

# An evidence-based approach to benchmarking the fairness of health-sector reform in developing countries

Norman Daniels,<sup>1</sup> Walter Flores,<sup>2</sup> Supasit Pannarunothai,<sup>3</sup> Peter N. Ndumbe,<sup>4</sup> John H. Bryant,<sup>5</sup> T. J. Ngulube,<sup>6</sup> & Yuankun Wang<sup>7</sup>

**Abstract** The Benchmarks of Fairness instrument is an evidence-based policy tool developed in generic form in 2000 for evaluating the effects of health-system reforms on equity, efficiency and accountability. By integrating measures of these effects on the central goal of fairness, the approach fills a gap that has hampered reform efforts for more than two decades. Over the past three years, projects in developing countries on three continents have adapted the generic version of these benchmarks for use at both national and subnational levels. Interdisciplinary teams of managers, providers, academics and advocates agree on the relevant criteria for assessing components of fairness and, depending on which aspects of reform they wish to evaluate, select appropriate indicators that rely on accessible information; they also agree on scoring rules for evaluating the diverse changes in the indicators.

In contrast to a comprehensive index that aggregates all measured changes into a single evaluation or rank, the pattern of changes revealed by the benchmarks is used to inform policy deliberation about which aspects of the reforms have been successfully implemented, and it also allows for improvements to be made in the reforms. This approach permits useful evidence about reform to be gathered in settings where existing information is underused and where there is a weak information infrastructure.

Brief descriptions of early results from Cameroon, Ecuador, Guatemala, Thailand and Zambia demonstrate that the method can produce results that are useful for policy and reveal the variety of purposes to which the approach can be put. Collaboration across sites can yield a catalogue of indicators that will facilitate further work.

**Keywords** Health care reform; Benchmarking; Social justice; Social responsibility; Health services accessibility; Evidence-based medicine; Program evaluation/methods; Developing countries; Cameroon; Ecuador; Guatemala; Thailand; Zambia (*source: MeSH, NLM*).

**Mots clés** Réforme domaine santé; Banc mesure performance; Justice sociale; Responsabilité sociale; Accessibilité service santé; Médecine factuelle; Evaluation programme/méthodes; Pays en développement; Cameroun; Equateur; Guatemala; Thaïlande; Zambie (*source: MeSH, INSERM*).

**Palabras clave** Reforma en atención de la salud; Benchmarking; Justicia social; Responsabilidad social; Accesibilidad a los servicios de salud Medicina basada en evidencia; Evaluación de programas/métodos; Países en desarrollo; Camerún; Ecuador; Guatemala; Tailandia; Zambia (*fuelle: DeCS, BIREME*).

Arabic

Bulletin of the World Health Organization 2005;83:534-540.

Voir page 539 le résumé en français. En la página 539 figura un resumen en español.

## Introduction

The Benchmarks of Fairness instrument is a method for evaluating the fairness of health-sector reforms. The concept of fairness in health systems is broad, integrating the goals of equity in access and financing, clinical and administrative efficiency, and

accountability. The Benchmarks address the complaint that “it is unfair” when the system treats some patients differently from others with similar needs, when some needs are not met because of administrative inefficiency, or when people have no say in how the system treats them. Fairness involves various claims about what people are owed as a matter of justice (1–3).

<sup>1</sup> Professor of Ethics and Population Health, Harvard School of Public Health, Department of Population and International Health, Building 1, Room 1104C, 665 Huntington Avenue, Boston, MA 02115, USA (email: ndaniels@hsph.harvard.edu). Correspondence should be sent to this author.

<sup>2</sup> Research Associate, Universidad de San Carlos de Guatemala, Guatemala City, Guatemala.

<sup>3</sup> Professor of Community Medicine, Centre for Health Equity Monitoring, Faculty of Medicine, Naresuan University, Phitsanulok, Thailand.

<sup>4</sup> Dean, Faculty of Medicine, University of Yaounde, Yaounde, Cameroon.

<sup>5</sup> Professor Emeritus, Aga Khan University, Karachi, Pakistan.

<sup>6</sup> Director, Centre for Health, Science, and Social Research, Lusaka, Zambia.

<sup>7</sup> Associate Professor, Kunming Medical College, Kunming, Yunnan, China.

