

14.04 Medição e comparação de resultados na prática assistencial



O estudo de resultados (outcomes research) avalia as consequências das decisões e intervenções de saúde em termos de seu produto final: a saúde individual e da população.

Autor: Salvador Peiró

Área de Pesquisa em Serviços de Saúde, Centro Superior de Pesquisa em Saúde Pública (CSISP), Valência.

Se recomienda imprimir 2 páginas por hoja

Citación recomendada:

Peiró S. Medição e comparação de resultados na prática assistencial [Internet]. Madrid: Escuela Nacional de Sanidad; Traducción 2015 [consultado día mes año]. Disponible en: direccion url del pdf.



TEXTOS DE ADMINISTRACION SANITARIA Y GESTIÓN CLINICA
by UNED Y ESCUELA NACIONAL DE SANIDAD
is licensed under a Creative Commons
Reconocimiento- No comercial-Sin obra Derivada
3.0 Unported License.



Resumo

A avaliação dos serviços da saúde pode incluir a avaliação da estrutura (acreditação, certificação), os processos (audit), os resultados ou as relações entre esses elementos (análise de produtividade, avaliação econômica, análise de eficiência ou efetividade). **A análise de resultados é utilizada para avaliar a eficácia ou efetividade das tecnologias** (testes clínicos, estudos observacionais) **ou para comparar as organizações de saúde** (indicadores de atividade e perfis de prática). As

comparações entre os fornecedores devem ajustar as diferenças na **gravidade** dos pacientes tratados por cada fornecedor.

Introdução

1. Marco conceitual

1.1. Manejo

1.2. Características

2. Aspectos metodológicos

2.1 Definição do objetivo do sistema

2.2. Seleção dos fatores de risco

2.3. Construção e desenvolvimento

2.4 Validação

2.5 Performance

3. Uso prático

4. Perspectivas

Referências bibliográficas

As medidas de **satisfação, estado funcional e qualidade de vida relacionada à saúde** são possíveis medidas de resultados, com vantagens e inconvenientes se comparadas com outras medidas clássicas (mortalidade, complicações, etc.). Os aspectos clinimétricos (fiabilidade, validade, sensibilidade às mudanças) são essenciais na escolha de uma ou outra medida.

Os **indicadores de atividade** são a forma de operacionalizar as medidas de processos e resultados em gestão de saúde. A sua utilização requer determinados requisitos, dentre os quais os mais importantes são a relevância, a capacidade de serem atribuídos aos fornecedores e a sua racionalidade clínica.

Introdução

Os elementos que melhor configuram o atual meio do atendimento da saúde incluem, provavelmente:

- uma maior preocupação pelo gasto com a saúde dos governos e das sociedades,
- a consciência de amplas variações na prática médica, não explicáveis por diferenças na morbidade das populações,
- a preocupação pela eficácia, efetividade e segurança das práticas de saúde, que aumentam na presença de uma fortíssima inovação tecnológica - em produtos, antes

que em processos - e que, no entanto, têm um aspecto decrescente sobre a saúde das populações, e

- O crescente interesse dos pacientes em obter informação sobre saúde e em participar das decisões que afetam a sua saúde.

Esse contexto favoreceu a popularidade de uma variante de pesquisa clínica, o **estudo de resultados de saúde** (*outcomes research*), e a rápida difusão de muitos de seus elementos à prática clínica habitual. **O estudo de resultados pretende avaliar as consequências das decisões e intervenções de saúde em termos de seu produto final, isto é, a saúde individual e da população e a satisfação dos usuários.**

Comparar resultados é a forma mais comum de avaliar a eficácia dos medicamentos, testes, diagnósticos e outras tecnologias ou intervenções de saúde (incluindo as educacionais, organizacionais, etc.). Tipicamente, essa comparação é realizada mediante testes clínicos com designação aleatória de pacientes a dois ou mais grupos; a ideia básica é que as diferenças - estatisticamente significativas - nos resultados para os pacientes terão sido causadas por diferenças na eficácia dos tratamentos, tecnologias ou intervenções.

A comparação de resultados de fármacos ou tecnologias na prática real (**efetividade**) e as comparações entre organizações de saúde (**qualidade**) foram áreas de pesquisa menosprezadas até que o desenvolvimento de os *Patient Outcomes Research Teams* (PORTs) e a publicação das taxas de mortalidade para mais de 5000 hospitais pela *Health Care Financing Administration* (HCFA) dos Estados Unidos despertassem uma atenção sem precedentes para a medição de resultados em condições reais.

Apesar de determinadas limitações, a análise dos resultados oferece importantes oportunidades para valorizar a efetividade e a eficiência das práticas médicas, tanto nos aspectos de avaliação de tecnologias médicas como de avaliação de organizações e centros de saúde, sejam esses hospitais ou outros centros de saúde, serviços ou médicos individuais.

Nesse sentido, tenta-se estabelecer as bases conceituais para a medição e comparação de resultados, com especial atenção a suas aplicações práticas.

- No ponto 1 apresenta-se um marco conceitual para a avaliação em atendimento de saúde, no qual se situa a avaliação de resultados, e descrevem-se os tipos de resultados utilizados com maior frequência.
- O ponto 2 é destinado a mostrar o raciocínio básico da utilização das medidas de resultado no atendimento da saúde, tanto para avaliar tecnologias como fornecedores.
- Enquanto os resultados clássicos (mortalidade, morbidade, complicações, etc.) estão bem aceitos no atendimento clínico, os resultados que dependem de escalas de medição são muito menos conhecidos. O ponto 3 dedica-se à medição da satisfação, o estado funcional e a qualidade de vida.
- O ponto 4 dedica-se à operação prática das medidas de resultado (e outras) em indicadores úteis em gestão clínica e de saúde.

1.- Marco conceitual para a medição de resultados

Embora, durante décadas, o interesse nos dados de mortalidade tenha estado muito vinculado ao mundo acadêmico e de pesquisa, a publicação na *Health Care Financing Administration* (HCFA), em 1986, das taxas de mortalidade hospitalar de, aproximadamente, 5000 hospitais norteamericanos provocou um interesse sem precedentes para a análise e interpretação das taxas de mortalidade hospitalares. A HCFA manteve a sua publicação, com importantes melhoras metodológicas, até 1993.

Embora inicialmente a maioria dos esforços tenham estado concentrados nas taxas ajustadas de mortalidade, posteriormente, e coincidindo com o desenvolvimento do chamado movimento para a medição de resultados (**outcomes movement**) e outras iniciativas de melhoria da qualidade na prática real focaram-se em outros resultados para detectar potenciais problemas de qualidade, sendo especialmente significativos os trabalhos que medem fenômenos como o da reinternação, as complicações, o estado funcional, a qualidade de vida.

A indagação na relação entre processos (o atendimento de saúde prestado) e resultados (as mudanças no estado de saúde dos

pacientes) tem sido complexa e problemática. Poucas práticas médicas aceitas estão solidamente fundamentadas em estudos bem controlados a partir dos quais fica estabelecida uma relação entre o processo de cuidados e o resultado clínico. Isso coloca o problema da atribuição dos resultados, que Donabedian explicou da seguinte maneira: *“a presença de uma relação causal válida entre um processo específico e um resultado significa que só é possível alcançar certo resultado sob condições específicas. Isso não significa que o resultado observado em uma situação dada tenha sido produzido, nesse caso, pelo processo citado (...)”*. Isto é, quando são utilizados resultados para fazer inferências sobre a qualidade dos cuidados, é necessário estabelecer que resultados podem ser atribuídos a esses cuidados (**validade da atribuição**). É preciso destacar que essa classe de validade depende do estabelecimento de uma relação causal entre processos e resultados sobre bases científicas e é - ainda hoje - o principal problema para a aplicação prática da avaliação de resultados no atendimento da saúde.

