

Uma avaliação crítica sobre o índice composto na avaliação dos sistemas de saúde no mundo*

A critical assessment of the compound rate for evaluating health systems around the world

Joaquim G. Valente

Professor do Departamento de Epidemiologia

Instituto de Medicina Social

Universidade do Estado do Rio de Janeiro

Pesquisador do Departamento de Epidemiologia e Métodos Quantitativos em Saúde

Escola Nacional de Saúde Pública

Fundação Oswaldo Cruz

Av. Leopoldo Bulhões, 1480, 8º andar

Bonsucesso – Rio de Janeiro, RJ

jvalente@ensp.fiocruz.br

* Artigo preparado para a Mesa Redonda "Indicadores: Aspectos Metodológicos" no V Congresso Brasileiro de Epidemiologia, Curitiba, PR, Brasil, 23-27 de março de 2002. As opiniões aqui expressas são de inteira responsabilidade do autor e não representam necessariamente pontos de vista institucionais da Fundação Oswaldo Cruz ou de seus diretores. O autor agradece a Célia Landman Szwarcwald, Gulnar Azevedo e Silva Mendonça, Cláudia Travassos, Claudio Salm e Alicia Ugá, pelos considerandos durante o processo de desenvolvimento da avaliação crítica.

Em 2000, a Organização Mundial de Saúde (OMS) liberou um relatório¹ avaliando os serviços de saúde nos 191 países-membros da OMS. O modelo teórico-conceitual incluía a avaliação em três domínios: saúde, responsividade (referente a aspectos não-médicos do atendimento) e justiça na contribuição social. Para os dois primeiros domínios, saúde e responsividade, o relatório apresentou resultados para o nível e a distribuição dos indicadores e, para a justiça na contribuição social, a trama conceitual incluía apenas a avaliação da desigualdade na distribuição da contribuição social.

O nível de saúde dos países-membros foi avaliado pela expectativa de vida ajustada para incapacidade, indicador chamado de DALE (*disability-adjusted life expectancy*). A distribuição do nível de saúde foi avaliada pela desigualdade interna dentro de cada país para a mortalidade na infância (entre os menores de cinco anos de idade). A avaliação dos aspectos não-médicos do atendimento incluía a mensuração do nível e da distribuição da responsividade, com pelo menos sete sub-áreas: rapidez do atendimento, dignidade, comunicação, autonomia, confidencialidade, liberdade de escolhas e qualidade das instalações.

A responsividade foi avaliada, no relatório mencionado, através de informantes-chaves selecionados em alguns países. No caso do Brasil, 50 participantes foram selecionados e convidados a responder a um questionário, o qual apenas 33 (67%) devolveram preenchido. Para os relatórios futuros de anos vindouros, a OMS pretende utilizar questionários a serem completados em inquéritos domiciliares.

Para a avaliação da distribuição da justiça na contribuição social, do total de remuneração de que as famílias dispunham, descontavam-se inicialmente os gastos com sua subsistência, excluindo-se apenas as despesas com alimentação. Do saldo remanescente, mensurava-se a proporção que foi considerada como *destinada à área da saúde*. Se uma família tivesse destinado mais de 50% desse saldo para a saúde dos seus componentes, essa família seria considerada como

tendo apresentado gastos *catastróficos* com saúde. Para a obtenção dos dados do relatório, obteve-se informações sobre a distribuição das famílias com gastos catastróficos em saúde, e a desigualdade da distribuição desse indicador foi estimada para cada país.

As cinco notas referentes ao nível e distribuição da saúde, nível e distribuição da responsividade e distribuição da justiça na contribuição social eram todas transformadas em valores de zero a 1,0 e somadas, ponderadamente, segundo pesos que serão apresentados e discutidos mais adiante, resultando em um indicador composto. Esse indicador composto era bastante sintético, resumia a situação da atuação dos serviços de saúde de cada país-membro em questão, e permitia o ranqueamento dos países da OMS, segundo o desempenho dos seus sistemas de saúde.