Ref. No. 04-016428

(Submitted: 15 July 2004 – Final revised version received: 4 January 2005 – Accepted: 31 January 2005)

The Benchmarks ask by how much reforms improve or worsen aspects of fairness within the health sector nationally or subnationally. They combine an ethical framework with familiar methods from operations research. Using appropriate indicators, changes are measured and evaluated relative to a baseline (the status quo at the time reforms are introduced). This evidence and evaluation enhances deliberations about improving reforms. The Benchmarks serve a different purpose than indices that aggregate all these changes into a single number for ranking health systems comparatively.

The Benchmarks aim to fill two crucial gaps. First, internationally supported reforms over two decades have failed to integrate the key goals of fairness. During the 1980s and early 1990s, international agencies such as the World Bank and International Monetary Fund pushed reforms that involved privatization, user fees, and decentralization, thus sacrificing equity (4–7) and other goals of fairness to the market-oriented pursuit of efficiency. Recently, the World Bank's attention has focused on health inequalities (8, 9) and better governance (10, 11) but has still failed to integrate these goals. Second is the lack of capacity in developing countries to undertake evidence-based policy analysis which is the result of weak information infrastructures and a lack of tools to maximize the use of existing information.

## Development and rationale

Teams from Colombia, Mexico, Pakistan and Thailand, using their own recent reforms as case studies, adapted a matrix for assessing the fairness of proposed American health insurance reforms (1, 12, 13) into a generic developing-country framework (14). Despite different cultural and social histories and levels of development in the collaborating sites, teams were able to agree on a generic matrix that included nine main Benchmarks (Box 1). Each Benchmark specifies a key objective of fairness through criteria that capture important elements and means of achieving these objectives (Box 2).

The nine generic Benchmarks integrate the goals of fairness as follows: Benchmarks 1–5 address equity, 6 and 7 consider efficiency, and 8 and 9 concern accountability (Box 1). Improving clinical and administrative efficiency can make a system fairer by allowing it to meet otherwise unmet needs. Though not all conflicts between efficiency and equity disappear, a fair process (included in Benchmark 8) can resolve disputes about them (15). Accountability is valued both intrinsically, as a matter of fairness in governance, and instrumentally, since it helps achieve efficiency or equity.

## Contrast with WHO framework for health-systems performance

The Benchmarks serve a different purpose from WHO's index of health systems performance (16, 17). The WHO index combines measures of health outcomes, system responsiveness, and the distribution of financial contributions into an overall index that ranks countries' performance comparatively. The dimensions measured are assigned weights by surveying appropriate experts. A commission evaluating the index proposed various methodological improvements (18–21).

In contrast, the Benchmarks make no cross-country comparisons, avoiding some methodological issues about weighting and aggregation that face the WHO index. Because there may

be various fair ways to weight and trade off changes within and across benchmarks, we propose that the different benchmarks and their component criteria should not be weighted in terms of their relative importance. The Benchmarks instead aim to reveal the complex pattern of the effects of reforms on different aspects of fairness. Understanding this pattern is more informative for local decision-makers than compressing information into an index. Changes in the various dimensions of fairness measured by the Benchmarks can still be evaluated.

## Methods: developing evidence to guide reform

The generic benchmarks and criteria must be adapted to serve a specific purpose by an interdisciplinary team. The team refines the generic criteria, specifies indicators appropriate to local conditions, and seeks agreement on how to evaluate changes in these indicators. Planners or community groups can then evaluate health policies in light of the evidence they have agreed is relevant. Under ideal conditions, policy-makers assessing reforms would benefit from systematically reviewed evidence (from natural or controlled social experiments, were they available) (22, 23). Developing countries implementing reforms need good evidence, based on local information, about the varied effects of actual reforms (24), and this is what the Benchmarks provide.

An ideal interdisciplinary team consists of policy-makers, academics, health-systems personnel, clinicians and civil society groups. The breadth of the Benchmarks compels people who have different training and work at various levels in the system to cross disciplinary boundaries and reconcile their perspectives. The adaptation process encourages stakeholders to take ownership of the results.

The team must consider the purpose of the application: for example, whether it is evaluating comprehensive reforms (e.g., Mexico (25)), measuring district variation in the implementation of reforms (e.g., Cameroon (P. Ndumbe, unpublished data, 2005) and Thailand (S. Pannarunothai et al., unpublished data, 2005)), measuring the impact of decentralization and financing reforms on delivery of public health services (Ecuador and Guatemala), evaluating the equity effects of rural insurance programmes (Yunnan, China), or the effects on the health sector of scaling-up AIDS treatment and prevention programmes (Zambia). The purpose determines how the team will modify the generic benchmark criteria.