1.1. A avaliação em atendimento de saúde

Donabedian (1966) conceitualizou os exercícios de mensuração-avaliação do atendimento de saúde na tríade estrutura, processo e resultados (**figura 1**).

- A **estrutura** inclui os suportes físicos, econômicos, de pessoal, de formação, organização etc., destinados ao atendimento de saúde, incluindo os conhecimentos existentes e as capacidades e habilidades do pessoal.
- Os **processos** dizem respeito àquilo que os fornecedores de serviços de saúde (hospitais, centros de saúde, médicos, etc.) fazem por e para os pacientes, e também como os pacientes buscam o atendimento médico e respondem às pautas terapêuticas recomendadas.
- O impacto desses processos sobre a saúde dos pacientes são os **resultados** ou **desfechos** (outcomes, endpoints), definidos como as **mudanças, favoráveis ou não, no estado de saúde atual ou potencial de pessoas, grupos ou comunidades, que podem ser atribuídos ao atendimento de saúde prévio ou atual**.

Figura 1.
Marco conceptual de la medición-evaluación de servicios sanitarios: modelo clásico

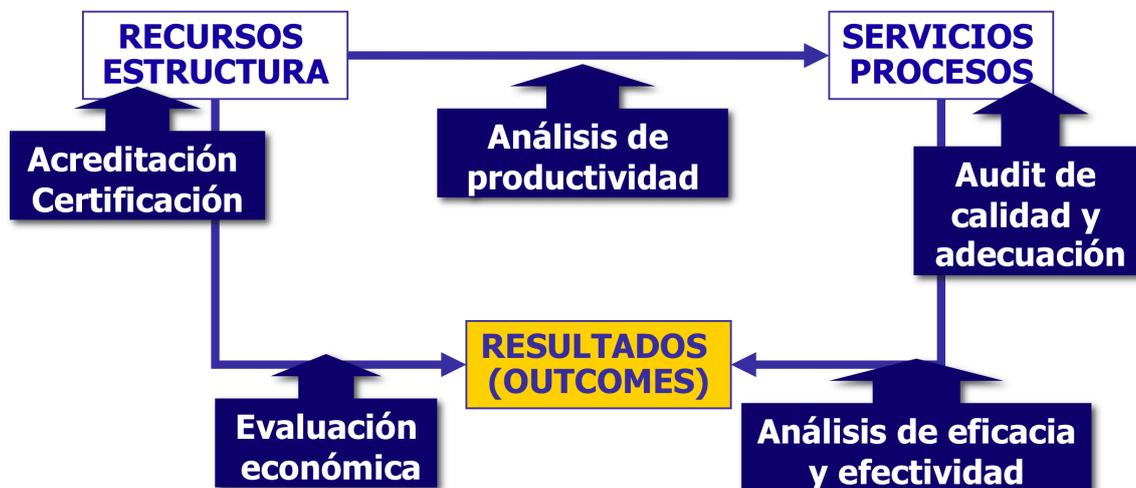


Sobre esse marco é possível situar as diversas modalidades de avaliação que corresponderiam a diversos tipos de medidas ou indicadores (**figura 2**). *As medidas de estrutura* têm pouco interesse do ponto de vista da avaliação de resultados (embora às vezes sejam utilizados devido à sua fácil disponibilidade), em parte porque os mais importantes são exigidos como requisito prévio nos procedimentos de autorização de centros da saúde, em parte devido à sua pouca correlação – pelo menos no âmbito dos países com sistemas hospitalares modernos - com os resultados para os pacientes.

As *medidas de processo*, incluindo as de utilização de serviços, podem incluir:

- taxas de eventos desejáveis (por ex. percentagem de mulheres de determinados grupos de idade que comparecem aos programas de triagem do câncer de mama),
- taxas de eventos normais mas que se forem elevadas sugerem problemas de qualidade (por ex. taxa de cesáreas), ou
- taxas de eventos não desejáveis (% de pacientes com mais de 2 meses de espera entre uma triagem positiva e a cirurgia de mama).

Figura 2.

Marco conceptual: modelo clásico

Os indicadores podem ser interpretados por meio de normas estatísticas, normas determinadas pelas autoridades de saúde ou profissionais e normas baseadas em evidências científicas.

As **normas estatísticas** são aplicadas a grupos e permitem avaliar se os fornecedores atuam de forma diferente, mas não a adequação do atendimento em pacientes individuais. Geralmente, derivam-se da prática observada em um determinado conjunto de fornecedores e - quando se adota a média como referência - tendem a reforçar o status quo pelo qual foi sugerido que as comparações deveriam ser realizadas com relação a centros de excelência em um determinado campo.

As **normas baseadas na evidência científica** representam critérios explícitos de cuidados apropriados baseados nos resultados dos testes clínicos. Geralmente, tomam a forma de porcentagem de preenchimento de uma recomendação normativa para o atendimento de uma situação concreta em um grupo definido de pacientes (por ex. trombose na hora seguinte à internação hospitalar em pacientes com infarto agudo do miocárdio e sem contraindicações para o fármaco; a porcentagem desses pacientes encaminhados por um médico que tinham recebido ácido acetilsalicílico).

As **normas baseadas em consensos** são semelhantes (por ex. porcentagem de diabéticos insulíndependentes que

recebem um controle de glicemia trimestralmente e um checkup oftalmológico anualmente).

As medidas **de processo são o eixo central dos indicadores de funcionamento** - muito mais do que os resultados - e, quando os testes clínicos demonstram que o processo medido está associado a melhores resultados, a devolução da informação para os clínicos é um dos instrumentos mais potentes de melhoria da qualidade.

Na relação entre os recursos e os resultados encontram-se as medidas de produtividade, entendidas como a relação entre os recursos (expressos em unidades físicas: médicos, enfermeiras, leitos, centros cirúrgicos, etc.) e determinados processos que são utilizados como produtos intermediários (expressos também em unidades físicas: estadias no hospital, intervenções, testes diagnósticos, etc.); embora a análise da produtividade nos serviços de saúde enfrente importantes problemas conceituais e práticos na definição do recurso e do produto, costuma tratar-se de medidas de fácil obtenção que formam a parte mais volumosa dos indicadores comuns de funcionamento hospitalar (estadia no hospital média, índice de rotatividade, funcionamento de centros cirúrgicos, etc.).

As medidas de eficiência estão centradas na relação entre recursos (medidos agora como unidades monetárias) e resultados de eficácia ou efetividade (por exemplo, custo por caso de hepatite evitada ou por ano de vida ganho); seu instrumental básico é a avaliação econômica em qualquer uma de suas modalidades, especialmente a análise custo-efetividade ou custo-benefício. Como indicadores de funcionamento costumam utilizar-se medidas indiretas derivadas de pesquisas que demonstraram a eficiência de alguma prática (por ex. porcentagem de prescrição de genéricos ou de cirurgia ambulatorial em determinadas intervenções).

Finalmente, as **medidas de resultados** estão focadas em aqueles aspectos realmente importantes para os pacientes: mortalidade, complicações, estado funcional, qualidade de vida, satisfação, etc. Apesar de sua crescente utilização, **são amplamente discutidas visto que os resultados não dependem exclusivamente do atendimento prestado**. Exemplos típicos desses indicadores são as taxas de mortalidade pós cirurgia cardíaca, as taxas de

complicações de feridas, de infecções hospitalares, as pesquisas de satisfação, etc. Atendendo o sentido das mudanças, os resultados podem ser considerados favoráveis (refletindo uma melhoria em uma ou mais das dimensões do estado de saúde) ou adversos (representando um problema).

