deve ser mais alto.

Interpretando o resultado de um teste qualitativo: a importância da probabilidade pré-teste

A primeira etapa, antes de solicitar um teste diagnóstico (que é um exame complementar), consiste na avaliação clínica do paciente para estabelecer possíveis hipóteses diagnósticas: queixa principal, história clínica, considerando exposição a possíveis fatores de risco, e o exame físico. A acurácia dessa etapa dependerá da experiência do profissional de saúde, que o faz, na maioria das vezes, de modo implícito. À medida que a avaliação clínica avança, hipóteses diagnósticas se fortalecem ou enfraquecem. Ao final da avaliação clínica, o profissional de saúde atribui uma probabilidade para cada um dos possíveis diagnósticos que buscará confirmar.

Essa probabilidade é chamada de probabilidade pré-teste é estimada por quem solicita o teste diagnóstico. A probabilidade pré-teste pode ser vista como a prevalência esperada da doença em um grupo de pessoas com as mesmas características do paciente para quem se solicita o teste diagnóstico complementar e será tão mais precisa quanto mais criteriosa for a avaliação clínica. Após a incorporação do resultado do teste diagnóstico solicitado, a probabilidade pré-teste passa ser a probabilidade pós-teste. As duas pro-

babilidades, pré e pós-teste, expressam a probabilidade da presença da doença antes e depois da inclusão de um determinado teste diagnóstico. A grande questão envolvendo a probabilidade pré-teste é que quanto maior ela for, maior será a probabilidade de um teste positivo ser verdadeiro e maior será a confiança que o profissional de saúde pode depositar num resultado positivo.

Como vimos acima, se o teste em avaliação for positivo e o indivíduo estiver doente, será um verdadeiro positivo (VP). Caso contrário, será um falso positivo (FP). O mesmo raciocínio se repete para o teste negativo: será verdadeiro negativo (VN) se o indivíduo for sadio e falso negativo (FN) se for doente.

Os Quadros 1.2.3 A e B apresentam exemplos para duas situações hipotéticas. Na situação **A** o paciente tem uma queixa e o profissional de saúde levanta imediatamente algumas hipóteses diagnósticas e orienta sua avaliação clínica. A história, o exame físico e os fatores de risco são compatíveis com o diagnóstico **X.** O profissional de saúde solicita um exame complementar e a partir da avaliação clínica, atribui a probabilidade de 70% para o diagnóstico **X.** O teste escolhido tem 90% de sensibilidade e 90% de especificidade. Imaginemos que 100 pessoas com a mesma condição clínica de seu paciente sejam examinadas. Segundo a estimativa dos 100 pacientes, 70 terão a doença X e 30 não terão. Aplicando a sensibilidade e especificidade de 90% chegamos ao Quadro 1.2.3.

Do total de indivíduos examinados 70% tem a doença (probabilidade pré-teste). Dos 100 indivíduos examinados, 66 tiveram o teste em avaliação positivo. Do total de testes positivos, 63 (95,45%) eram doentes. Assim, o valor preditivo positivo (VPP) deste teste é de 95,45%. O valor preditivo positivo corresponde a probabilidade de um teste positivo ser de um doente. Também é chamada de probabilidade pós teste positiva.

Na situação **B,** o profissional de saúde acredita que a probabilidade daquele paciente estar com o a doença **X** é muito baixa pela ausência de fatores de risco e por alguns elementos da história clínica. Ele atribui uma probabilidade de 10%. Segundo a estimativa dos 100 pacientes, 10 terão a doença **X** e 90 não terão. Aplicando a sensibilidade e especificidade de 90% obtém-se o quadro ilustrado acima na opção B.

Do total de indivíduos examinados 10% tem a doença (probabilidade pré-teste). Dos 100 indi-

Quadro 1.2.3. Relação entre o resultado de um teste diagnóstico e a ocorrência da doença. A) Relação com probabilidade pré-teste ou prevalência de 70%. B) Relação com probabilidade pré-teste ou prevalência de 10%

	Teste ouro		
	Doente	**Sadio**	**Total do teste em avaliação**
A			
Teste em avaliação (+)	63% verdadeiro positivos	3% falso positivos	66% positivos
Teste e avaliação (–)	7% falso negativos	27% verdadeiro negativos	34% negativos
Total do teste ouro	**70%**	**30%**	**100%**
B			
Teste em avaliação (+)	9% verdadeiro positivos	9% falso positivos	18% positivos
Teste e avaliação (–)	1% falso negativos	81% verdadeiro negativos	82% negativos
Total do teste ouro	**10%**	**90%**	**100%**

víduos examinados, 18 tiveram o teste em avaliação positivo. E dentro do total de testes positivos, 9 (50%) eram doentes. Assim o valor preditivo positivo (VPP) do teste passou a ser de 50%. O valor preditivo positivo traduz a probabilidade de o teste positivo ser de um doente ou a proporção de pacientes com resultado do teste positivo que foram corretamente diagnosticados.

VPP = verdadeiro positivo/ total de resultados positivos

VPP situação A = 63/66 = 95,45%

VPP situação B = 9/18 = 50%

Em outros momentos, é necessário afastar um diagnóstico. Nesse caso, o raciocínio é o inverso. O interesse está na probabilidade de um resultado negativo ser de um indivíduo sadio. Essa probabilidade é chamada de valor preditivo negativo (VPN). Examinando os quadros das situações **A** e **B,** observa-se que o VPN será maior na situação B, onde a probabilidade pré-teste é muito baixa.

Na situação **A**, dos 34 testes com resultados negativos, 27 (79,4%) eram verdadeiros negativos. Já na situação **B**, dos 82 testes com resultados negativos, 81 (98,78%) eram de um indivíduo sadio. O valor preditivo negativo (VPN) traduz a probabilidade de o teste negativo ser de um sadio ou a proporção de pacientes com resultado do teste negativo que foram corretamente diagnosticados. O VPN é igual 1- probabilidade pós-teste negativa, que veremos a seguir.

VPN = verdadeiro negativo/ total de resultados negativos

VPN situação A = 27/34 = 79,4%

VPN situação B = 81/82 = 98,78%

Para mostrar mais claramente a influência da probabilidade pré-teste na interpretação do resultado de um teste diagnóstico, observe a Figura 1.2.4.