Box 1. The nine main Benchmarks of Fairness and their corresponding key objective of fairness

Benchmark	Objective of fairness
B1 Intersectoral public health	Equity
B2 Financial barriers to equitable access	
B3 Non-financial barriers to access	
B4 Comprehensiveness of benefits and tiering	
B5 Equitable financing	
B6 Efficacy, efficiency and quality improvement	Efficiency
B7 Administrative efficiency	
B8 Democratic accountability and empowerment	Accountability
B9 Patient and provider autonomy	

Box 2. Examples of main criteria and their corresponding benchmarks

Benchmark	Selected main criteria
B1 Intersectoral public health	<ol style="list-style-type: none"> <li>1. Degree to which reform increases percentage of population (demographically differentiated) benefiting from basic nutrition, clean water, literacy</li> <li>2. Development of information infrastructure for monitoring inequalities in health status</li> <li>3. Reform engages in intersectoral efforts</li> </ol>
B3 Non-financial barriers to access	<ol style="list-style-type: none"> <li>1. Reduction in geographical maldistribution of services, supplies, etc.</li> <li>2. Gender issues affecting access: degree of reproductive autonomy, access to resources, authority in family regarding decisions, sensitivity of services</li> <li>3. Cultural barriers: language, attitudes towards disease</li> <li>4. Discrimination (race, religion, class, sexual orientation, disease stigma)</li> </ol>
B8 Democratic accountability and empowerment	<ol style="list-style-type: none"> <li>1. Explicit, public procedures for evaluating services</li> <li>2. Explicit, deliberative procedures for resource allocation with transparency and stakeholder involvement in agreement on rationales</li> <li>3. Fair appeals procedures</li> <li>4. Adequate protections for privacy and patients' rights</li> <li>5. Measures to strengthen civil society</li> </ol>

In our case, the Guatemalan team ignored some benchmarks, combined features of others, and developed criteria that concentrated on the delivery of public health services. Familiarity with the specifics of local reform mechanisms helps in selecting criteria. Developing criteria for improvements in efficiency or accountability requires knowing how community structures are involved in health-unit management (as in Cameroon or Zambia) (Box 3).

The team must pay attention to the sources and quality of information used for different indicators. There often is good data for traditional indicators bearing on some measures of health outcomes and some kinds of utilization, but indicators bearing on accountability, intersectoral cooperation, and quality measures, require non-standard sources of information or the use of qualitative techniques.

Teams must define an approach to evaluating changes from a measured baseline. Most of our sites have not completely addressed this evaluative task, concentrating instead on developing an instrument capable of measuring a baseline. Most teams plan to develop scoring rules that assign a qualitative value to a degree of measured change. This evaluation might be explicitly qualitative (a scale of poor, fair, good, excellent or colour-coded stars); alternatively, a team may use ordinal numbers on a scale (e.g., from -5 to 5), provided it avoids using these ordinals for arithmetic operations like averaging or aggregating.

Evaluation requires deliberation. To illustrate: some team members might think that if 50% of the population lacked access to clean water, then each additional 10% of the population that gained access should count for an additional point on an ordinal scale ranging from 0 to 5. Others might want to give weight to the difficulty of reaching remote populations and propose a different scoring rule. Whatever rule the team agrees on operationalizes its view of fairness on this dimension. Some other team or group may disagree. There is no effort to capture a general public or expert consensus since the goal is not a universal ranking of fairness. Nevertheless, the result is this kind of objectivity: disagreements focus on the evidence about change and the reasons for considering some amount of change as being more or less fair. This approach provides a proper basis for evidence-based evaluation and promotes better deliberation.

In Thailand (see below) the problem of evaluation was approached differently. Instead of developing scoring rules, diverse focus groups were informed about changes in selected indicators and judgements about the fairness of these changes were elicited. This approach engages a wider group in evaluation than the use of scoring rules, but it is less clear how evaluations are connected to specific changes. As country-level work proceeds, we shall have to grapple with the strengths and weaknesses of these different approaches to evaluation.