Visto que os processos de cuidados afetam diferentes dimensões do estado de saúde e que os resultados podem estar inter-relacionados, é necessário diferenciar entre **desfechos** (*endpoints*) e **desfechos substitutos** (*surrogate endpoints*). Os desfechos são aqueles que têm relevância clínica para os pacientes e são os que realmente se deseja conseguir com as intervenções de saúde. A redução da mortalidade é o exemplo clássico de desfecho, mas também é um desfecho a melhoria no estado funcional ou na qualidade de vida. A pesquisa clínica frequentemente utiliza desfechos substitutos (por exemplo, controle da pressão arterial) como evidência de sucesso no tratamento ou no processo de cuidados, mas esse tipo de resultados deve ser visto com cautela, porque o que realmente se deseja reduzir é a mortalidade (ou a mortalidade cardíaca, ou a incidência de infartos) e não em todos os grupos de pacientes a associação entre *surrogate endpoints* e *endpoints* tem a mesma importância.

1.2. Tipos de resultados

Na [tabela 1](#) apresentam-se os resultados mais utilizados no atendimento de saúde com alguns exemplos de operação desses resultados. A **mortalidade**, um dos resultados mais utilizados por sua natureza decisiva e fácil disponibilidade, costuma ser utilizada em diferentes janelas temporais com relação ao tipo de patologia ou com o processo de interesse (por exemplo, os cirurgiões estão especialmente interessados na mortalidade 10 e 30 dias após a intervenção cirúrgica), por causas específicas (mortalidade neonatal, materna, etc.), ou como sobrevivência após o diagnóstico (frequente em oncologia). É um indicador de pouca utilidade para avaliar processos com baixa mortalidade prevista (por exemplo, a cirurgia sem internação), embora nesses supostos possa haver interesse na análise detalhada dos casos concretos. Tampouco é muito útil para avaliar centros de atendimento primário, e não muito por limitações de tipo estatístico, mas sim pela dificuldade de atribuir as diferenças

em mortalidade às diferenças em qualidade do atendimento primário.

As **complicações e outros eventos adversos**, como as reinternações, as infecções hospitalares e algumas complicações específicas (de ferimento cirúrgico, recidivas, falhas de intervenção), são cada vez mais utilizadas junto a - ou em lugar de - a mortalidade por sua maior frequência (maior poder estatístico) e por uma suposta maior relação com problemas de qualidade no atendimento passíveis de correção. Nesse terreno cabe assinalar a crescente utilização desse tipo de medidas com relação à prescrição farmacêutica (**reações adversas a medicamentos**). Foi proposta a utilização de alguns tipos de complicação que podem requerer atendimento hospitalar em emergências ou mediante internação (por exemplo, descompensações diabéticas) para avaliar a qualidade do atendimento primário a partir dos - mais consolidados - registros hospitalares. As propostas de utilizar as "hospitalizações evitáveis" para monitorar a qualidade do atendimento primário seriam feitas nesse sentido.

Tabla 1. Resultados de la atención de salud

<i>Resultados</i>	<i>Ejemplos</i>
Supervivencia o mortalidad	Mortalidad intrahospitalaria; Mortalidad a los <i>n</i> días de la intervención; supervivencia a los <i>n</i> años del diagnóstico.
Otros sucesos adversos	Reingresos en determinados periodos o por causas específicas; infecciones nosocomiales; complicaciones post-quirúrgicas; complicaciones de la herida quirúrgica, reacciones adversas a fármacos; reintervenciones inesperadas; recidivas a los <i>n</i> años de la intervención.
Estabilidad de los sistemas fisiológicos	Normalización/mejora de signos clínicos: pulso, temperatura, u otros; normalización/mejora de parámetros bioquímicos; normalización mejora de otros parámetros.
Satisfacción	Encuestas de satisfacción y otros instrumentos
Estado funcional	Capacidad para el esfuerzo; capacidad para las actividades de la vida diaria; capacidades específicas: visuales, movilidad, y otras similares; normalización/mejora en pruebas funcionales; reincorporación a la vida laboral.
Funcionamiento psico-social	Estado cognitivo; escalas de ansiedad, depresión, estrés, y otras similares.
Dolor	Escalas analógico-visuales, instrumentos específicos de medición del dolor.
Calidad de vida relacionada con la salud	Medición de utilidades; índices de estado de salud; perfiles de salud genéricos o específicos; instrumentos diagnóstico-específicos de medición de calidad de vida.

A **estabilidade ou melhoria dos sistemas fisiológicos** é utilizada como medida intermediária de resultado em muitas

patologias crônicas (diminuição da pressão arterial em mmHg, redução dos níveis hemáticos de lipídios, controle da glicemia ou da hemoglobina glicada, etc.) que costumam ser controladas no atendimento primário. Também é utilizada nas unidades de terapia intensiva onde muitas dessas variáveis são críticas na avaliação de resultados imediatos (normalização gases, frequência cardíaca, ...).

A **capacidade funcional** (muito utilizada em especialidades como a cardiologia ou pneumologia, com classificações históricas - como a da insuficiência cardíaca da *New York Heart Association (NYHA)* utilizada desde 1972- e testes funcionais objetivos amplamente utilizados como os respiratórios, de esforço e outros), assim como o **funcionamento psicossocial** têm-se tornado peças chave para medir resultados importantes para os pacientes e, em algumas especialidades (por exemplo, a Cirurgia Ortopédica ou a Reabilitação) são a medida básica de resultados. A capacidade funcional implica uma avaliação feita por alguém **externo ao paciente** e com **critérios objetivos** dos efeitos da doença sobre a vida do indivíduo; a avaliação funcional exclui os aspectos subjetivos. As medições concretizam-se como pontuações em **escalas de avaliação funcional genéricas** (índices das atividades da vida diária) ou **específicas das patologias a avaliar** (por exemplo: *Knee Society Score*, escalas de acuidade visual...). O essencial desses instrumentos é que permitem quantificar os resultados clínicos relevantes em muitas patologias (ou serviços) de um modo viável, uniforme, objetivo e reproduzível.

A medição da **qualidade de vida relacionada à saúde** (HRQoL) pressupõe um esforço para medir de forma abrangente os resultados importantes para os pacientes. De certa forma, surge como uma extensão das medidas de capacidade funcional às quais são incorporadas uma série de componentes, dentre os quais se destacam a **percepção subjetiva por parte do paciente sobre o impacto da doença** (aspecto que requeria um foco multidimensional e não apenas de função) e a **relação com os cuidados da saúde**. Esse último aspecto é transcendental porque o termo qualidade de vida é excessivamente geral e as abordagens úteis no contexto do atendimento médico concentram-se naqueles aspectos da experiência vital das pessoas relacionadas com a saúde e a assistência de saúde (aspecto que é descrito com maior precisão pelo termo qualidade de vida relacionada à saúde, *health related quality of life*).

Finalmente, a **satisfação dos pacientes** acompanhando as transformações sociais e a mudança do papel dos usuários dos serviços de saúde tem-se tornado um resultado importante, para cuja medição numerosos marcos conceituais e pesquisas específicas forma desenvolvidos. Cabe assinalar, no entanto, que esse tipo de pesquisa é utilizada mais como instrumento de melhoria de qualidade do que como medida de resultados.