Apesar de todos os avanços que poderiam resultar da avaliação dos sistemas de saúde de nos vários países da OMS, e do apoio que se poderia dar a tal iniciativa importante e positiva de se avaliar a atuação dos serviços de saúde do conjunto de países da OMS, podem-se detectar alguns problemas, que vamos apresentar e discutir.

O primeiro grande problema dessa iniciativa da OMS foi relativa à falta de informações disponíveis para as estimativas. Dos 191 países-membros da OMS, 94 países não tinham qualquer informação em nenhum dos cinco componentes; 32 países tinham informação em apenas um dos componentes; e, 38 países tinham informação em dois componentes. Somente 20 países apresentaram resultados para três componentes e 7 países tinham dados sobre quatro ou cinco indicadores. A OMS pretende equacionar esse problema no futuro, obtendo dados diretamente das populações dos países-membros, através de inquéritos de saúde domiciliares, onde serão usados questionários incluindo perguntas do tipo “Com que frequência médicos, enfermeiros ou outros profissionais de saúde pediram sua autorização antes de iniciar um exame ou tratamento?” (autonomia) ou então “Com que frequência as conversas com seu médico, enfermeiro ou outro pro-

fissional de saúde ocorreram de maneira privativa, sem que outros pudessem ouvir o que se dizia?” (confidencialidade). Resta definir com que frequência esses inquéritos seriam aplicados, na medida em que a OMS pretende avaliar bianualmente os serviços de saúde no mundo. É também importante a discussão sobre a interferência que esses inquéritos periódicos poderiam exercer sobre os próprios serviços de saúde nas suas atividades de coleta de dados de rotina.

Um segundo problema digno de nota diz respeito à não replicabilidade dos resultados obtidos pela OMS. Um bom exemplo pode ser a comparação entre os resultados obtidos pela França e pela Austrália. Aplicando-se as fórmulas sugeridas no WHR-2000 às notas originais de cada um dos cinco componentes para os serviços de saúde desses dois países, obtém-se um total de 91.810 pontos para a França e 91.855 para a Austrália, indicando que os serviços de saúde deste último país tiveram um melhor desempenho do que aqueles da França. Entretanto, no WHR-2000, a França se apresenta em 6º lugar no ranqueamento dos países, com 91.925 pontos, e a Austrália em 12º lugar, com 91.330. Se repetirmos esses cálculos para os quinze países com sistemas de saúde com melhor desempenho no mundo, observaremos três alterações de ranqueamento.

Podemos agora percorrer a tabela divulgada pelo Relatório da OMS, onde os países foram colocados em ordem decrescente de desempenho dos seus serviços de saúde, e comparar os resultados apresentados na tabela com aqueles resultantes da aplicação das fórmulas sugeridas no WHR-2000. Observa-se que, à medida em que se desce na tabela, distanciando-se cada vez mais dos países com serviços de saúde com melhor desempenho, detecta-se uma discordância cada vez maior entre o número de ordem publicado no relatório e o número de ordem obtido através dos cálculos e aplicações das fórmulas sugeridas no próprio Relatório. Assim, entre os países situados entre as trinta primeiras posições, 16 manteriam o seu número de ordem, 13 alterariam até quatro posições, e um alteraria

cinco ou mais posições. Já para os trinta países seguintes (31ª até 60ª posição), apenas dois manteriam o seu posto, 17 alterariam até quatro posições, e 11 alterariam cinco posições ou mais. Para os países situados no relatório entre as posições 61 e 90, e 91 a 120, nenhum manteria o seu número de ordem; 11 e 7, respectivamente, em cada um dos grupos (61-90 ou 91-120), alterariam até quatro posições; e 19 e 23 alterariam cinco posições ou mais.