### Applying the locally-adapted benchmarks

In some middle-income developing countries, the application of the Benchmarks made robust use of locally available information of reasonable quality that had not been used previously to inform ongoing reform efforts (Ecuador, Guatemala, Mexico, Thailand). In low-income developing countries, they provide useful evidence even if crude proxies have to be substituted for better quality indicators that are not locally available (for example, in Cameroon, Yunnan Province in China, and Zambia).

#### Guatemala

A team of academics, members of the ministry of health, and a nongovernmental organization (NGO) (CARE) adapted the Benchmarks to assess reforms aimed at decentralizing, financing and improving access to public health services. To measure equity, the team developed three indices that draw on good quality data available at the district level. An Index of the Priority of Health-Care Services combines the coverage for three basic services (immunization, antenatal care and supervised deliveries). The larger the gap between actual and ideal coverage — that is, the greater the need for resources — the higher the coefficient (between 0.01 and 0.99). The Index of Resource Distribution (IRD) is a weighted measure of health expenditure for primary care, health personnel and health facilities, using districts with the most resources for comparison with other districts. The higher the score (between 0.01 and 0.99) on the IRD, the more resources per capita are available in a district. An index of human resources measures the availability of community volunteers at the district level.

## Box 3. Examples of locally-developed indicators for adapted benchmark criteria

Benchmark and criterion <sup>a</sup>	Country that developed indicator and definition of indicator
B1 1. Degree to which reform increases percentage of the population (demographically differentiated) benefiting from basic nutrition, clean water, literacy, etc.	Cameroon: percentage of health districts monitoring iodized salt in markets
B3 1. Reduction in geographical maldistribution of services	Cameroon: percentage of health units with more than 10% of population living more than 1 hour's walk from health unit
B3 1. Reduction in geographical maldistribution of services	Guatemala: district ratio of index of priority for immunization coverage (IPHS) to index of resource distribution (IRD) <sup>b</sup>
B3 3. Language barriers	Guatemala: percentage of health personnel by category who speak an indigenous language
B8 5. Measures to strengthen civic society	Cameroon: percentage of health units with budget approved through community dialogue

<sup>a</sup> See reference 14 for a full listing of the Benchmarks and their criteria.

<sup>b</sup> See discussion of Guatemala in text for further details.

Equity requires that districts with greater unmet need receive more resources. Fig. 1 (web version only, available at: <http://www.who.int/bulletin>) shows the mismatch in coefficients for districts. The ministry hoped to rely on community health workers to compensate for the lack of resources in higher priority districts, but the results showed that high priority districts also had worse scores on the index of human resources. Other indicators showed variations among districts in the degree to which patients had access to personnel who spoke their native language (Box 3), revealing inequities in a non-financial barrier to services. Ministry participants, impressed with these results, plan to use a benchmarking approach to monitor the implementation of recent reforms.

### Ecuador

Learning from the experience in Guatemala, the Ecuador team built a coalition that included representatives from the provincial level of the ministry of health, local government, NGOs and civil society organizations, in order to better sustain pressure on the ministry and local government to pursue fairness in reforms. Drawing on and adding to the Guatemalan indicators, the team concentrated on evaluating the ministry of health's policy of providing free care to mothers and children younger than 5 years of age by reimbursing public facilities for their care. The Benchmark on financial barriers to care showed that health facilities in 4 of 6 districts surveyed charged patients US\$ 1–4.00 for treating an acute respiratory infection. Acute respiratory infections were used as an indicator for services that were supposed to be free for the targeted women and children. These findings were presented to health authorities. Some of the civil society groups requested an investigation into the charges and that immediate action be taken to stop them.

To evaluate improvements in the quality of services for target groups, the team used the ratio of administrative staff to medical staff in health facilities, which is admittedly a crude indicator of efficiency and quality. On average, public facilities have 7 administrative staff for every 10 medical staff whereas facilities run by NGOs have a ratio of 3 to 10. Public hospitals varied considerably, with ratios ranging from 1 to 5 to 1 to 2. The coalition secured agreement from provincial authorities to study public staffing issues and modify policy where necessary.

### Thailand

A team of Thai researchers applied an adapted version of the Benchmarks to examine how different groups judged the fairness of reforms (S. Pannarunothai, unpublished data, 2005). They presented the results from 81 health-outcome and process indicators that measured the geographical variations in the effects of reform to eight focus groups in each of 10 provinces. The groups, differing by urban and rural location and further divided into groups concerned with health care and its delivery (managers, providers, advocates) and comparable groups not primarily concerned with health, were asked to evaluate the fairness of the changes. Non-health groups tended to give lower estimates of fairness than health groups. The approach revealed the instrument had a reasonable sensitivity to different levels of performance in the system, moving the ministry of health to plan to develop this approach for further use in evaluating health-sector reforms under decentralization.