Nas três situações a sensibilidade e a especificidade é a mesma, de 50%. Na primeira, a probabilidade pré-teste é de 50%, havendo a mesma proporção de VP, FP, FN e VP. Na situação B, a probabilidade pré-teste é baixa e observa-se que a probabilidade de um resultado positivo ser de um doente (VP) também é baixa. Já a probabilidade de um resultado negativo ser de um indivíduo sadio (VN) é alta. Na situação C a probabilidade pré-teste é alta e a probabilidade de um resultado positivo ser de um doente (VP) também é alta. Já a probabilidade de um resultado negativo ser de um indivíduo sadio (VN) é baixa. Podemos concluir que quanto maior a probabilidade pré-teste, maior o VPP e quanto menor a probabilidade pré-teste, maior o VPN.

Há autores que estabelecem uma regra de ponto de corte para iniciar um tratamento. Se a probabilidade pré-teste for inferior a 25%, não faça novo teste e não trate. Nessa situação, o ideal é acompanhar a evolução do paciente. Se a probabilidade pré-teste for superior a 65%, não faça novo exame e trate. Diante de valores de pré-teste maiores que 25% e inferiores a 65%, solicite o exame (Sacket et al, 2003). Essa regra não se aplica para testes feitos para avaliar a gravidade de uma doença, ou para ter parâmetros para o acompanhamento do paciente. É uma regra exclusivamente para o diagnóstico.

Para ilustrar a relação entre a probabilidade pré-teste e o valor preditivo positivo (ou probabilidade pós teste positiva) e valor preditivo negativo (1-probabilidade pós-teste negativa)

A)
Probabilidade pré-teste = 50%

VP	**FP**
FN	**VN**

B)
Probabilidade pré-teste < 50%

VP	**FP**
FN	**VN**

C)
Probabilidade pré-teste > 50%

VP	**FP**
FN	**VN**

FIGURA 1.2.4 – Relação entre a probabilidade pré-teste (prevalência) e interpretação do resultado do teste.

apresentamos dois gráficos (Figura 1.2.5 A e B) para diferentes níveis de sensibilidade (S) e especificidade (E), que nesse exemplo, são iguais.

Em **A** observa-se que à medida que a probabilidade pré-teste aumenta, há um aumento na probabilidade pós-teste positiva, ou valor predito positivo (VPP). Quanto menor a sensibilidade ou especificidade, maior é essa correlação. Em **B** observa-se a relação inversa, ou seja, quanto menor a probabilidade pré-teste, maior a probabilidade pós-teste negativa (1-VPP) e essa correlação também é maior quanto menor a sensibilidade ou especificidade.

Do exposto conclui-se que um novo teste diagnóstico acrescenta informação àquela que os testes realizados anteriormente forneceram.

Da observação sobre a relação entre a probabilidade pré-teste e os valores preditivos positivos, três questões precisam ser enfatizadas:

1. Antes de se solicitar um teste é importante estimar a probabilidade do paciente ter o diagnóstico para o qual o teste está sendo pedido;
2. O teste será mais útil quando o profissional de saúde está em dúvida, quando a probabilidade pré-teste gira em torno de 50%;
3. Ao se aplicar os valores do VPP e VPN em estudos publicados é essencial comparar o perfil dos pacientes incluídos nessas pesquisas e o perfil dos pacientes que estão sendo avaliados pelo profissional de saúde.

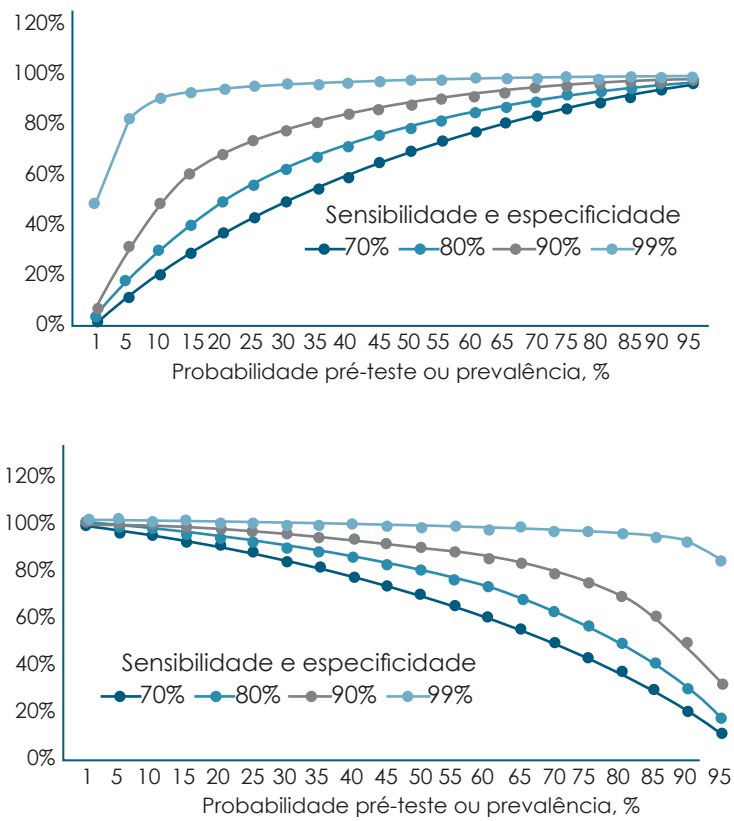


FIGURA 1.2.5 – Correlação entre a probabilidade pré-teste (ou prevalência). **A)** Probabilidade pós-teste positiva (ou VPP); **B)** Probabilidade pós-teste negativa (ou VPN).

Relação entre sensibilidade, especificidade, probabilidade pré-teste VPP e VPN

Vimos que a sensibilidade e a especificidade são propriedades do teste que não se alteram com a probabilidade pré-teste. Vimos também que quanto maior a sensibilidade, menor percentual de falsos negativos e maior o valor preditivo negativo. E ainda que, quanto maior a especificidade, menor o percentual de falsos positivos e maior o valor preditivo positivo.

Por outro lado, mostramos que quanto mais alta a prevalência ou probabilidade pré-teste, menor o percentual de falsos positivos e por isso, maior o VPP. Mostramos também que quanto mais baixa a prevalência ou probabilidade pré-teste menor o percentual de falsos negativos e, consequentemente, o VPN.