Por que se detecta essa não replicabilidade? Porque os resultados divulgados são oriundos de mil simulações para cada um dos cinco componentes do indicador composto, e para cada um dos 191 países. No final do processo de simulações, apenas as medidas pontuais foram divulgadas no Relatório. A OMS não divulgou nem disponibilizou na rede internacional de computadores os parâmetros que foram utilizados, nem as próprias mil simulações de cada nota de cada país. Só assim os pesquisadores poderiam ter acesso aos parâmetros utilizados e verificar (ou replicar) os resultados divulgados. Em outras palavras, a OMS divulgou um resultado pouco transparente e irreplicável sobre a avaliação dos serviços de saúde no mundo, pois não foram disponibilizados os parâmetros necessários para suas estimativas.

Como terceiro problema, poderia ser mencionado o excesso de precisão utilizado nas estimativas do desempenho dos sistemas de saúde. Por exemplo, mensurou-se a expectativa de vida ajustada por incapacidade no Japão com treze casas decimais, como sendo 74.5475712298641 anos de vida, o que significa dizer que a expectativa naquele país seria de 74 anos, 6 meses, 17 dias, 3 horas, 0 minutos, 55 segundos e 533 milésimos de segundo, e ainda deixaria resto.

Verifica-se também que a Holanda, com 72.0145769233091 anos de expectativa de vida ajustada pela incapacidade está melhor, em termos do nível de saúde resultado do desempenho de seus serviços de saúde, do que o Canadá, com 72.0189543173123 anos por exatamente 1 dia, 13 horas, 49 minutos e 14 segundos de diferença na expectativa de vida,

com uma precisão de 2 por cem milhões de anos de vida.

Apesar de a OMS ter alegado que esse excesso de precisão é, na verdade, um artefato do programa utilizado (Excel)², chamamos a atenção para a alteração de ranqueamento de alguns países, quando se arredondam os valores relativos à expectativa de vida ajustada por incapacidade. Mas esse fato só pode ser detectado na classificação produzida a partir do cálculo do indicador composto utilizando os valores pontuais divulgados pela OMS e aplicando-se as fórmulas recomendadas pelo WHR-2000. No documento mencionado há pouco, a OMS afirma que não há variação entre os países, avaliada através das mil simulações, o que não se pode verificar, pois, como já foi mencionado, não estão disponibilizados os dados para a comunidade científica. Porém, se calcularmos os pontos totais que cada país obteve no índice composto, assim como o respectivo ranqueamento de cada país, utilizando as treze casas decimais ou arredondando para três casas, alguns países teriam diferentes ranqueamentos nos dois métodos mencionados.

Nas comparações, realizadas segundo métodos diferentes, não foi possível utilizar o ranqueamento divulgado pela OMS, por não estarem disponíveis as informações sobre as simulações. Foram sempre comparadas as classificações obtidas através da aplicação das fórmulas sugeridas no WHR-2000, com os valores pontuais divulgados com e sem o procedimento que se está mencionando.

Como quarto problema que justifica menção, observa-se que a incerteza final no ranqueamento é menor do que as várias incertezas para cada componente. O Brasil, ficou na 125ª posição no ranqueamento, segundo o Relatório divulgado em 2000. O intervalo de incerteza sobre esta classificação obtida para o Brasil, divulgado também no mencionado relatório, variava entre a 118ª até a 133ª posição. Acontece que, quando se calcula o desempenho global dos serviços de saúde no Brasil, ou o indicador composto, observa-se que o nosso país ocuparia a

135ª posição, isto é, estaria fora do intervalo de incerteza. Ou seja, o próprio cálculo direto da posição do Brasil, utilizando-se os valores pontuais que a OMS divulgou, a partir das simulações realizadas, e a simples aplicação das fórmulas sugeridas, deixou o país em uma posição fora do intervalo de incerteza divulgado no Relatório, que teria sido obtido, acreditamos, a partir das simulações realizadas.