2. Utilidade da mensuração de resultados

Comparar resultados é uma maneira normal de avaliar a eficácia dos medicamentos, testes diagnósticos e outras tecnologias. Tipicamente, essa comparação é realizada por meio de testes clínicos com atribuição aleatória de pacientes a dois ou mais grupos. A ideia básica é que as diferenças de resultados para os pacientes tenham sido causadas por diferenças na eficácia dos tratamentos ou das tecnologias. A atribuição aleatória dos pacientes é essencial nesse tipo de estudos, visto que garante que aqueles que receberam um determinado tratamento não eram diferentes daqueles que receberam o tratamento alternativo e, excluindo as diferenças entre pacientes como causa das diferenças em resultados, permite atribuir os melhores resultados à superioridade de um tratamento sobre a alternativa.

O problema fundamental ao comparar os resultados fora de situações experimentais está em que, diferentemente dos testes clínicos, os pacientes não forma atribuídos aleatoriamente para o tratamento por um ou outro centro, mas sim, pelo menos em boa parte, têm sido atribuídos em função de sua própria condição clínica. Nessa situação, não podemos pressupor que as diferenças em resultados devam-se exclusivamente à qualidade dos centros da saúde, já que estarão muito influenciados pelas características dos pacientes tratados em cada lugar. As características dos pacientes que definem seu risco de piores resultados agrupam-se sob o nome de **gravidade** (*severity*).

As características dos pacientes que configuram a gravidade incluem fatores como a idade, o sexo, o diagnóstico principal e sua gravidade, a co-morbilidade (doenças concomitantes não relacionadas com o diagnóstico principal) incluindo sua gravidade e extensão, as complicações (patologias surgidas durante ou após o processo de atendimento, que não estavam presentes na

hora da internação), a estabilidade clínica, o estado funcional, o estado psicológico, cognitivo e psicossocial, e outras variáveis.

A forma de abordar o problema de se as diferenças em resultados devem-se à diferente gravidade dos pacientes ou as diferenças em qualidade é habitual em epidemiologia quando se quer isolar o efeito de fatores de confusão: ajustar (geralmente mediante a modelização multivariante) os resultados obtidos em cada centro por aquelas características dos pacientes que definem seu risco prévio com relação ao resultado que é utilizado.

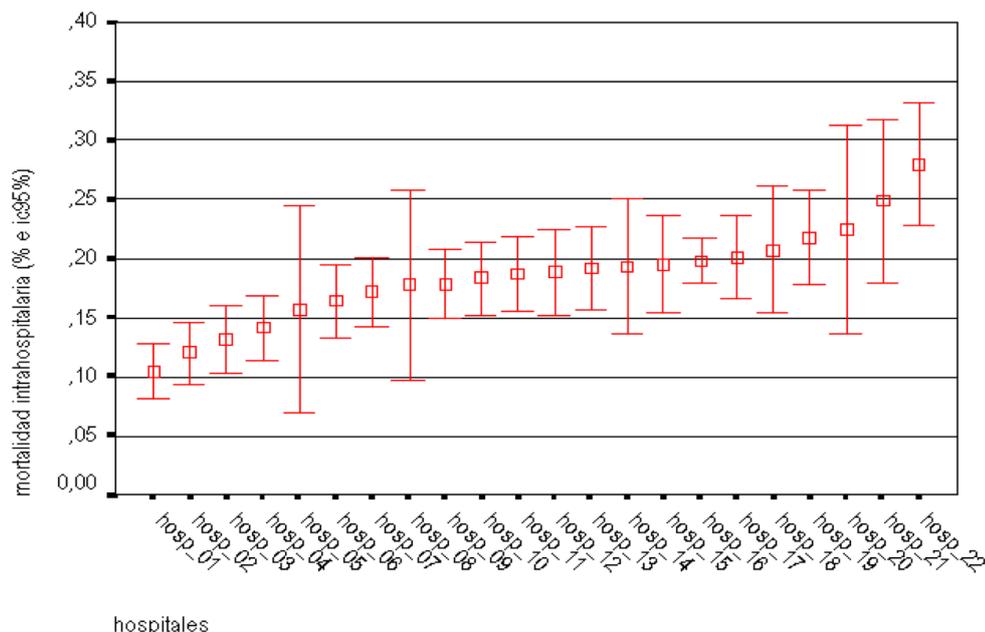
Se avaliarmos as diferenças em qualquer resultado (mortes, internações, complicações, qualidade de vida) entre tratamentos ou hospitais, podem ser construídos modelos matemáticos com as variáveis de interesse (idade, comorbilidade, diagnóstico, estabilidade clínica, etc.) para cada um desses resultados; dessa maneira pode ser obtido um resultado ponderado pela gravidade dos pacientes atendidos em cada centro que, até certo ponto, pode ser interpretado como o resultado que teria tido cada centro se seus pacientes tivessem sido iguais nas variáveis incluídas (em outras palavras, o resultado teórico que teriam tido os hospitais se tivessem atendido pacientes hipotéticos com a idade média de todos os hospitais, o sexo médio, um diagnóstico médio, etc.). Essa é a base para a construção de **sistemas de classificação de pacientes** ou - generalizando, visto que há sistemas que não constroem grupos de pacientes - **sistemas de ajuste de riscos** (*risk adjustment systems*) que, simplificando, não são outra coisa que métodos de quantificar a probabilidade que os pacientes têm de obter um determinado resultado. Esses sistemas serão abordados em outro tópico.

Na **figura 3** apresentam as taxas de mortalidade intrahospitalar, incluindo seus intervalos de confiança de 95%, no acidente vascular cerebral (exceto isquemia transitória) em 22 hospitais públicos da Comunidade Valenciana (1993-1995). É possível observar que essas taxas, para o conjunto dos 3 anos incluídos, variavam segundo os hospitais de 10% a 28%, e que as diferenças eram estatisticamente significativas (os intervalos de confiança de muitos hospitais não são escondidos).

Figura 3. Medidas de resultado: calidad de los proveedores

MORTALIDAD INTRAHOSPITALARIA EN EL GDR 14 (ACCIDENTE VASCULO-CEREBRAL excepto ISQUEMIA TRANSITORIA).

Hospitales públicos del País Valenciano (1993-95)



Essas diferenças, aplicando o raciocínio prévio de inferir a partir das diferenças em resultados para a qualidade do atendimento prestado, implicariam graves problemas de qualidade. Mas, na prática real, a situação é mais complexa, visto que os hospitais sem serviço de neurocirurgia encaminharam os casos mais graves (derrames hemorrágicos, com mortalidade geralmente superior a 30%) para os hospitais que dispunham desses serviços, enquanto retiveram os acidentes trombóticos (com mortalidade em torno de 9%). No caso apresentado, no hospital com a menor taxa de mortalidade, só 25% das altas deviam-se a hemorragias cerebrais, enquanto que esse número era de 57% no hospital com maior mortalidade. Logicamente, a mortalidade foi maior nesse último, mas esse excesso de mortalidade não ocorreu tanto devido a um problema de qualidade como ao fato que os “piores” pacientes foram encaminhados para os “melhores” centros.