Vamos rever o exemplo descrito no Box 1.2.1, cujos dados para o ponto de corte do BNP de ≥ 41,5 pg/mL para o diagnóstico de ICFER foram apresentados na Figura 1.2.3. Revendo os resultados:

- Probabilidade pré-teste = 4,6%
- Sensibilidade = 79,3%
- Especificidade = 86,1%
- VPP = 21,5%
- VPN = 98,9%

Observamos que a sensibilidade e a especificidade têm valores próximos, o que é mais comum de ser encontrado, no entanto, os valores de VPP e VPN são muito diferentes: 21,5% e 98,9% respectivamente. O que provocou tal diferença? A prevalência da ICFER que foi de 4,6%.

A prevalência de uma condição numa população, ou a probabilidade pré-teste (probabilidade atribuída a uma condição frente a um paciente, história e exame clínico) pode variar muito mais que a sensibilidade em relação a especificidade. Por isso, a influência da probabilidade pré-teste na interpretação de um resultado é mais importante.

No caso do exemplo do Quadro 1.2.2, um BNP ≥ 41,5 pg/mL tem a probabilidade de 21,5% de ser de um indivíduo doente (havia 78,5% de falsos positivos). Já um resultado inferior a esse ponto de corte, tem a probabilidade de quase 99% ser de um indivíduo sadio.

Avaliando a acurácia de um teste independentemente da probabilidade pré-teste

A razão de verossimilhança é outra foram para avaliar o desempenho de um teste diagnostico. A vantagem da sua utilização é que combina a sensibilidade e a especificidade e não depende dos valores da probabilidade pré-teste. A razão de verossimilhança pode ser positiva ou negativa e descreve a capacidade do teste de discriminar doentes de sadios.

A razão de verossimilhança, tanto a positiva quanto a negativa, varia de zero a infinito. Trata-se de uma chance pois ambas relacionam duas probabilidades (P): sucesso (P) e fracasso (1-P). Nos dois casos, a razão expressa a chance de o resultado ser de um indivíduo doente (P) em relação a um indivíduo sadio (1-P).

A razão de verossimilhança positiva (RV+) expressa a chance de um resultado positivo ser de um doente em relação a um sadio.

RV+ = sensibilidade/1-especificidade

Quando a RV+ é igual a 1, a chance de um resultado positivo ser de um indivíduo doente é a mesma de ser de um indivíduo sadio. Quanto maior a RV+ maior certeza de que um resultado positivo é de um doente e mais informativo é o teste.

A razão de verossimilhança negativa (RV-) expressa a chance de um resultado negativo ser de um doente em relação a um sadio.

RV- = sensibilidade/ especificidade

Quanto menor a RV- maior certeza de que um resultado negativo é de um indivíduo sadio.

Problemas com a interpretação da razão de verossimilhança

Os conceitos de chance e de razão de verossimilhança não são intuitivos. Na verdade, traduzimos mentalmente a chance como probabilidade, o que não é verdade. No exemplo descrito no Box 1.2.1, a probabilidade de se encontrar ICFER na população geral foi de 4,6% (29/633), mas a chance foi um pouco maior: 4,8% (29/ (633-29). Quanto mais elevada a prevalência ou probabilidade pré-teste, maior é a diferença entre a chance e a probabilidade.

Uma RV+ de 10 não significa que a probabilidade da presença da doença é 10 vezes a probabilidade de sua ausência, mas sim que a chance é 10 vezes maior.

Como avaliar as razões de verossimilhança

Retomando o exemplo do Box 1.2.1, apresentado no Quadro 1.2.2, a RV+ do BNP ≥ 41,5 pg/mL foi igual a 0,793/ 1-0,860 = 5,67. Ou seja, a chance de um resultado de BNP ≥ 41,5 pg/mL ser de um indivíduo doente é 5,67 vezes maior que a chance de ser um indivíduo saudável. Já a RV– foi igual a 1-0,793/ 0,860 = 0,25, assim, a chance de um resultado de BNP < 41,5 pg/mL ser de um doente é de ¼ (0,25) desse resultado ser de um sadio.

Há alguns pontos de corte propostos para a avaliação das RV (Sackett *et al.*, 2003):

- RV+ com valores entre > 1 e 2: pouca utilidade para a confirmação de um diagnóstico;
- RV+ com valores entre > 2 e 10: utilidade moderada para a confirmação de um diagnóstico;
- RV+ > 10: utilidade forte para a confirmação de um diagnóstico;
- RV– entre 0,5 e 1: pouca ou nenhuma utilidade para descartar um diagnóstico;
- RV– entre < 0,5 e 0,1: utilidade moderada para descartar um diagnóstico;
- RV– < 0,1: utilidade forte para descartar um diagnóstico.

Voltando ao exemplo. A RV+ do BNP no ponto de corte de ≥ 41,5 pg/mL foi de 5,67, tendo uma utilidade moderada para confirmar o diagnóstico de ICFER. A RV - do BNP no ponto de corte de < 41,5 pg/mL foi de 0,25, tendo também uma utilidade moderada para descartar o diagnóstico de ICFER.

A razão de verossimilhança para testes quantitativos

Quando se tem acesso aos resultados de um teste em sua forma contínua pode-se estimar a RV tanto positiva, quanto negativa, a cada ponto de corte. No tópico seguinte está descrito como se estima a sensibilidade e a especificidade para os diferentes pontos de corte de um teste diagnóstico (curva ROC) em sua forma contínua. Feito isso, basta aplicar as fórmulas descrita no tópico anterior e se chega aos valores das RV.

Na Tabela 1.2.1 são apresentadas as dosagens de BNP e as respectivas RV+ e RV– para o diagnóstico de ICFER, a partir dos dados do ESTUDO DIGITALIS (Box 1.2.1).

Para esse cálculo utilizou-se os dados da sensibilidade e especificidade a cada ponto de BNP, apresentados na Tabela 1.2.2. Observa-se que a cada ponto de BNP a RV+ aumenta. Isso significa que à medida que o BNP aumenta, maior é a chance de o resultado ser de um indivíduo doente. O menor ponto de BNP descrito foi de 10,5 pg/mL, quando a chance deste resultado ser de um indivíduo doente foi de 1,28, muito próxima da unidade que implica em iguais chances do resultado ser de um indivíduo doente ou sadio. Com o BNP de 20,5 pg/mL a chance de o resultado ser de um doente foi o dobro de ser de um sadio (RV+ = 2), já com o BNP de 30,5 pg/mL a chance foi o triplo (RV+ ≅ 3) e assim por diante, chegando a uma chance maior que 100, para BNP de 279,5 pg/mL.