O quinto problema inclui uma alteração no ranqueamento em decorrência da modificação na escala de transformação para a expectativa de vida ajustada pela incapacidade. Para se obter o indicador composto, foi necessário que todas as notas tivessem sido transformadas em uma escala de zero a um. O WHR-2000 propôs que, para que a expectativa de vida ajustada pela incapacidade pudesse ser transformada em uma nota de zero a um, se considerasse 20 anos como a menor expectativa e 80 anos como a máxima expectativa possível. Se se considerar a amplitude da expectativa de vida ajustada pela incapacidade como sendo de 22 até 80 anos, ao invés de 20-80 anos, 13 (6,8%) países dentre os 191 alterariam sua posição no ranqueamento em até duas posições.

Murray *et al.*³ referem, em um documento divulgado em 2000, que *nem o ranqueamento nem o indicador composto eram sensíveis à variação na escolha dos pesos*, mais ainda que *os resultados na figura 5 demonstram firmemente que o indicador composto é robusto a grandes variações nos pesos atribuídos aos cinco componentes*. Os autores mencionam que os pesos obtidos em um inquérito que conduziram indicaram uma preferência de se utilizar um peso de 0,24 para o nível de saúde (expectativa de vida ajustada por incapacidade), de 0,25 para a distribuição da saúde (mortalidade na infância), de 0,13 para o nível de responsividade, 0,16 para a distribuição da responsividade e, finalmente, 0,22 para a distribuição da justiça na contribuição social.

Para que a definição do índice composto fosse mais fácil de entender, os autores mencionam, no já referido documento, que *os resultados desses inquéritos foram arredon-*

dados para o um oitavo mais próximo. Assim, foram sugeridos 0,250, para o nível de saúde, 0,250, para a distribuição da saúde, 0,125, para o nível de responsividade, 0,125, para a distribuição da responsividade, e 0,250 para a distribuição da justiça na contribuição social. Os resultados obtidos pelas simulações, sem dúvida que na dependência dos parâmetros de variação utilizados, podem não sofrer influência dos pesos, mas tal não é verificado nos resultados calculados a partir dos valores pontuais divulgados pela OMS. Por exemplo, quando se deseja dar um pouco mais de peso à desigualdade em saúde do que ao próprio nível de saúde, utilizando-se 0,24 para o nível de saúde e 0,26 para a distribuição, 60 (31,4%) países mudam o seu ranqueamento.

Entretanto, se aumentássemos o peso para o nível de saúde para 0,30, e o diminuíssemos para a distribuição de saúde para 0,20, 126 (66,0%) países alterariam seus ranqueamentos, em até oito posições, com 41 (21,5%) países alterando pelo menos três posições. Vamos agora ignorar a questão das ponderações, simplesmente somando todas as cinco notas, independentemente de pesos, e portanto atribuindo 0,20 de peso a cada um dos componentes do indicador composto. Nesse caso, 167 (87,4%) países alterariam o seu ranqueamento em até 17 posições, com 47 (24,6%) países modificando pelo menos cinco posições no seu ranqueamento.

O sétimo problema refere-se ao fato de as notas de cada indicador terem apresentado médias diferentes entre si. Por que haveria diferenças nas médias dos vários componentes? Porque talvez a Humanidade estivesse em estágios de desenvolvimento diferentes em algumas áreas em relação a outras. Ou seja, pode haver pouco a fazer em um domínio que esteja sendo avaliado no mundo, e muito a ser ainda realizado em outro domínio. Se tal fosse o caso, seria esperado de fato médias diferentes nos vários domínios avaliados no desempenho global dos serviços de saúde. E os dados publicados no WHR-2000 mostram exatamente essa diferença nas médias dos indicadores.

De fato, o nível de saúde no mundo apre-

sentou uma média de 61,4; a distribuição de saúde, uma média de 76,3; a responsividade apresentou uma média de 52,7 para o nível e 90,3 para a distribuição; já a justiça na contribuição social apresentou uma média de 91,7 pontos. Como os próprios autores mencionam no documento já citado³, a ocorrência de médias diferentes significa que *países na amostra estão relativamente mais longe do seu valor teórico mínimo para o índice com maior valor médio*. Ou seja, considerando-se os resultados encontrados, vale dizer que muito ainda precisa ser feito no mundo para melhorar o nível da saúde (média de 61,4), mas que sua distribuição (média de 76,3) está melhor do que o próprio nível em si encontrado. Ou que muito precisa ser feito pelo nível de responsividade (média de 52,7), mas que sua distribuição, com uma média de 90,3, já está muito boa no mundo, visto se situar bastante próxima do valor teórico máximo (no caso 100). Na verdade, a distribuição da contribuição social obteve a maior média, 91,7, indicando que ser este o parâmetro onde o mundo está mais “avançado”, ou mais próximo do valor teórico máximo.