Essa situação, comum (e desejável) no atendimento de saúde prática (como norma será feita uma tentativa para que os pacientes sejam atendidos no lugar com maiores recursos

e experiência na patologia em questão) faz com que as comparações de resultados entre organizações de saúde devam ser realizadas com grande cautela, seja qual for a medida de resultados utilizados, e inclusive quando utilizam-se sistemas de ajuste de riscos. Como norma, os resultados desse tipo de análise (estudos observacionais, e geralmente sobre dados tomados retrospectivamente) nunca terão a mesma solidez para inferir desde os resultados à qualidade dos centros da saúde, que a que demonstram os testes clínicos com atribuição aleatória para inferir desde os resultados à eficácia de uma tecnologia médica concreta. Apesar disso, é o tipo de análise que pode ser utilizada para avaliar os resultados dos fornecedores e, na prática e aplicado com prudência, pode ser extremamente útil para melhorar a qualidade do atendimento.

3. Estudo especial das medidas de satisfação, estado funcional e qualidade de vida

3.1. A medida da satisfação

A medida da satisfação dos pacientes capta um tipo de informação - a avaliação pessoal dos usuários com relação às suas preferências e expectativas - que não pode ser obtida mediante a observação externa e, portanto, além da informação que poderia obter o fornecedor utilizando instrumentos objetivos. Desde os trabalhos iniciais de Hulka, tanto os sociólogos como os pesquisadores em serviços de saúde desenvolveram diversas teorias e modelos para definir os "pontos de vista" dos pacientes sobre o atendimento que recebem, sem que ainda exista uma definição aceita e uma clara conceptualização da "satisfação" dos pacientes.

Uma revisão da literatura focada na satisfação com o atendimento hospitalar que incluía 112 publicações, identificou uma taxonomia de componentes com relação à experiência dos pacientes cuja inclusão era extraordinariamente diferente entre os diferentes instrumentos. Essa revisão mostrava, além disso, uma enorme variabilidade nas formas de descrever como essas experiências eram avaliadas nos diferentes estudos (excelente ou regular,

satisfatória ou insatisfatória, agradável ou desagradável, disposição ou não a retornar ou recomendar o serviço, etc.), as taxas de resposta, o custo e aceitabilidade para os possíveis usuários das pesquisas, a confiabilidade coletada em um único trabalho dos 112 revisados), as formas de avaliar a validade utilizadas - com grandes insuficiências - e os possíveis vieses (não resposta, forma de administração, tempo de alta, formato das respostas, viés de contexto, respostas como familiares).

Essa situação de discrepâncias na definição e conceptualização da satisfação dos pacientes, junto a uma grande difusão das pesquisas formando parte de estratégias de *marketing* ou mediáticas (as pesquisas de satisfação não são apenas um instrumento de pesquisa, mas que estão incorporadas à gestão rotineira de muitos centros e organizações de saúde), tem propiciado a existência de centenas - provavelmente milhares - de instrumentos de medida da satisfação diferentes cuja escolha está baseada mais no acaso, a moda ou aspectos conjunturais, que no conhecimento de suas propriedades clinimétricas ou à possibilidade de algum modelo conceitual definido.

Hulka conceitualizou sua abordagem como a **medição da satisfação dos pacientes como uma medida de resultado**, fundamentalmente dirigida a compreender as atitudes para o sistema de saúde em diferentes segmentos da população. Nesse sentido, Hulka assinalava que a satisfação dos pacientes devia ser avaliada como um resultado *per se* do atendimento de saúde que, adicionalmente, influenciaria na utilização e conduta dos pacientes (adesão, fundamentalmente). Ware, no entanto, considera a **satisfação como uma avaliação pessoal – subjetiva - dos serviços de saúde e dos fornecedores**. Embora reconheça que um nível determinado de satisfação contribui com informação do ponto de vista das preferências do paciente, relacionando desse modo com a perspectiva do *marketing*, não transforma a definição e medida da satisfação em uma análise de *gaps*, de elementos a corrigir, senão em uma abordagem à qualidade dos processos de um sistema de saúde concreto. Parasuman, do campo de marketing de serviços analisou as **diferenças (*gaps*) entre o serviço recebido e o serviço esperado, tanto por parte dos médicos como do usuário**. Esses estudos buscam a identificação desses *gaps* para adaptar os serviços às expectativas dos “clientes”, em geral, em três 3 áreas: expectativas do usuário frente a experiências do

usuário; expectativas do usuário frente a percepções profissionais das expectativas do usuário, e experiências do usuário frente a percepções profissionais das experiências do usuário. No atendimento de saúde, o desenvolvimento mais importante desta perspectiva é o **SERVQUAL** que mede a qualidade dos serviços como diferenças entre as expectativas e as percepções dos pacientes.

É preciso assinalar que, embora a difusão e aplicação real dos três modelos (Hulka, Ware, Parasuman) tenha sido diferente, os três continuam atualmente em uso. A pesquisa original da Hulka utiliza-se pouco, mas existem numerosas adaptações da mesma, especialmente em atendimento primário e são a base da maior parte das pesquisas utilizadas nos hospitais. O trabalho do Ware e sua equipe teve uma ampla divulgação e aplicação no sistema de saúde norte-americano, dando origem ao *Consumer Satisfaction Survey*, um instrumento bem estudado que pretende medir a satisfação dos usuários com os Planos de Assistência de Saúde com a finalidade de que os financiadores dos serviços - basicamente as empresas que asseguram seus trabalhadores - possam estabelecer uma comparação entre eles. O SERVQUAL, por sua vez, é possivelmente a escala de medição de qualidade mais conhecida e utilizada nos Estados Unidos e, embora SERVQUAL não seja apenas um instrumento de mensuração de satisfação, no atendimento de saúde tem-se focado nesse aspecto e no de qualidade percebida que são utilizados por numerosos hospitais, embora possivelmente ainda predominem as baseadas no modelo de Hulka.

Uma característica dos instrumentos de mensuração da satisfação dos pacientes é que **os resultados tendem a estar enviesados positivamente**. De fato, por volta de 85%-95% dos entrevistados escolheram uma resposta ou opinião positiva. Uma explicação para isso é o efeito do "interesse social", no qual os entrevistados acham muito difícil expressar críticas sobre a assistência de saúde. Por esse motivo, os instrumentos desenvolvidos mais rigorosamente tiveram que superar um número de ciclos de teste dos itens nos quais são eliminadas as perguntas que produziam menor variabilidade. De qualquer jeito, esse viés positivo torna as pesquisas de satisfação em um instrumento muito pouco sensível às mudanças e pouco útil, na prática, como medida de resultado.

3.2. A medida do estado funcional

O estado funcional é conceitualizado como a capacidade para procurar os **próprios cuidados e desenvolver as atividades físicas**. Até o surgimento das escalas de avaliação funcional, na década dos 50, a capacidade funcional somente foi utilizada como medida de resultados.

O objetivo básico dessas escalas é permitir a avaliação inicial da função após a identificação da patologia, fazer o acompanhamento do processo, avaliar os resultados uma vez finalizado o processo, comparar resultados com outros profissionais que tenham utilizado a mesma intervenção terapêutica ou comparar diferentes terapias diante de um mesmo problema e, resumindo, proporcionar uma base para a comunicação científica e para evidenciar de forma objetiva a eficácia dos tratamentos.

Reforça-se o **aspecto objetivo da avaliação funcional** que, à diferença da CVRS, sempre é avaliada por uma pessoa externa ao paciente. No entanto, algumas escalas incluem a avaliação da dor, aspecto em boa medida subjetivo e que as aproxima das escalas de qualidade de vida.

