# Benchmarks of fairness for health care reform: a policy tool for developing countries

Norman Daniels,[1] J. Bryant,[2] R.A. Castano,[3] O.G. Dantes,[4] K.S. Khan,[5] & S. Pannarunothai[6]

Teams of collaborators from Colombia, Mexico, Pakistan, and Thailand have adapted a policy tool originally developed for evaluating health insurance reforms in the United States into ''benchmarks of fairness'' for assessing health system reform in developing countries. We describe briefly the history of the benchmark approach, the tool itself, and the uses to which it may be put. Fairness is a wide term that includes exposure to risk factors, access to all forms of care, and to financing. It also includes efficiency of management and resource allocation, accountability, and patient and provider autonomy. The benchmarks standardize the criteria for fairness. Reforms are then evaluated by scoring according to the degree to which they improve the situation, i.e. on a scale of –5 to 5, with zero representing the status quo. The object is to promote discussion about fairness across the disciplinary divisions that keep policy analysts and the public from understanding how trade-offs between different effects of reforms can affect the overall fairness of the reform. The benchmarks can be used at both national and provincial or district levels, and we describe plans for such uses in the collaborating sites. A striking feature of the adaptation process is that there was wide agreement on this ethical framework among the collaborating sites despite their large historical, political and cultural differences.

**Keywords:** health care reform; health services accessibility; benchmarking, standards; health policy; developing countries; United States.

*Voir page 748 le résumé en français. En la página 749 figura un resumen en español.*

## A new tool for policy analysis

We report here on progress towards developing the benchmarks of fairness (*1*) into a policy tool that will be useful in developing countries for analysing the overall *fairness* of health care reforms. Fairness is a many-sided concept, broader than the concept of equity (*2–4*). Fairness includes equity in health outcomes, in access to all forms of care and in financing. Fairness also includes efficiency in management and allocation, since when resources are constrained their inefficient use means that some needs will not be met that could have been. For the public to have influence over health care, fairness must also include accountability. Finally, fairness also includes appropriate forms of patient and provider autonomy. The benchmarks help the integrated examination of objectives that often involve trade-offs with each other, which requires looking across disciplinary boundaries in a systematic way.

When originally developed and presented in the United States, the benchmarks had an ethical rationale that appealed to a theory of justice and health care (*1, 5*). The central thought is that disease and disability reduce the opportunities open to individuals, and that the principle of equal opportunity provides a basis for regulating a health care system. The same theory can also be extended to look beyond the point of delivery of health care to the social determinants of health (*6*).

The objection might be raised that this liberal democratic, rights-based account is too culturally limited to provide an international framework for the benchmark approach. Nevertheless, in our work in four developing country sites, which differ considerably in their political, cultural and religious backgrounds, we found a wide agreement on the benchmarks without extensive discussion of an underlying ethical framework. Participants were introduced to the equal-opportunity theory but it played no explicit role in producing agreement on benchmarks. Because of our focus on fairness, we also avoided some culturally sensitive issues, such as abortion, euthanasia, and the use of human and fetal tissues or organs. We did discuss the fact that the weight or priority given to different benchmarks might vary in different countries depending on some cultural beliefs. In our workshops, these variations were not significant. We deliberately refrained from giving benchmarks an equal weighting in all countries.

[1] Professor, Department of Philosophy, Tufts University, Medford, MA 02468, USA (email: ndaniels@emerald.tufts.edu). Correspondence should be addressed to this author.

[2] President, Council for International Organizations of Medical Sciences (CIOMS), Geneva, Switzerland, and Emeritus Professor of Community Health Sciences, Aga Khan University, Karachi, Pakistan.

[3] Ministry of Health, Bogata, Colombia (participating as an individual and not as a formal representative of the Ministry of Health).

[4] Director of Health Policy and Planning, Center for Health Systems Research, National Institute of Public Health, Cuernavaca, Mexico.

[5] Associate Professor, Community Health Sciences, Aga Khan University, Karachi, Pakistan.

[6] Head, Centre of Health Equity Monitoring, Faculty of Medicine, Naresuan University, Phitanulok, Thailand.

Ref. No. **00-0583**

  

The benchmarks are relevant, because there is rapid reform of health care systems around the world as a result of changes in economic and political systems, economic growth, or previous failures to meet population needs. External agencies have played a large role in offering incentives to privatizing and decentralizing reforms. In all these contexts, however, reforms are usually debated without a systematic evaluation of their impact on the fairness of the resulting system. Privatizing and decentralizing efforts may aim at adding new resources and circumventing inefficient bureaucracies. The private sector, however, often competes with and weakens the public sector, and it requires strong and efficient regulation if it is not to undermine equity. Promoting some kinds of efficiency without attention to other dimensions of fairness will not improve fairness and may undercut it. The benchmarks provide a framework for evaluating the effects on fairness of these and other strategies.

The aim of the benchmarks is to encourage debate on the specific, interacting effects of the reforms being compared, not simply to produce a "report card" with numerical "grades." Consequently, for the necessary objectivity it is essential that a rationale, containing reasons and evidence, be provided for the score on each relevant criterion. Rationales might not be needed if we only included criteria with measurable magnitudes, such as the proportion of the population receiving some particular service or having some particular health status. Many critical components of fairness are not so directly measurable, and satisfaction of their criteria requires judgement.