Primeiro, qual seria esse *valor teórico máximo* e como se calcularia? Segundo, matematicamente, o que se está fazendo é dar um peso maior (pois a média é maior) exatamente aos pontos onde o mundo está “mais desenvolvido”, em relação a um valor teórico máximo, e menos ponderação aos pontos onde ainda há muito a ser feito. Como se espera que os países que tenham atingido os valores mais altos sejam os países mais desenvolvidos em termos de serviços de saúde, esse método de adição das cinco notas para a obtenção do índice composto tenderia a privilegiar, na dependência também da variância dos indicadores, exatamente os países mais desenvolvidos, em oposição a um método que ajustasse as médias.

Para se reduzir todas as notas para a mesma média, teria sido necessário utilizar outros pesos que não aqueles que foram realmente empregados. Significa dizer que, pela influência das diferenças de médias, foi como se tivessem sido utilizados os pesos de

20,41 para o nível de saúde; de 25,36 para a distribuição da saúde; de 8,76 e 15,00 para o nível e distribuição da responsividade, respectivamente; e um peso de 30,48 para a justiça na contribuição social. São pesos bem diferentes daqueles mencionados no WHR-2000.

Empregando-se uma técnica de padronização para reduzir o efeito das diferenças das médias, por exemplo, calculando-se os z-escores das notas de cada componente de cada país, e somando-se depois os escores z, ponderadamente, pelos pesos sugeridos no WHR-2000, observa-se que apenas 18 países manteriam sua posição no ranqueamento, com 173 (90,6%) países movendo-se para cima ou para baixo em até 32 posições, dentre os quais 51 (26,7%) países modificariam, no mínimo, nove posições no seu ranqueamento.

Por outro lado, caso se utilizasse os mesmos escores z, mas sem ponderação, o que significa dizer que os cinco componentes seriam ponderados pelo mesmo peso de 1/5 ou 0,20, apenas 12 países manteriam o seu ranqueamento, com 179 (93,7%) países alterando o seu ranqueamento em até 33 posições, sendo que 74 (38,7%) países alterariam, no mínimo, nove posições.

O oitavo problema encontrado refere-se a alguns resultados estatísticos encontrados para as estimativas pontuais divulgadas no WHR-2000, para as quais não há aparentemente explicação razoável imediata, e sobre as quais a OMS não se pronunciou no documento citado, em resposta ao que a Delegação Brasileira apresentou para o Comitê Executivo da OMS, em Genebra, em janeiro de 2001.

Ao se utilizar a regressão linear múltipla, tendo, como variável resposta, a nota calculada para cada país-membro para o indicador composto, observa-se um coeficiente de determinação de 1,00. Na análise de variância, obtém-se uma soma dos quadrados dos resíduos insignificante (0,0209), uma estatística *f* muito elevada (46220382,72), coeficientes beta muito próximos dos pesos utilizados para o cálculo, 0,24994, 0,24996, 0,12501, 0,12525 e 0,24978, respectivamente, para ní-

vel e distribuição de saúde, nível e distribuição de responsividade, e justiça da contribuição social. Os valores dos f parciais são bem semelhantes entre si: 8,5, 5,0, 6,3, 7,0 e 4,4, respectivamente. O intercepto é quase zero (0,00222).