As **atividades da vida diária** (AVD) foram o primeiro parâmetro funcional amplamente aceito e reconhecido e são definidas como **aquelas funções físicas, diárias e rotineiras, necessárias para viver**. Incluem atividades de cuidado pessoal (roupas, banho, higiene pessoal, uso do banheiro, transferências, alimentação, continência, deslocamento em interiores) e, embora exista um acordo geral naquilo que constitui o núcleo central das AVD, as chamadas AVD básicas (AbVD), existem variações nas atividades incluídas em diferentes escalas. As escalas de avaliação das AbVD mais utilizadas são o **índice de Atividades da Vida Diária** (*Index of Independence in Activities of Daily Living*) e o **índice de Barthel**. A principal limitação das escalas do AbVD é sua escassa sensibilidade para detectar pequenas mudanças clínicas, embora essas sejam relevantes. Para corrigir esse aspecto foram desenvolvidas as **escalas de avaliação das AVD instrumentais** (AiVD), que incluem atividades mais complexas (necessárias para viver independente em casa e na comunidade: preparação dos alimentos, cuidado da casa, lavagem de roupa, capacidade para utilizar o telefone, uso do transporte público, responsabilidade com o uso dos medicamentos

e administração do dinheiro, assim como o cuidado com as crianças e o deslocamento fora de casa) e são mais sensíveis às mudanças. Entre as escalas de AiVD mais utilizadas encontram-se a *Physical Self-Maintenance Scale*, mais conhecida como **Escala de Lawton**.

Embora a capacidade para realizar as AVD tenha sido utilizada como sinônimo de função ou rendimento físico, somente pressupõe uma parte da função física global que englobaria também as dimensões mental, emocional e social. Dentro desse contexto foram elaboradas escalas de avaliação funcional globais ou mistas que incluem aspectos da função física, da comunicação, cognitivos, sociais e psicossociais, entre as quais se destacam o **Functional Status Index**.

Igualmente, muitas escalas são específicas de um órgão (por exemplo: escalas de acuidade visual) ou de uma articulação (por exemplo, a Escala de Harris Hip) e, em alguns casos, os resultados de alguns testes são equiparáveis à medida da função de um órgão (por exemplo: audiometria, testes funcionais respiratórios, ergometria...). Assinalar, finalmente, que as escalas de avaliação globais incluem aspectos sobre os quais não é possível a intervenção médica, pelo qual, na prática, em muitos casos são selecionados para uso só aqueles aspectos de uma escala mais relevantes para a avaliação de um processo concreto.

3.3. A medida da qualidade de vida relacionada à saúde

A qualidade de vida é um conceito fácil de entender intuitivamente, mas complexo de definir e sobre o qual não existe um acordo universal. Em medicina costuma-se utilizar um conceito de qualidade de vida relacionado à saúde (*health related quality of life* - HRQoL) que mede **o impacto da doença e seus tratamentos sobre o bem-estar do paciente, incorporando - além das medições clínicas - a perspectiva do próprio paciente**.

A HRQoL é um **conceito subjetivo que só pode ser avaliado pelo próprio paciente**. A avaliação feita pelos profissionais não é apenas inadequada, mas também -como tem demonstrado vários trabalhos nos quais essas avaliações diferiam amplamente

das realizadas pelos pacientes- inexata. É preciso assinalar que a subjetividade das medidas da HRQoL não implica falta de reprodutibilidade que, em geral, é semelhante à das medidas facilitadas por instrumentos objetivos.

A HRQoL é um conceito multidimensional (para incorporar os diferentes aspectos do bem-estar), subjetivo (para apresentar a perspectiva do paciente) e dinâmico (muda ao longo do tempo, mas também em função do meio).

As primeiras medidas de qualidade de vida por sua proximidade - quando não confusão - com as medidas de capacidade funcional enfatizavam os componentes relacionados com a função física, mas a consideração multidimensional da HRQoL obriga a identificar e operacionalizar os componentes que vão integrá-la. Esses componentes denominam-se **dimensões ou domínios** e, no mínimo, incluem a dimensão física, a psicológica e a social. Embora inicialmente as medidas da HRQoL estejam apoiadas principalmente na dimensão física, evolutivamente tem sido incluído um maior número de dimensões pertencentes à vida psicológica e social, e tem sido dada uma maior especificação dos domínios que formavam cada uma delas.

As dimensões de maior interesse poderiam ser descritas assim:

- 1. Função física.** Semelhante ao conceito de capacidade funcional. Diz respeito à energia, força, e capacidade para realizar uma atividade costumeira. Contempla fundamentalmente 2 domínios: 1) A capacidade para as **atividades da vida diária** e, 2) A **capacidade para o esforço**, ou energia ou atividade positiva.
- 2. Função psicológica.** Diz respeito a indicadores cognitivos e afetivos. A **função afetiva**, referida ao distresse psicológico tem produzido o desenvolvimento de numerosas escalas orientadas principalmente à avaliação da depressão e ansiedade. A **função cognitiva** é também considerada um importante componente do bem-estar psicológico, visto que a memória e a capacidade para realizar tarefas intelectuais têm uma grande importância para as pessoas.
- 3. Interação e função social.** Diz respeito à capacidade do paciente para desenvolver as interações pessoais que

formam o núcleo da vida social. Essas interações estão tradicionalmente estruturadas de uma forma hierárquica: família, amigos, colegas e comunidade.

4. Sensações somáticas, limitações e alterações. Reúne as características de ordem fisiológica derivadas da doença e/ou tratamento. Em geral consideram-se os sintomas ou as moléstias relatados pelo paciente, as sensações desagradáveis (dor, náuseas e dispneia entre outros) e os sinais obtidos da exploração clínica, os resultados ou medidas fisiológicas provenientes de testes diagnósticos. São fundamentais nos instrumentos de qualidade de vida desenhados para doenças específicas.

5. Percepções de saúde. Baseadas nas próprias percepções do paciente, incluem as avaliações que o indivíduo faz dos efeitos que sua saúde tem sobre outros aspectos da vida.

Apesar da importância da conceptualização das dimensões incluídas em cada instrumento de medição, atualmente continua sem existir uniformidade na quantidade, significados, e importância com a qual contribuem na qualidade de vida. Por outro lado, a influência de valores socioculturais na qualidade de vida percebida confirma a necessidade de utilizar componentes segundo o contexto a ser avaliado.

A informação sobre qualidade de vida pode ser obtida por meio de entrevistas e perguntas abertas e/ou fechadas. Essa informação pode ser muito útil para o manejo de um paciente individual, mas, se poder ser quantificada em alguma medida, a sua utilidade para transmitir informação científica é bem pequena. Por essa razão, a informação sobre a HRQoL é obtida por meio de questionários estruturados que incluem itens (perguntas ou afirmações com as quais se está ou não de acordo) para coletar as diferentes dimensões. As respostas a essas perguntas são quantificadas em pontuações para cada uma das dimensões ou em uma única pontuação global.

Os instrumentos de mensuração da HRQoL podem ser classificados conforme diferentes critérios como o tipo de população a qual estão dirigidas, os domínios e/ou conceitos que cobrem, o sistema de pontuação mediante o qual reflete o nível individual e/ou global dos componentes e algum outro. No entanto, para

diferenciar o conjunto de medidas que incluem diversos aspectos relacionados com a qualidade de vida, o critério mais utilizado é aquele que diz respeito à sua aplicação geral (a diferentes populações) ou específica (doenças concretas), conhecidas comumente como **instrumentos genéricos** e **instrumentos específicos**. Na **tabela 3** encontram-se à disposição alguns endereços de internet com informação sobre esses instrumentos.