Como vimos acima, RV+ acima de 10 são considerados altas e praticamente, confirmariam o diagnóstico. No exemplo descrito, isso equivale a um BNP de 60,5 pg/mL ou mais.

Tabela 1.2.1. Razão de verossimilhança positiva e negativa para diferentes pontos de corte do BNP

BNP (pg/mL)	RV+	RV–
10,5	1,28	0,42
20,5	2,12	0,33
30,5	3,54	0,27
40,5	5,51	0,24
41,5	5,71	0,24
42,5	5,98	0,28
51,0	7,70	0,30
60,5	10,18	0,40
70,0	10,46	0,44
82,5	9,66	0,54
90,0	11,23	0,54
100,0	11,83	0,61
110,0	16,48	0,64
123,0	17,25	0,67
130,0	15,50	0,70
201,0	30,13	0,77
279,5	103,50	0,79
> 279,5	0	1

Fonte: ESTUDO DIGITALIS

Tabela 1.2.2. Sensibilidade e especificidade para diferentes pontos de corte do BNP plasmático

BNP (pg/mL)	Sensibilidade	Especificidade	1-Especificidade
10,5	0,862	0,328	0,672
20,5	0,793	0,626	0,374
30,5	0,793	0,776	0,224
40,5	0,793	0,856	0,144
41,5	0,793	0,861	0,139
42,5	0,759	0,873	0,127
51,0	0,724	0,906	0,094
60,5	0,621	0,939	0,061
70,0	0,586	0,944	0,056
82,5	0,483	0,950	0,05
90,0	0,483	0,957	0,043
100,0	0,414	0,965	0,035
110,0	0,379	0,977	0,023
123,0	0,345	0,980	0,02
130,0	0,310	0,980	0,02
201,0	0,241	0,992	0,008
279,5	0,207	0,998	0,002

Fonte: ESTUDO DIGITALIS.

Os valores de RV-, ainda no nosso exemplo, oscilaram entre 0,42 e 0,79 não contribuindo muito para a exclusão na presença de ICFER nesses pacientes.

Da razão de verossimilhança à probabilidade pós-teste

Como dissemos acima, depois de incorporada à informação do resultado do teste diagnóstico solicitado, a probabilidade de o paciente ter a doença se modifica. Essa probabilidade é chamada de probabilidade pós-teste. A probabilidade pós-teste também pode ser positiva ou negativa. Como também já comentado, uma vez incorporada, essa probabilidade deve ser considerada a probabilidade pré-teste para os próximos passos.

Há algumas maneiras de se chegar à probabilidade pós-teste, uma delas, inclusive, não incluindo em seu cálculo, a probabilidade pré-teste.

■ Da razão de verossimilhança à probabilidade pós-teste, incorporando a probabilidade pré-teste no cálculo

Há uma sequência de cálculos que partem da probabilidade pré-teste para se chegar à probabilidade pós-teste.

1) Transforme a probabilidade pré-teste em chance pré-teste:

Chance pré-teste = probabilidade pré-teste/(1- probabilidade pré-teste)

2) Transforme a chance pré-teste em chance pós-teste, positiva ou negativa:

Chance pós-teste positiva = chance pré-teste × RV+

Chance pós-teste negativa = chance pré-teste × RV–

3) Transforme a chance pós-teste em probabilidade pós-teste:

Probabilidade pós-teste = chance pós-teste/ 1 + chance pós-teste

No caso do teste quantitativo a probabilidade pré-teste ou prevalência deve ser calculada a cada ponto de corte. E é essa probabilidade pré-teste a cada ponto de corte que será utilizada para os cálculos subsequentes.

A Tabela 1.2.3 apresenta os dados do ESTUDO DIGITALIS (Box 1.2.1), incorporando as RV+ e RV– apresentadas na Tabela 1.2.1.

Vale notar que a probabilidade pós-teste positiva tem o mesmo valor do valor preditivo positivo e a probabilidade pós-teste negativa tem o mesmo valor de 1 – valor preditivo negativo.

Da razão de verossimilhança à probabilidade pós-teste: avaliando através de um nomograma

A segunda forma é utilizando o nomograma de Fagen (Figura 1.2.6). A ideia geral de um no-

Tabela 1.2.3. Probabilidade pré-teste e probabilidade pós-teste positiva e negativa para diferentes pontos de corte do BNP plasmático. sensibilidade e especificidade para diferentes pontos de corte do BNP plasmático

BNP (pg/mL)	Probabilidade pré-teste	Probabilidade pós-teste	
		Positiva	Negativa
10,5	0.632	0.809	0.265
20,5	0.948	1.989	0.312
30,5	0.948	3.277	0.252
40,5	0.948	5.006	0.229
41,5	0.948	5.177	0.227
42,5	1.106	6.264	0.304
51,0	1.264	8.974	0.384
60,5	1.738	15.257	0.697
70,0	1.264	16.820	0.825
82,5	1.738	18.993	1.273
90,0	1.896	21.423	1.264
100,0	2.370	24.610	1.605
110,0	2.844	32.537	1.967
123,0	3.002	34.802	2.176
130,0	3.160	33.586	2.590
201,0	3.476	52.031	2.806
279,5	3.633	79.602	2.801

Fonte: ESTUDO DIGITALIS.

mograma é ter escalas de 3 variáveis para permitir que traçando uma reta entre o 1º e o 2º valor, se chegue ao 3º valor, sem necessidade de cálculos. Em 1975, em uma carta ao *New England Journal*, Dr Fagan propôs o uso do nomograma, que utiliza o teorema de Bayes. A partir da probabilidade pré-teste (1ª variável) e da RV+ (2ª variável) se chegaria a probabilidade pós-teste, ou seja, qual a probabilidade do teste ser de um doente, dada a probabilidade pré-teste e à RV.

A Figura 1.2.7 apresenta o nomograma de Fagan utilizando o exemplo do Box 1.2.1 e dados das Tabelas 1.2.1 e 1.2.3. Para o ponto de corte do BNP de 70 pg/mL, a prevalência ou probabilidade pré-teste foi de 1,26% (Tabela 1.2.3) e a RV+ foi de 10,46 (Tabela 1.2.1). Ligando os dois pontos chegamos a uma probabilidade pós-teste cujo valor ficou valor próximo ao encontrado pelas fórmulas: 16,82 (Tabela 1.2.3).