### Mexico

A team in Mexico used the Benchmarks to evaluate retrospectively the fairness of changes from 1995, when reforms were introduced, until 2000. They had to restrict their study to indicators available at both times (25). As in other middle-income countries, the benchmarking approach used good quality, underutilized data in a novel way. A newer Mexican adaptation allows for ongoing monitoring and evaluation of cervical cancer screening programmes (26).

### Cameroon

In 2002, a workshop of district medical officers, ministry of health officials and academics adapted the Benchmarks to assess recent national reforms, assigning the task of refining indicators to a team of district medical officers. The adaptation focused on seven benchmarks and 70 criteria, for which 73 indicators were selected (see examples in Box 3). The restrictions were that indicators had to involve information that could be collected by medical students during their sixth-year district rotations, could not require special surveys, and could not put the students at risk. After training the students, two cycles of data collection were undertaken involving 70 medical students in 20 of 155 districts.

These field tests revealed errors in some indicators. Students could not, for example, collect information about testing

for the presence of iodized salt in local markets (Box 3) because it is available only at the provincial level. Field tests nevertheless showed that students, assisted by properly trained district medical officers could gather reliable evidence with the benchmarking approach, although appropriate surveys were also needed. The full dataset is being analysed to provide a baseline against which national reforms can be monitored. The use of students is not only low cost, but has educational value, since it produces a large number of doctors trained to assess the fairness of reforms. (A project in Pakistan also incorporates the Benchmarks into medical school curriculum.)

### Zambia

In June 2003, a workshop involving some members of parliament, representatives from the ministry of health and community members from four districts, proposed adapting the Benchmarks for use in monitoring the equity of new antiretroviral treatment and prevention programmes for HIV/AIDS.

In March 2004, a team of community representatives and health providers from four districts adapted the benchmarks to that purpose by constructing a table with four key questions about the effects of scaling-up treatments: what do you want to see get better? How would you judge or tell it was getting better? What information would you need to make that judgement? Where would you get that information? For example, the team wanted a fairer process to be used in selecting treatment sites and patients. They would be able to tell whether the process was fair if appropriate stakeholders were involved, and this information could be obtained from minutes of meetings at different levels.

A regional meeting in Malawi, drawing on the results of the Zambian workshop as well as on other proposals about monitoring the introduction of antiretroviral treatments in the southern African region, proposed additional indicators that will be incorporated into the Zambian work. These indicators included the percentage of treatments delivered at different facility levels within the health system (to measure integration in the system) and the percentage of treatments by gender compared with prevalence (to measure gender equity). Sharing benchmarking ideas across the region will accelerate the adaptation process.

### Yunnan Province, China

A team including academics from Kunming Medical School, various medical professionals, and health authorities used a benchmarking approach to measure the effects of a new rural insurance programme in selected districts and municipalities in Yunnan Province. They found significant financial barriers to enrolment of the poorest farmers, who could not afford a contribution of 10 yuan per year and who had limited access to subsidies that were supposed to reach the very poor. They found lower enrolments among better-off farmers, who thought the care being provided would not meet their needs. They found overcharging for drugs in some districts. This initially promising effort encountered an obstacle we have faced in a number of settings, namely, inadequate funding to conduct monitoring and evaluation on a sustained basis.

## Discussion: lessons learnt

Several important lessons can be drawn from these illustrations of the benchmarking approach despite the fact that it is a work in progress.

### Adapting to local purposes and local conditions

One important lesson is that interdisciplinary teams can agree on what changes make a system more or less fair in specific ways and adapt the generic Benchmarks to meet specific purposes and local conditions. Because such agreement involves deliberation about values and trade-offs among them, it is a key step towards using evidence to evaluate policy. Whereas the generic Benchmarks show that considerable agreement can be reached across cultural differences, the country-specific projects allow different sites to emphasize aspects of fairness in a way that reflects local deliberation.