Tabela 3. Webs com informação sobre qualidade de vida relacionada à saúde

Título	Endereço	Comentários
EORTC Quality of Life Group	http://groups.eortc.be/qol/	Web da Quality of Life Unit do EORTC Data Center. Dedicado à qualidade de vida em câncer. Incorpora informação sobre o QLQ-30 e outros instrumentos desenvolvidos pela EORTC.
EuroQol Group	http://www.euroqol.org	Informação sobre o EuroQol Group e o EQ-5D.
Health and Quality of Life Outcomes.	http://www.hqlo.com/	A revista científica sobre qualidade de vida publicada pelo BioMed Central (gratuita).
International Society for Quality of Life Research	http://www.isoqol.org	Sociedade científica internacional dedicada a assuntos de qualidade de vida relacionada à saúde.
International Society for Quality-of-Life Studies	http://marketing.cob.vt.edu/isqols/	Instituição dedicada a estudos de qualidade de vida (não de saúde).
Mapi Research Institute	http://www.mapi-research-inst.com/index02.htm	Inclui informação de numerosos instrumentos, versões em diferentes idiomas e um jornal.
Medical Outcomes Group	http://www.outcomes-trust.org	Recolhe informação sobre numerosos instrumentos desenvolvidos ou administrados pelo Medical Outcomes Trust, incluindo o Sickness Impact Profile, a Quality of Well-Being Scale, o Duke Health Profile, o SF-36 Health Survey e o SF-12 Health Survey.
Quality of Life Instruments Database	http://www.proqolid.org/	A Quality of Life Instruments (QOLID) database, desenvolvida pelo Mapi Research Institute inclui mais de 700 questionários. Recurso essencial.
BiblioPRO	http://iryss.imim.es/iryss/BiblioPRO.asp	Biblioteca virtual de questionários de Qualidade de vida relacionada à saúde e outros resultados recebidos pelos pacientes (Patient Reported Outcomes, PRO) em espanhol de acesso livre e gratuito.

O objetivo dos **instrumentos genéricos** é que sejam aplicáveis a uma grande variedade de populações. Dividem-se perfis e índices de utilidade. Os **perfis de saúde** medem diferentes aspectos da qualidade de vida utilizando um sistema de pontuação comum, que pode corresponder a um número reduzido de itens que formam uma dimensão, ou algumas vezes

também permitem obter uma pontuação global, a partir da soma de valores alcançados em cada dimensão. A importância e o número de item outorgados a cada domínio pode ser muito diferente segundo os perfis. Assim, existem instrumentos que exploram as dimensões físicas, emocionais e sociais de modo equitativo, enquanto outros carregam sobre alguns domínios um número maior de itens, enfatizando assim a sua avaliação.

Como principais vantagens, os perfis costumam ter estudos de confiabilidade e validade bem estabelecidos, com frequência dirigidos a diferentes populações. Para objetivos discriminativos, pode-se examinar e determinar a área de disfunção que afeta um indivíduo ou grupo. A identificação dessas áreas, permite por sua vez, conduzir a elaboração de instrumentos específicos e orientar a intervenção clínica para onde produza um maior impacto na qualidade de vida. Quando são aplicadas com propósitos de avaliação, também permitem determinar os efeitos da intervenção sobre diferentes aspectos da qualidade de vida sem necessidade de aplicar múltiplos instrumentos. Por outro lado, facilitam a realização de estudos comparativos em diferentes populações. Por último, aqueles que proporcionam uma pontuação global, podem ser utilizados nas análises de custo-benefício sobre as intervenções clínicas. Entre as desvantagens, podem focar a sua atenção em diferentes aspectos da qualidade de vida que não sejam os mais interessantes para o objetivo de um determinado estudo, perdendo sensibilidade quando são aplicados a populações com doenças específicas.

Os perfis de saúde mas conhecidos incluem o **Medical Outcomes Study Short Form 36 (MOS-SF36)**, provavelmente o instrumento mas utilizado na atualidade e do qual existem algumas versões reduzidas, o **Perfil de Saúde do Nottingham** (Nottingham Health Profile, NHP), o **Perfil de Impacto da Doença** (Sickness Impact Profile, SIP) e os quadros **COOP-WONCA**, utilizados principalmente no atendimento primário. Todos eles dispõem de numerosos estudos de validação em diferentes populações, adaptações em espanhol (e também em alguns outros idiomas oficiais na Espanha, assim como versões no espanhol falado em diferentes países da América Latina), estudos de validação em população espanhola e, ocasionalmente, valores de referência em população geral. Em grande parte, trata-se de instrumentos registrados cuja utilização requer a autorização de seus proprietários (para usos em pesquisa esta

autorização costuma ser gratuita e ocasionalmente pode ser obtida pela internet).

Os **índices** são medidas derivadas da teoria econômica da decisão. A qualidade de vida é medida de modo global com um simples número que representa um valor em um contínuo, geralmente de 0 (morte ou pior situação possível) a 1 (saúde completa). Para a sua aplicação em estudos clínicos, existem duas abordagens fundamentais. Na primeira, pergunta-se ao paciente uma série de questões sobre suas funções físicas, sociais, etc. Com base nessas respostas, classifica-se o paciente numa determinada categoria que, por sua vez, está associada a um benefício ou utilidade, que tem sido previamente estabelecido em outros grupos. Essa abordagem caracteriza os instrumentos conhecidos como escalas de bem-estar e o mais utilizado em nosso meio é, provavelmente, o **EuroQol** (também dispõe de adaptação em espanhol e pesquisas de validação). A segunda abordagem coleta numa simples razão ou num valor numérico as preferências do paciente com relação a diferentes estados de saúde. Aqui os métodos mais utilizados foram a "loteria" padrão (**standard gamble**) e a preferência temporária (**time trade-off**).

A maior vantagem dos índices é a sua adequação às análise custo-benefício na avaliação econômica. De forma simples, trata-se de ponderar o tempo de vida ganho com cada intervenção pela qualidade de vida medida pelo índice para construir **Anos Ajustados por Qualidade de Vida** (*Quality Adjusted Life Year, QALY*). Por exemplo, se um transplante hepático permite prolongar a vida de uma pessoa em 10 anos, os três primeiros meses com baixa qualidade de vida (vamos supor um valor de 0,30), os 9 seguintes com uma qualidade intermediária (0,50) e o restante com uma qualidade aceitável (0,90), o acréscimo desses valores $[(0,25 \text{ anos} * 0,30) + (0,75 \text{ anos} * 0,50) + (9 \text{ anos} * 0,90)] = 8,55 \text{ QALY}$ nos dá o número de QALY ganhos.

Em teoria, os QALY podem ser úteis para atribuir recursos para aqueles programas de saúde ou intervenções que maior bem-estar produzam (medido em função das preferências das pessoas pesquisadas). Entre as limitações, as utilidades podem variar em função de como foram obtidas, e é questionada a validade de uma simples medida agregada. Além disso, e para seu uso na clínica, não permitem determinar que aspectos da qualidade de

vida são responsáveis pelas mudanças e são menos sensíveis que os perfis para detectar pequenas mudanças clínicas.

Outra alternativa na medida da qualidade de vida é focar-se em aspectos que sejam considerados relevantes para um determinado problema de saúde e aumentar a sensibilidade do instrumento sobre essa doença em concreto. Esse é o objetivo dos **instrumentos específicos** que podem partir de uma relação previamente estabelecida, ou bem focar-se nos problemas que relata o indivíduo.