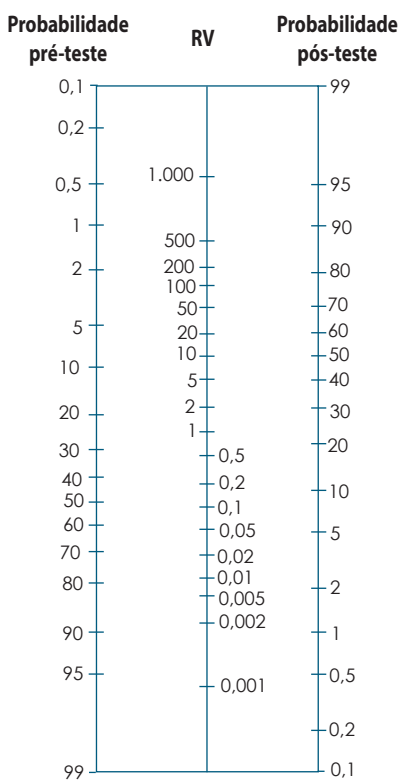


FIGURA 1.2.6 – Nomograma de Fagan.

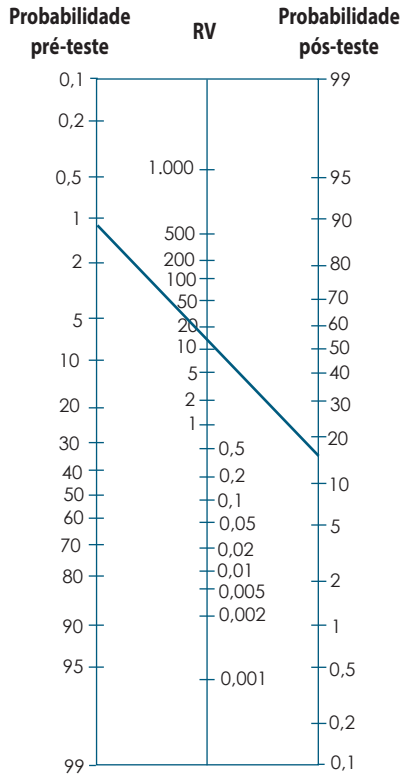


FIGURA 1.2.7 – Nomograma de Fagan para ponto de corte de BNP de 70 pg/mL para diagnóstico de ICFER, ESTUDO DIGITALIS.

■ Da razão de verossimilhança à probabilidade pós-teste: avaliando por uma tabela calculada

Uma alternativa a esses cálculos é o uso de pontos de corte proposta por McGee (2002). Utilizando a RV+ ele estima as mudanças que devem ocorrer na probabilidade pós-teste, como exposto no Quadro 1.2.4.

Avaliando a acurácia de um teste quantitativo: a curva ROC

Quando os testes são quantitativos o resultado não é apresentado como positivo ou negativo, ou ainda reagente ou não regente, mas é apresentado o valor encontrado. De um modo geral, o resultado vem acompanhado de valores

Quadro 1.2.4. Relação entre a verossimilhança e a probabilidade pós-teste	
Razão de verossimilhança	**Mudança aproximada na probabilidade da doença (%)**
Valores entre 0 e 1 diminuem a probabilidade da doença	
0,1	-45
0,2	-30
0,3	-25
0,4	-20
0,5	-15
1	0
Valores maiores de 1 aumentam a probabilidade da doença	
2	+15
3	+20
4	+25
5	+30
6	+35
7	
8	+40
9	
10	+45

Fonte: McGee, 2002.

de referência, ou seja, um limite considerado "normal".

Uma das técnicas utilizadas para se a chegar a um valor acima do qual há risco de uma lesão, ou o prognóstico não é bom, é a construção de uma curva chamada curva ROC. Em português seria curva COR (característica de operação do receptor). Ela é construída a partir das características do teste nos diferentes pontos de corte.

Ainda utilizando os dados do ESTUDO DIGITALIS, construiu-se a Tabela 1.2.2 com diferentes dosagens de BNP e a cada valor de BNP calculou-se a sensibilidade e a especificidade para o diagnóstico de ICFER, tendo como padrão ouro o EDT associado à presença de sinais e sintomas da doença. Para efeitos didáticos são apresentados alguns valores.

Para o cálculo da sensibilidade considera-se como verdadeiros positivos os doentes com valores de BNP iguais ou maiores do que o ponto de corte em questão. Por exemplo, para o cálculo da sensibilidade no ponto de corte de 10,5 pg/ml, considera-se todos os casos com valores iguais ou maiores que 10,5 pg/mL e isso é repetido a cada ponto de corte. Para o cálculo da especificidade, os verdadeiros negativos serão os indivíduos sadios com valores inferiores ao ponto de corte em questão. No nosso exemplo, seriam considerados verdadeiros negativos todos os sadios com valores de BNP inferiores a 10,5 pg/mL.

Observa-se que à medida que os valores de BNP aumentam, a sensibilidade diminui e a especificidade aumenta, o que fica mais claramente explicado no tópico *trade-off*, já apresentado. O objetivo da curva ROC é chegar-se ao ponto de corte que maximize as duas propriedades. Ou seja, que alcance a melhor sensibilidade com a menor perda da especificidade.

As coordenadas da curva são os valores de sensibilidade e os valores de 1-especificidade para cada dosagem do que está sendo medido. No nosso exemplo, percebe-se que a sensibilidade se manteve em 79,3% do ponto de corte de 20,5 pg/mL à 41,5 pg/mL e a especificidade aumentou nesse intervalo, passando de 62,6% a 86,1%. Após 41,5 pg/mL a sensibilidade e a especificidade se distanciam. O melhor ponto de corte de BNP para discriminar portadores de ICFER de não portadores seria de 41,5 pg/mL.

A Figura 1.2.8 apresenta a curva ROC para o exemplo acima. O melhor ponto de corte é aquele com maior sensibilidade (mais alto) e menor perda de especificidade (menor 1-especificidade) que equivale ao ponto mais próximo ao eixo da sensibilidade. É chamado de "ombro" da curva.