Quando se utiliza uma regressão linear múltipla, tendo como variável resposta o valor calculado do indicador composto publicado no WHR-2000 a partir das mil simulações para cada país, alguns resultados são bastante discordantes e difíceis de explicar. O coeficiente de determinação é bem próximo ao anterior (0,98), porém os resíduos são bem mais elevados (685,87) e a estatística f bem menor (1523,51). Os coeficientes beta não correspondem aos pesos utilizados, tendo sido encontrados valores de 0,2351, 0,2944, 0,2130, 0,0364 e 0,2554, respectivamente. Os valores dos f parciais são bem diferentes entre si: 230,4, 210,7, 58,9, 1,8 e 141,3, respectivamente. O intercepto não se aproxima de zero (-1,862). A distribuição da responsividade não mostra, assim, associação estatística independente com o indicador composto.

Repetindo-se essas regressões apenas para os 32 países da América Latina e Caribe, os resultados são ainda mais inquietantes. Tomando-se, inicialmente, o resultado calculado para o indicador composto de cada país, observa-se que o coeficiente de determinação foi de 1,00. De novo, a soma dos quadrados dos resíduos foi mínima (0,0033) e a estatística f foi muito grande (1739231,39); os coeficientes beta aproximaram-se muito dos pesos utilizados (0,2501, 0,2502, 0,1248, 0,1254 e 0,2498) e os valores dos f parciais foram relativamente próximos e bem elevados (44000, 43000, 53854, 92423, 61000, respectivamente); e o intercepto, mais uma vez, aproximou-se de zero (-0,026).

Entretanto, ao se rodar a regressão linear múltipla apenas para os 32 países da América Latina e Caribe, e tomando-se como variável resposta a nota do indicador composto divulgada no próprio WHR-2000, como resultante das mil simulações, observa-se, inicialmente, um coeficiente de determinação de 0,94 e uma soma dos quadra-

dos dos resíduos mais elevada (56,5), com uma estatística f bem menor (87,0) do que aquela encontrada na regressão anterior. Os coeficientes beta não têm qualquer relação com os pesos utilizados no cálculo do índice composto, pois foram obtidos valores de 0,1052, 0,4689, 0,1177, -0,0448, e 0,1974, respectivamente, tendo sido, inclusive, negativo o valor para a distribuição da responsividade. O intercepto está longe do zero (10,96) e os valores dos f -parciais são muito heterogêneos, 4,6, 88,7, 2,8, 0,7 e 22,1, indicando, inclusive, não significância na associação do índice composto com o nível de saúde nem com o nível nem com a distribuição da responsividade. Mesmo forçando-se o intercepto para zero, não há mudança importante nos resultados obtidos.

Observando-se agora a matriz de correlação do índice composto e seus cinco componentes, observa-se que a justiça na contribuição social apresentou correlações bem mais baixas do que os demais componentes, tanto com o próprio índice composto quanto com os demais componentes (Tabela 1).

Atente-se, também, que a regressão linear múltipla, tendo como variável resposta o índice composto divulgado no WHR-2000, e como variáveis independentes o logaritmo natural do produto interno bruto (PIB) per capita e o coeficiente gini, apresentou um coeficiente de determinação de 0,87, com f parciais de 488 e 28,6, respectivamente, para PIB per capita e coeficiente gini. Se se acrescentam a expectativa de vida ao nascer (no sexo masculino) e a proporção de gastos públicos com a saúde, obtém-se um coeficiente de determinação de 0,99. Esse resultado indica, com clareza, que o que se está divulgando como *desempenho dos serviços de saúde* é, na verdade, resultado da distribuição da renda, da velha expectativa de vida (não ajustada pela incapacidade) e da renda per capita de um país, e não um resultado, direto ou indireto, da atuação do sistema de saúde dos países-membros da OMS.

Finalmente, reforça-se mais uma vez a necessidade de um organismo internacional do porte e da importância da Organização Mundial de Saúde divulgar os resultados das

Tabela 1 - Correlação entre os vários componentes e o índice composto. Avaliação do desempenho dos sistema de saúde no mundo. WHR-2000.