Teams are resourceful in respecting constraints on local information yet using it fruitfully. Mexico's retrospective evaluation required restricting the analysis to data available at the inception of reforms. Cameroon selected indicators that could be collected at the district level by medical students. Zambia also restricted indicators to those that could be collected locally. Guatemala and Ecuador, middle-income countries with better information sources, developed more complex indices of available information that had not been used previously to evaluate reforms. One obstacle to the process is ensuring that teams know what information can be accurately collected locally. Cameroon's field tests revealed that some indicators had to be revised. Zambia proposed exit surveys that were too difficult. A strength of the adaptation process is that it builds capacity to understand the process of monitoring and evaluating reform: community members in Zambia and civic society groups in Ecuador mastered the task of deciding what should be improved and determining how to tell if it had improved.

### Process and use of results

Another central lesson of these applications is that stakeholders — the ministry of health, providers and civic society groups — take ownership of the adaptation primarily through their involvement in the process. When this is achieved, some of the political tensions that arise during evaluation can be overcome. For example, although the ministry of health staff involved in the Guatemalan project were initially resistant to making certain information available, when they realized that it would enable them to improve the delivery of services, they produced the information, participated in the analysis, and later took the results seriously. The team in Ecuador, aware of the Guatemalan experience, included representatives of civic society groups from the start in order to sustain pressure to use the results.

The adaptation process requires a significant investment of labour by a diverse team; it can take months, even a year, before it yields results. This factor is further complicated by the difficulty of securing funding for early-stage projects as well as for sustained monitoring and evaluation.

One compensating strength of the process is that it develops capacity to analyse and deliberate about improving reform efforts. It is also possible to learn across sites, as Ecuador did, improving on the process and using indices already developed in Guatemala. The work in Zambia was enhanced by having district medical officers from the Cameroon project help conduct the workshops. Sharing experiences from Zambia in the Malawi regional meeting advanced the work in several sites.

### Evaluation

Although these sites are at a relatively early stage of using the evidence acquired from benchmarking for evaluation, most teams plan to develop scoring rules. Thailand is exploring the

Norman Daniels et al.

use of surveys among focus groups. The methods underlying the evaluation of the Benchmarks are simpler and less controversial than the weighting and aggregation required in constructing cross-country indices. A clear statement of the strengths of different approaches must await further work in various sites.

### Conclusions

We conclude with two general points: first, cooperative research can more systematically share the lessons learnt across sites. We plan to assemble a network of researchers to create a catalogue of appropriate indicators for non-standard criteria and contexts in which information is poor. The catalogue will enable new sites to carry out adaptations more easily and effectively; it will be incorporated in a training manual posted on the Benchmarks web page Available from: <http://www.hsph.harvard.edu/benchmark/>).

Second, there is growing interest in this method at the local level. As a result of presentations at the ministerial summit in Mexico and at a WHO consultation on equity in reproductive health in Geneva in December 2004, Costa Rica and the

Philippines are both initiating Benchmarks projects in 2005. If funds can be secured, projects in Viet Nam and several other countries, as well as the southern African region, will move forward. The Benchmarks are attractive because they focus on building capacity to use locally available evidence to improve the fairness of systems whatever a country's starting point. ■

### Acknowledgements

We thank CARE, the Ministry of Health-Baja Verapaz (Guatemala), ALDES, and CLEPS (Ecuador) for support in data collection and data analysis. We also thank the Health Systems Research Institute, WHO/Cameroon and the University of Yaounde. We also thank Marin Levy for editorial assistance.

**Funding:** The authors thank the Rockefeller Foundation for generous, overall support of this project (research grant No. HE 070), as well as the European Union (research Grant No. ICA4-CT-2000-30037).

**Competing interests:** none declared.

## Résumé

### Evaluation, à partir de données factuelles, des performances en matière d'équité des réformes du secteur de la santé dans les pays en développement

Les Benchmarks of Fairness sont un instrument fondé sur des données factuelles, mis au point sous forme générique en 2000 pour aider les pouvoirs publics à évaluer les effets des réformes du système de santé sur l'équité, l'efficacité et la responsabilisation. Cette méthode, qui mesure tous ces effets sur l'objectif central qu'est l'équité, comble une lacune qui freine les réformes depuis plus de deux décennies. Des projets mis en œuvre ces trois dernières années dans des pays en développement de trois continents ont adapté la version générique de cette évaluation pour l'utiliser aux niveaux national et infranational. Des équipes interdisciplinaires de gestionnaires, prestataires de services, universitaires et personnes militantes conviennent des critères pertinents pour évaluer les éléments de l'équité et, selon les aspects de la réforme qu'ils souhaitent évaluer, choisissent des indicateurs appropriés fondés sur des informations accessibles ; ces équipes fixent également des règles de notation pour évaluer les divers changements concernant ces indicateurs.