Quanto maior a área sob a curva, melhor a acurácia do teste como um todo. A curva do teste ideal acompanharia o eixo da sensibilidade, com uma área correspondente a 1. A cada ponto de corte a sensibilidade cresceria, chegando

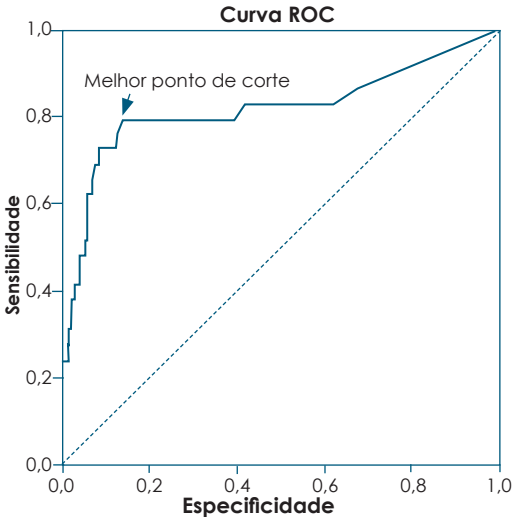


FIGURA 1.2.8 – Coordenadas para a curva ROC para avaliação do melhor ponto de corte para o BNP plasmático no diagnóstico de ICFER. Fonte: ESTUDO DIGITALIS.

a 100%, sem perda da especificidade. O oposto seria um teste no qual cada aumento da sensibilidade corresponderia a uma perda proporcional na especificidade, representado pela linha tracejada no gráfico. No nosso exemplo, a área sob a curva é de 0,818 com um intervalo de confiança de 95% de 0,710 a 0,926.

Curvas ROC são formas particularmente valiosas para comparar diferentes testes em um mesmo diagnóstico.

Testes múltiplos

A maioria dos testes diagnósticos não tem uma acurácia que permita o profissional de saúde utilizá-lo como um recurso definitivo para confirmação ou exclusão de uma hipótese diagnóstica. Quando o diagnóstico pode levar a uma intervenção invasiva, ou ainda em situações que o diagnóstico se refere a uma doença muito grave, ou que pode representar uma carga muito grande para o paciente, não é satisfatório recorrer-se a exames com sensibilidade e especificidade que não sejam próximas a 100%. Nessas situações a conduta é seguir na investigação, adotando a estratégia dos testes múltiplos. Há duas abordagens principais para a realização de testes múltiplos: realizá-los em paralelo ou de forma sequencial. Nas duas abordagens, a combinação dos resultados muda a sensibilidade e a especificidade dos testes iniciais.

Os testes em paralelo são utilizados principalmente em situações de emergência, cuja intervenção precoce seja definitiva para a sobrevida do paciente. Para apresentação e discussão dos testes em paralelo, utilizamos o exemplo do diagnóstico de infarto agudo do miocárdio (IAM) na emergência, a partir dos marcadores bioquímicos de necrose miocárdica. As Diretrizes da Sociedade Brasileira de Cardiologia sobre Angina Instável e Infarto Agudo do Miocárdio sem Supradesnível do Segmento ST (II Edição, 2007 – Atualizado em 2013) recomendam que os marcadores bioquímicos de necrose miocárdica devem ser mensurados em todos os pacientes com suspeita de IAM, na admissão e repetidos pelo menos uma vez 6 a 9 horas após o início dos sintomas, caso a primeira dosagem seja normal ou discretamente elevada. Os testes de escolha são o CK-MB massa e as troponinas. Do ponto de vista de marcadores bioquímicos de necrose miocárdica, o diagnóstico de IAM deve ser feito de acordo com os seguintes critérios:

1. Troponina T ou I: aumento acima do percentil 99 em pelo menos uma ocasião nas primeiras 24 horas de evolução;
2. Valor máximo de CK-MB, preferencialmente massa, maior do que o limite superior da normalidade em duas amostras sucessivas; valor máximo de CK-MB acima de duas vezes o limite máximo da normalidade em uma ocasião durante as primeiras horas após o evento.

Seguiremos tomando como exemplo a troponina I cardíaca TnIc e a CK-MB massa acima de duas vezes o limite máximo da normalidade.

Por que a Diretriz recomenda duas alternativas para o diagnóstico, alteração da Troponina ou da CK-MB? Como dissemos acima, a sen-

sibilidade e a especificidade destes métodos não são de 100%. A sensibilidade da TnIc gira em torno de 90% (Christenson *et al.*, 1999) e a sensibilidade CK-MB massa varia em torno de 97%. As especificidades são aproximadamente de 97% (TnIc) e 90% (CKMB massa).

Com a sensibilidade de 97% (CKMB massa), a cada 100 exames, 3 são falsos negativos, ou seja, pacientes que seriam liberados apesar de terem um IAM. A situação diante do TnIc é mais grave, a cada 100 exames, 10 casos de IAM seriam liberados.

Daí a proposição de realizar dois testes em paralelo, simultaneamente. O diagnóstico é considerado positivo se um dos exames for positivo. Ao se adotar essa estratégia a sensibilidade resultante aumenta na medida em que os resultados positivos dos dois testes são combinados: (sensibilidade do TnIc) + (sensibilidade do CKMB massa) – (sensibilidade comum aos dois exames). Estima-se a sensibilidade comum aos dois testes multiplicando-se as sensibilidades individuais (Figura 1.2.9).

Sensibilidade de testes em paralelo = (sensibilidade teste A + sensibilidade teste B) – (sensibilidade teste A × sensibilidade teste B)

No nosso exemplo teríamos: Sensibilidade testes em paralelo = 0,90 + 0,97 – (0,90 × 0,97) = 0,997%.

Assim, com a estratégia do teste em paralelo em que será considerado doente aquele que apresentar um ou outro teste positivo, a sensibilidade de 90% ou 97% passaram a ser de 99,7%. Com a estratégia do teste em paralelo, a especificidade diminui porque seriam considerados negativos somente os casos em que os dois testes forem negativos. Ou seja, a especificidade combinada seria a especificidade de um teste, multiplicada pela especificidade do outro teste (Figura 1.2.10).

Especificidade de testes em paralelo = (especificidade teste A × especificidade teste B)

No nosso exemplo teríamos: Especificidade testes em paralelo = 0,97 × 0,90 = 0,873%.

A especificidade diminuiu o que implicaria um maior percentual de falsos positivos, ou seja, 13 indivíduos sadios, a cada 100 examinados, seriam tratados. A escolha da estratégia de testes em paralelo pelos cardiologistas para o início da abordagem ao paciente com suspeita de IAM se justifica pela gravidade desta condição e pelo benefício de uma intervenção precoce.