Table 1 - Correlation between various components and the compound rate. Evaluation of the performance of Health Systems around the world. WHR - 2000.

	Indicador Composto	Nível de Saúde	Distribuição da Saúde	Nível da Responsividade	Distribuição da Responsividade	Distribuição da Contribuição Social
Indicador Composto	1,000					
Nível de Saúde	0,939	1,000				
Distribuição da Saúde	0,948	0,890	1,000			
Nível de Responsividade	0,855	0,761	0,802	1,000		
Distribuição da Responsividade	0,832	0,782	0,805	0,771	1,000	
Distribuição da Contribuição Social	0,433	0,282	0,282	0,340	0,325	1,000

Fonte/Source: WHR-2000.

simulações que foram realizadas, para que se possa entender com mais clareza os procedimentos aplicados, pois um dos pressupostos básicos da Ciência é a possibilidade de replicação dos resultados encontrados por um pesquisador.

Concluindo, seria ainda importante chamar a atenção para alguns pontos:

- o indicador composto implica, na verdade, em um excesso de reducionismo, tentando reduzir a mensuração da saúde e o complexo desempenho do sistema de saúde de um país a cifras únicas, o que não parece ser útil;
- na proposição teórico-conceitual desse novo indicador, ignorou-se toda a produção científica anterior da área de avaliação de serviços de saúde;
- na construção do modelo teórico proposto, foram utilizados e publicados documentos internos da OMS, sem revisão crítica por outros pesquisadores de áreas correlatas, inclusive com divulgação limitada, e algumas vezes até sem qualquer divulgação, pois não se pode obtê-los na *home page* da OMS;
- não estavam disponíveis, no mundo, dados para as estimativas necessárias nem para a construção do indicador composto;
- não se ouviu os países em relação às estimativas que foram produzidas;
- apesar de ter sido um exercício acadê-

mico interessante, os muitos cálculos realizados mostraram-se impróprios para uma Congregação de Países e, principalmente, para serem utilizados pelos Governos no acompanhamento da evolução do desempenho dos seus serviços de saúde;

- os resultados mostraram-se como não-replicáveis, o que contraria o princípio científico básico e essencial para o avanço da Ciência;
- o indicador composto mostrou-se sensível ao re-escalonamento de um de seus componentes;
- o indicador composto é sensível à mudança de pesos atribuídos a cada um dos seus componentes;
- a precisão utilizada nos cálculos das estimativas foi incompatível com a realidade dos dados disponíveis para aquelas estimativas;
- a variabilidade final não corresponde ao nível de precisão possível de se obter com os dados existentes;
- os dados não sofreram adequado processo de standardização que remove-se as influências específicas das diferenças de avanço que podem estar ocorrendo em cada um dos diversos componentes utilizados para o cálculo do indicador composto;
- os resultados obtidos nas análises multivariadas e nas correlações não são

compatíveis com as simulações realizadas;

- não se tinha precisão suficiente para ranquear todos os países do mundo, um a um;
- poderiam ter sido definidos níveis de desenvolvimento para os sistemas de saúde dos vários países, ao invés do ranqueamento de todos os países-membros da OMS, por exemplo, decis ou

quintis, dentro dos quais não se tivesse obtido um valor sintético único (índice composto) e, portanto, não teria sido viável hierarquizar os países em nenhum tipo de ranqueamento;

- a obsessão pelo ranqueamento dos países obscureceu o grande avanço de se tentar avaliar os sistemas de saúde no mundo de uma maneira mais global e sistematizada.

Referências

1. World Health Organization. Health Systems: Improving performance. World Health Report. Geneva; 2000.
2. World Health Organization. The Methods and Data used in the World Health Report 2000: A response to the commentary by the Brazilian Delegation to the Executive Board, 17th and 19th January 2001. Geneva; 2001.
3. Murray C, Christopher JL, Lauer J, Tandon A & Frenk, J. Overall Health System Achievement for 191 countries. Discussion Paper Series no. 28. EIP/GPE, World Health Organization, 2000.