Contrairement à un indice global regroupant tous les changements mesurés dans le cadre d'une seule évaluation ou d'un seul classement, le schéma des changements mis en évidence par les évaluations est utilisé pour aider les responsables à déterminer les aspects des réformes qui ont abouti, et pour améliorer les réformes. Cette méthode permet également de recueillir des données factuelles utiles sur la réforme, là où les informations existantes sont insuffisamment utilisées et où l'infrastructure du système d'information est faible.

Les brèves descriptions des premiers résultats obtenus au Cameroun, en Equateur, au Guatemala, en Thaïlande et en Zambie montrent que cette méthode peut donner des résultats utiles aux responsables politiques et révéler la diversité de ses utilisations possibles. Un catalogue d'indicateurs établi en collaboration par les différents sites facilitera les travaux ultérieurs.

## Resumen

### Enfoque basado en la evidencia para comparar la equidad de las reformas del sector sanitario en los países en desarrollo

El instrumento Criterios de Comparación de la Equidad es una herramienta de política basada en pruebas desarrollada de manera genérica en 2000 para evaluar los efectos de las reformas de los sistemas de salud en la equidad, la eficiencia y la responsabilización. Integrando las medidas de estos efectos en la meta central de equidad, este enfoque llena un vacío que ha obstaculizado los esfuerzos de reforma durante más de dos decenios. Durante los últimos tres años, diversos proyectos emprendidos en países en desarrollo de tres continentes han adaptado la versión genérica de esos criterios para uso tanto nacional como subnacional. Equipos interdisciplinarios de gestores, proveedores, universitarios y abogados acuerdan los criterios pertinentes para evaluar los componentes de equidad y, según los aspectos de las reformas

que deseen evaluar, seleccionan indicadores apropiados que dependen de la información accesible; también se ponen de acuerdo en torno a las reglas de puntaje para evaluar los cambios de los indicadores.

A diferencia de un índice global que agregase todos los cambios medidos en una sola evaluación o categoría, las pautas de diferencias que revelan los criterios se usan para orientar las deliberaciones de política encaminadas a determinar qué aspectos de las reformas se han aplicado satisfactoriamente, y permiten además introducir mejoras en las reformas. Esta manera de proceder permite reunir evidencia de utilidad sobre las reformas en los entornos caracterizados por un uso insuficiente de la información existente y por la debilidad de la infraestructura de información.

Una somera descripción de los resultados preliminares logrados en Camerún, Ecuador, Guatemala, Tailandia y Zambia demuestra que el método puede arrojar resultados de utilidad para las políticas y revela la diversidad de objetivos con que se

puede emplear el enfoque. Mediante la colaboración entre sitios, es posible obtener un conjunto de indicadores que facilite la realización de nuevos trabajos.