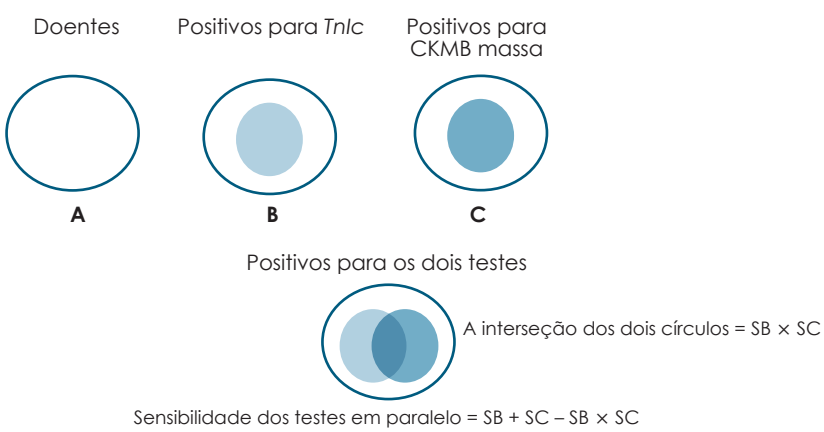


FIGURA 1.2.9– Combinação da sensibilidade de dois testes em paralelo.

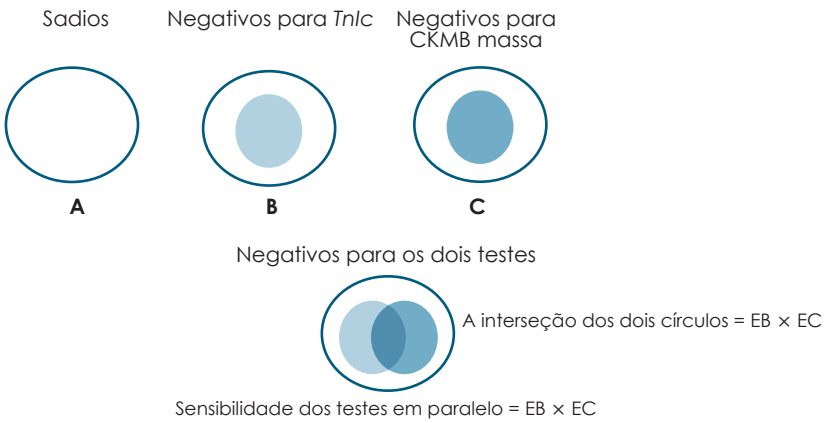


FIGURA 1.2.10 – Combinação da especificidade de dois testes em paralelo.

Os testes em série são utilizados principalmente em situações ambulatoriais, nas quais o diagnóstico pode ser feito em um intervalo maior de tempo e naquelas em que a sequência da abordagem da condição implica em um teste invasivo ou em um tratamento com possíveis efeitos colaterais, ou ainda, quando a simples comunicação ao paciente da positividade pode trazer danos, como é o caso de doenças de grande letalidade ou estigma.

Utilizaremos o exemplo do diagnóstico da sorológico da infecção por HIV. Desde o início da epidemia do HIV, o Ministério da Saúde preconiza que o diagnóstico seja realizado com pelo menos dois testes, um para triagem e um segundo, mais específico, para confirmar o resultado da triagem (Ministério da Saúde, 2016). Assim o exame só será considerado reagente se os dois testes forem positivos. O segundo teste só é realizado se o primeiro for positivo. A estratégia proposta tem o objetivo de aumentar o valor preditivo positivo (VPP) de um resultado reagente no teste inicial. Sabemos que para aumentar o VPP é necessário aumentar a especificidade do teste, diminuindo o percentual de falsos positivos. A estratégia de testes em série aumenta a especificidade dos testes empregados. Após o primeiro teste, que deve ser o mais sensível para diminuir os falsos negativos, os indivíduos cujo resultado foi negativo são liberados. É importante ressaltar que a sensibilidade para o diagnóstico da infecção pelo HIV é igual à sensibilidade do primeiro ensaio utilizado. Somente os positivos (verdadeiros e falsos) serão retestados com um teste com maior especificidade para reduzir os falsos positivos. Tal estratégia implica na combinação das especificidades.

Especificidade de testes em série = (Especificidade teste A + Especificidade teste B) – (Especificidade teste A × Especificidade teste B)

No fluxograma apresentado na Figura 1.2.11, o teste de imunoensaio de 4ª geração (IE4ªG) é o teste de triagem, e um teste molecular (TM) é o complementar para amostras reagentes na triagem. Tanto a sensibilidade quanto a especificidade destes testes são superiores a 99%. A cada 100 exames realizados há um falso negativo e um falso positivo. Com a estratégia dos testes em paralelo o objetivo é diminuir ainda mais os resultados falsos positivos. Assim, a especificidade combinada dos testes em série, seria: Especificidade dos testes em série = 0,99+0,99-0,9=1

Com a estratégia dos testes em paralelo, a especificidade que era de 99% para cada um dos testes, chega a 100%.

Nota: na persistência de casos negativos, persistindo a suspeita de infecção pelo HIV, uma nova amostra deverá ser coletada 30 dias após a data da coleta desta amostra devido à janela imunológica. Coletar uma segunda amostra para repetir IE 4ªG para concluir o resultado.

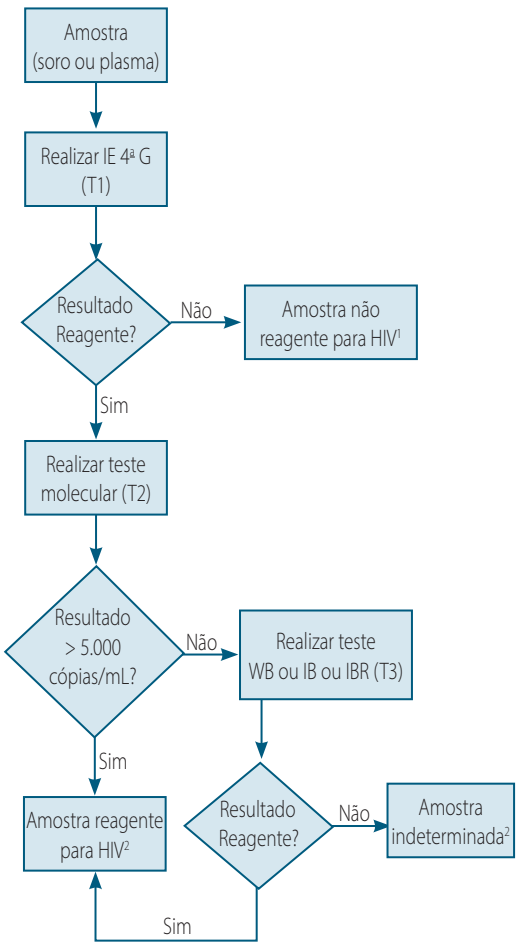
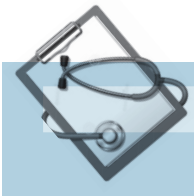


FIGURA 1.2.11 – Fluxograma de testes em série realizados em diagnóstico de HIV.