A especificidade de um instrumento pode referir-se a um tipo de doença (dirigido a patologias oncológicas, cardíacas ou respiratórias), mas também pode orientar-se a uma determinada população (como os elaborados para avaliar a qualidade de vida dos idosos que apresentam uma grande variedade de problemas), a uma função em especial (questionários que examinam a função emocional ou sexual), ou bem ser específicos de uma condição ou problema comum para diferentes doenças (por exemplo, a dor). De todas essas abordagens, a que tem motivado um maior desenvolvimento de instrumentos é a que está dirigida a medidas específicas de doença que tem proporcionado uma considerável utilidade em estudos clínicos. Suas maiores desvantagens são a impossibilidade de utilizá-los em condições diferentes para as quais foram elaboradas (dada a deliberada ausência de diversos aspectos de qualidade de vida), devido a isso, em certas ocasiões são aplicados junto a outro genérico.

Atualmente, é possível encontrar centenas de instrumentos específicos. Entre os de maior interesse, com versões validadas em espanhol, cabe citar o Questionário de Qualidade de Vida da EORTC QLQ-30 (dirigido a pacientes com câncer, que inclui um módulo geral, para todos os cânceres, e módulos específicos para câncer de mama, pulmão, esôfago, cabeça e pescoço, cuidados paliativos, colo-retal e outros).

4. Da avaliação à gestão: operacionalizando as medidas de resultado

Os indicadores de atividade ou funcionamento (*performance indicators*) são, em essência, **critérios de avaliação aplicados a um conjunto de organizações ou fornecedores da saúde.**

Embora cada conjunto particular de indicadores implique uma concepção explícita das expectativas sobre o funcionamento do sistema de saúde ou de alguns de seus componentes (hospitais ou outros fornecedores, áreas de atendimento, etc.), existe um amplo consenso em não separar as medidas de produtividade, de qualidade e de resultados clínicos, de modo que os indicadores de funcionamento não se limitem aos de atividade e utilização de serviços, mas que incluam uma extensa gama de indicadores clínicos.

A utilidade básica dos indicadores de funcionamento **é fornecer informação útil aos diferentes agentes do sistema de saúde para facilitar suas escolhas e tentar que suas expectativas sejam atendidas**. Dado que os diferentes agentes - e grupos sociais - têm expectativas diferentes sobre os serviços de saúde, os indicadores desejáveis podem variar ostensivamente. Há pacotes de indicadores que são relevantes apenas para algum componente do sistema (geralmente, e dada a natureza técnica do atendimento de saúde, instituições de saúde e profissionais).

No entanto, a tendência atual é o desenvolvimento de pacotes estruturados de indicadores orientados - em algum grau - para os problemas dos pacientes e as comunidades, e que tentam ser tanto instrumentos essenciais das políticas de melhoria, como mecanismos para prestar contas à sociedade da eficiência, efetividade e qualidade com a qual são gerenciados os serviços de atendimento de saúde.

A utilização básica dos indicadores consiste em comparar fornecedores de serviços - ou outros componentes do sistema - em um determinado momento ou ao longo do tempo. Em sua forma mais interessante, um indicador de funcionamento não é mais do que a taxa de um determinado evento - numerador - que faz parte do atendimento prestado pelos fornecedores de saúde às populações definidas - denominador - em um tempo determinado.

Os eventos podem ser unidades monetárias, fazer parte dos processos de cuidados - incluindo os serviços prestados - ou tratar-se de resultados (*outcomes*) do atendimento médico. Os fornecedores podem ser médicos individuais ou, mais frequentemente, centros de saúde, hospitais, organizações de saúde ou territórios sob uma mesma autoridade de saúde. A população pode estar definida em função de sua localização

geográfica, embora o habitual seja defini-la também em função de alguma característica específica (diabéticos, hipertensos, etc.) ou do risco com relação a um resultado de interesse.

As comparações transversais de fornecedores utilizando indicadores são conhecidos como **perfis de fornecedores (profiling)**, termo que define *a aplicação de métodos epidemiológicos à descrição das práticas assistenciais, o monitoramento de resultados do atendimento de saúde e a avaliação da qualidade e a eficiência dos cuidados, com os objetivos de proporcionar a gestores, compradores, usuários e responsáveis pela tomada de decisões, informação para comparar o custo, a utilização e a qualidade dos fornecedores, e proporcionar aos próprios fornecedores uma referência de seus custos e a qualidade com relação a outros fornecedores, ou com relação à sua evolução temporária.*

Por exemplo, as taxas de administração de ácido acetilsalicílico (fato) por diferentes equipes de atendimento primário (fornecedores), em pacientes encaminhados para o hospital por suspeita de infarto agudo do miocárdio (população) em um ano determinado (tempo), definem o perfil dos fornecedores do atendimento primário com relação a um aspecto da assistência no infarto agudo, e permite elaborar estratégias de melhoria. Essa mecânica é aplicável a indicadores de resultados (taxa de complicações de ferida em intervenções de cirurgia maior programada, reintervenções não planejadas, reações adversas a medicamentos, etc.), de produtividade (permanência média ajustada por case-mix), de custos (por processo, por habitante de um território), de satisfação, de acessibilidade, etc.

O **benchmarking**, embora precise da realização de perfis, inclui a **identificação das melhores práticas para dar uma ideia dos benefícios que poderiam ser alcançados se as diversas organizações se situassem no nível das melhores**. Embora seja possível incorporar análises mais detalhadas para mostrar o que outros centros fazem melhor e facilitar o aprendizado, a análise de indicadores costuma-se limitar a estabelecer um padrão para cada critério definido pelos centros que demonstram os melhores resultados nesse critério (por exemplo, utilizar como norma desejável de duração da permanência em um processo concreto, a média de 5 ou 10% dos centros com melhor comportamento, em lugar da média de todos os centros).

Referências bibliográficas

1. Alonso J. La medida de la calidad de vida relacionada con la salud en la investigación y la práctica clínica. **Gac Sanit.** 2000; 14:163-7. [<http://www.elsevier.es/en/node/2008072>]
2. Badía X, Salamero M, Alonso J. La medida de la salud. Guía de escalas de medición en español (3ª edición). Barcelona: Edimac, 2002.
3. Brook RH, McGlynn EA, Shekelle PG. Defining and measuring quality of care: a perspective from US researchers. **Int J Qual Health Care.** 2000;12:281-95. [<http://intqhc.oxford-journals.org/content/12/4/281.long>]
4. Donabedian A. The role of outcomes in quality assessment and assurance. **Qual Rev Bull.** 1992;18:356-60.
5. Eddy DM. Performance measurement: problems and solutions. **Health Aff (Millwood).** 1998;17:7-25. [<http://content.healthaffairs.org/content/17/4/7.long>]
6. Fitzpatrick R, Fletcher A, Gore S, et al. Quality of life measures in health care. I: Applications and issues in assessment. **BMJ.** 1992; 305:1074-77. [<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1883623/?tool=pubmed>]
7. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. **Ann Intern Med.** 1993;118:622-9.
8. Guyatt GH, Naylor CD, Juniper E, et al. Users' guides to the medical literature: XII. How to use articles about health-related quality of life. **JAMA.** 1997;277:1232-7.
9. Kassirer JP. The use and abuse of practice profiles. **NEJM.** 1994; 330:634-6. [<http://www.nejm.org/doi/full/10.1056/NEJM199403033300910>]

10. Peiró Moreno S. *De la gestión de lo complementario a la gestión integral de la atención de salud: gestión de enfermedades e indicadores de actividad. En: Ortun V, ed. Gestión clínica y sanitaria: de la práctica a la academia, ida y vuelta. Barcelona: Masson; 2003:17-87.*

11. Testa MA, Simonson DC. *Assessment of quality of life outcomes. NEJM. 1996;334:835-40. [<http://www.nejm.org/doi/full/10.1056/NEJM199603283341306>]*