---

## Arabic

---

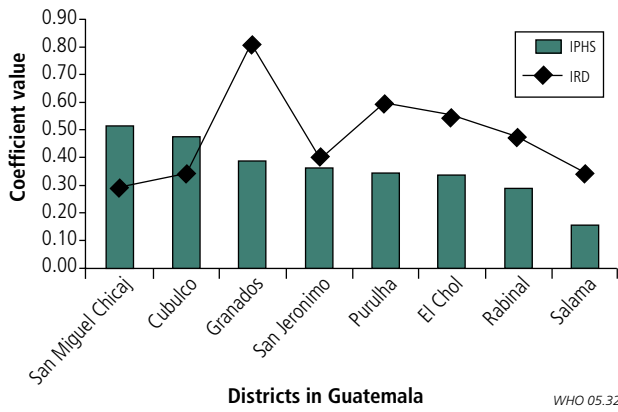


---

### References

- Daniels N, Light D, Caplan R. *Benchmarks of fairness for health care reform*. New York: Oxford University Press; 1996.
- Daniels N. *Just health care*. New York: Cambridge University Press; 1985.
- Daniels N. Justice, health and health care. *American Journal of Bioethics* 2001;1:3-15.
- Gilson L. *Implementing and evaluating health reform processes: lessons from the literature*. Washington, DC: US Agency for International Development, Partnerships for Health Reform Project; 1997 (Major Applied Research 1, Working Paper No. 1).
- Bennett S, McPake B, Mills A, editors. *Private health providers in developing countries: serving the public interest?* London: Zed Books; 1997.
- Berman PA, Bossert, TJ. *A decade of health sector reform in developing countries: what have we learned?* Data for Decision Making Project, International Health Systems Group, Harvard School of Public Health. Available from: <http://www.hsph.harvard.edu/ihsq/topic.html#5>
- Berman PA, editor. *Health sector reform in developing countries: making health development sustainable*. Cambridge (MA): Harvard University Press; 1995 (Harvard Series on Population and International Health).
- Evans T, Whitehead M, Diderichsen F, Bhuiya A, Wirth M. *Challenging inequities in health: from ethics to action*. New York: Oxford University Press; 2001.
- Roberts MJ, Hsiao W, Berman P, Reich MR. *Getting health reform right: a guide to improving performance and equity*. New York: Oxford University Press; 2004.
- Loewenson R. *Participation and accountability in health systems: the missing factor in equity?* (Equinet Discussion Paper No. 1); 2000. Available from: <http://www.equinet africa.org/bibl/docs/DIS1gov.pdf>
- World Bank Group, Public Sector Group. *Reforming public institutions and strengthening governance: a World Bank strategy*. Washington DC: World Bank; 2000. Available from: <http://www.worldbank.org/publicsector/strategy.htm>
- Daniels N, Brock D. Ethical foundations of the Clinton Administration's proposed health care system. *JAMA* 1994;271:1189-96.
- Daniels N. Ethics and health care reforms: a global view. In: Bankowski Z, Bryant JH, Gallagher J, editors. *Ethics, equity and health care for all*. Geneva: CIOMS; 1998:86-94.
- Daniels N, Bryant J, Castano RA, Dantes OG, Khan KS, Pannarunothai S. Benchmarks of fairness for health-care reform: a policy tool for developing countries. *Bulletin of the World Health Organization June 2000*;78:740-50.
- Daniels N, Sabin JE. *Setting limits fairly: can we learn to share medical resources?* New York: Oxford University Press; 2002.
- World Health Organization. *The world health report 2000 – Health systems: improving performance*. Geneva: WHO; 2000.
- Murray CJL, Evans DB, editors. *Health systems performance assessment: debates, methods and empiricism*. Geneva: World Health Organization; 2003.
- Braveman P, Starfield B, Geiger HJ. World Health Report 2000: how it removes equity from the agenda for public health monitoring and policy. *BMJ* 2001;323:678-81.
- Navarro V. Assessment of the World Health Report 2000. *Lancet* 2001;356:1598-1601.
- Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. *Health Affairs* 2001;20:10-20.
- Anand S, Ammar W, Evans T, Hasegawa T, Kissimova-Skarbek K, Langer A, et al. *Report of the scientific peer review group on health system performance assessment*. Geneva: World Health Organization; 2002.
- Niessen L, Grijseels E, Rutten F. The evidence-based approach in health policy and health care delivery. *Social Science and Medicine* 2000;51:859-69.
- Davies P. *Is evidence-based government possible?* Available from: <http://www.policyhub.gov.uk/home/jerrylecture1202041.pdf>
- Mackenbach P. Tackling inequalities in health: the need for building a systematic evidence base. *Journal of Epidemiology and Community Health* 2003;57:162.
- Gómez-Dantes O, Gómez-Juaregui J, Indian C. La equidad y la imparcialidad en la reforma del sistema mexicano de salud [Equity and fairness in the reform of the Mexican health system]. *Salud pública de México* 2004;46:399-416.
- Gómez-Juaregui J, Daniels N, Reichenbach L. Cervical cancer screening program in Mexico: equity and fairness implications. Cambridge, MA: Harvard Center for Population and Development; In press.

Fig. 1. Index of Priority for Health-Care Services (IPHS) versus Index of Resource Distribution (IRD) for sample districts in Guatemala. Vertical bars are measures of priority: the higher the value, the greater the need for resources to fill a gap in coverage. Points represent the index of resource distribution, with higher values corresponding to greater allocation per district. The mismatch is a reflection of inequity in meeting needs



WHO 05.32