De onde vem a informação sobre a acurácia dos testes?

O conhecimento da acurácia de um teste diagnóstico é proveniente de evidências científicas. Uma boa interpretação de artigos científicos sobre avaliação da acurácia de testes diagnósticos é de grande importância para os profissionais de saúde.

Alguns critérios são essenciais para se apreciar a validade de estudos científicos. O primeiro diz respeito à seleção dos participantes do estudo. Os indivíduos selecionados devem apresentar diferentes níveis de gravidade da doença em questão incluindo ainda aqueles com sintomatologia, mas sem a doença. Para garantir a validade dos resultados é obrigatório ainda que os profissionais que avaliam a presença da doença pelo padrão ouro não participem da avaliação do teste em estudo. Por último, quando se for avaliar a acurácia do teste, deve-se garantir que seja independente, isto é, que não integre os componentes do padrão ouro. Por exemplo, o BNP é um elemento necessário para o diagnóstico de insuficiência cardíaca com fração de ejeção preservada (ICFEP). Os parâmetros do *Ecodoppler* tecidual não são suficientes para a conclusão do diagnóstico. Logo a acurácia do BNP não pode ser avaliada para o diagnóstico de ICFER.



CASO CLÍNICO

■ Caso 1

Em um estudo, resultados de um teste para a doença X realizado pela metodologia de imunocromatografia, mostraram que das 88 amostras realizadas, 72 foram positivas, dois eram negativos e que 14 eram limítrofes. Em contraste, entre 297 indivíduos sem a doença X, 4 foram positivos, 18 foram borderline e 275 foram negativos.

1) O que encontramos se calcularmos a sensibilidade e especificidade excluindo os resultados limítrofes?
2) E se forem incluídos os resultados limítrofes considerando as seguintes hipóteses: se os resultados *borderline* foram considerados positivos ou se forem considerados negativos?
3) Qual delas você escolheria para realizar testes de *screening* em doadores de sangue?

Bibliografia consultada

1. Christenson RH, Duh SH. Evidence based approach to practice guides and decision thresholds for cardiac markers. Scand J Clin Lab Invest Suppl. 1999; 230: 90-102.
2. Fagan TJ. Nomogram for Bayes's Theorem. N Engl J Med 1975; 293:257.
3. Jorge AL, Rosa ML, Martins WA, Correia DM, Fernandes LC, Costa JA, Moscavitch SD, Jorge BA, Mesquita ET. The Prevalence of Stages of Heart Failure in Primary Care: A Population-Based Study. J Card Fail. 2016 Feb;22(2):153-7.
4. McGee S. Simplifying Likelihood Ratios. Gen Intern Med. 2002;17(8):647–50.
5. Ministério da Saúde, Secretaria de Vigilância em Saúde Departamento de DST, Aids e Hepatites Virais. Manual Técnico para o Diagnóstico da Infecção pelo HIV. 3ª edição. Brasília - DF 2016.
6. Nicolau JC, Timerman A, Marin-Neto JA, Piegas LS, Barbosa CJ, et al. Diretrizes da Sociedade Brasileira de Cardiologia sobre Angina Instável e Infarto Agudo do Miocárdio sem Supradesnível do Segmento ST (II Edição, 2007) – Atualização 2013/2014.Arq Bras Cardiol. 2014 Mar;102(3 Suppl 1):1-61.
7. Piroozmand A, Soltani B, Razavizadeh M, Matini AH, Gilasi HR, Zavareh AN, Soltani S. Comparison of the serum and salivary antibodies to detect gastric Helicobacter pylori infection in Kashan (Iran). Electron Physician. 2017 Dec 25;9(12):6129-6134.
8. Rosa Maria Luiza G, Mesquita Evandro T, Jorge Antonio José L, Correia Dayse MS, Lugon Jocemir R, Kang HC, et al. Prevalence of chronic diseases in individuals assisted by the family health program in niteroi, brazil: evaluation of selection bias and protocol. Int J Med Res Health Sci.2015;4(3):587-596.
9. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Medicina Baseada em Evidências - Prática e Ensino - 2ª Ed. ArtMed. 2003.



DISCUSSÃO DO CASO CLÍNICO

Caso 1

1) Primeira hipótese: Excluindo os resultados limítrofes:

Imunocromatografia	Doença X		Total
	Presença	Ausência	
Positivo	72	4	**76**
Negativo	2	275	**277**
Total	**74**	**279**	**353**

Sensibilidade = 72/74 × 100 = 97%
Especificidade = 275/279 × 100 = 99%

2) Segunda situação – considerando os resultados limítrofes como positivos:

Imunocromatografia	Doença X		Total
	Presença	Ausência	
Positivo	86 (72 + 14)	22 (4 + 18)	**108**
Negativo	2	275	**277**
Total	**88**	**297**	**385**

Sensibilidade = 86/88 × 100 = 98%
Especificidade = 275/297 × 100 = 93%

Segunda situação – considerando os resultados limítrofes como negativos:

Imunocromatografia	Doença X		
	Presença	Ausência	Total
Positivo	72	4	**76**
Negativo	16 (2 + 14)	293 (275 + 18)	**309**
Total	**88**	**297**	**385**

Sensibilidade = 72/88 × 100 = 82%
Especificidade = 293/297 × 100 = 99%

3) Um teste de *screening* deve se caracterizar pela alta sensibilidade, assim, o melhor teste seria aquele onde os resultados limítrofes são considerados como positivos, já que neste caso, o valor da sensibilidade é 98%, de maior do que as outras situações, 97 e 82%, para as situações de exclusão dos limítrofes ou os considerando como negativos.