We begin with a short history of the benchmarks approach, comment briefly on the benchmarks and their scoring, note some preliminary findings from their use, and conclude by explaining how the benchmarks supplement, rather than compete with, alternative ways of measuring equity and indexing health system performance.

## History of the benchmarks approach

The original "benchmarks of fairness" were developed to assess and promote discussion about comprehensive medical insurance reforms proposed in the United States in the first Clinton Administration (*1, 7, 8*). These benchmarks focused heavily on the needs in reforming a technologically advanced but inefficient and inequitable system that lacked universal coverage. Despite this specific focus, the original benchmarks addressed basic questions that must be asked about *any* reform:

- does it reduce barriers to access to public health measures and medical services?
- does it provide health care services appropriate to the needs of the population?
- does it distribute the burdens of paying for health protection fairly?
- does the reform promote clinical and administrative efficiency?
- does it make institutions publicly accountable for their decisions?
- how does it affect the choices people can exercise?

To adapt the benchmarks for use in health systems in countries at different levels of development (*9*), teams of collaborators from four countries — Colombia, Mexico, Pakistan, and Thailand — were formed. During 1999, these teams held two week-long workshops in Cuernavaca (for both the Colombian and Mexican teams), Bangkok and Karachi, with representation from each Asian site participating in the other Asian workshops. Members of the country teams had various backgrounds: university faculty members, representatives of donor agencies supporting health care reform, members of health services research teams working on reform options, and persons involved in policy-making at the national level.

Teams used each country as a "case study" for which appropriate benchmarks were developed. By successively reviewing the work of previous workshops across sites, the teams produced a modifiable scheme of benchmarks appropriate to all countries.

In each workshop, the process included:

- seminar presentations and discussion about the salient problems facing each system, including a history and critical evaluation of recent reform efforts;
- a seminar presentation and discussion about the original benchmarks and how they had been applied to reform efforts in the United States;
- a discussion of whether new benchmarks were needed to address local issues that were not addressed by the original set, or by the provisional set developed by previous workshops;
- a critical review and revision of each of the original benchmarks, or of the results provided by the preceding workshops;
- an attempt to link the detailed discussion of problems and reforms to specific criteria for each benchmark;
- "testing", including field testing in Thailand (*10*), of the provisional benchmarks by using them to score actual and proposed reforms in each country;
- refinement and revision of the criteria in light of these scoring attempts;
- development of specific plans for disseminating the benchmarks for actual use in each site.

The Asian workshops included field trips to villages and urban areas of high poverty to examine the delivery system and provide first-hand experience of the problems requiring reform.

## The revised benchmarks

There are nine benchmarks, each of which contains various criteria for evaluating specific aspects of the fairness of reform proposals. Here, we highlight key features of each benchmark.

Benchmark 1. **Intersectoral public health**

I. Degree to which reform increases the percentage of population, demographically differentiated where relevant and possible, receiving the following
- Basic nutrition
- Housing
  - Crowding
  - Homelessness
  - Physical adequacy
- Environmental factors
  - Clean water (and water treatment)
  - Sanitation (vector control)
  - Clean air
  - Reduced exposure to workplace and environmental toxins
- Education and health education
  - Literacy
  - Basic education
  - Health literacy
    - Nutritional education
    - Sex education and promotion
    - Substance abuse education
    - Anti-smoking education
    - Anti-drug and alcohol abuse education
- Public safety and violence reduction
  - Vehicular accident reduction
  - Violence reduction (homicide, rape)
  - Domestic abuse (women, children)

II. Development of information infrastructure for monitoring health status inequalities
- Provision for regular measurement of health status inequalities, using appropriate indicators
- Research into interventions most likely to reduce health status inequalities

III. Degree to which reform has actively engaged intersectoral efforts at local, regional, and/or national level to improve social determinants of health, and the degree to which vulnerable groups have been involved in defining these efforts.

## Benchmark 1: intersectoral public health

The rationale for this benchmark is that social determinants (*6, 11*) and other risk factors "upstream" from the point of health care delivery affect population health. The first criterion in Benchmark 1 asks for estimates of the degree to which a population benefits from reductions in exposure to various risk factors as a result of the reforms under consideration. Though not all reforms will affect all these factors, the comprehensive list is included because reforms would make a system more fair the more inequalities in exposure they eliminated. The criterion encourages gathering information on these exposures.

The second criterion calls for developing an information infrastructure needed to measure and monitor health inequalities and to carry out research about the most effective ways to reduce these. The third criterion evaluates reforms for their coverage across sectors and their involvement of communities and vulnerable groups in these efforts. Country-specific differences require country-specific adjustments of the benchmarks. For example, in some countries it may be crucial to focus on the reduction of violence or accidents, and in others on clean water or other factors.

Benchmark 2. **Financial barriers to equitable access**

I. Informal sector coverage
- Universal access to the most appropriate package of basic services, and improvement of packages over time
- Examples of packages of varying scope
  - 12 Mexican interventions (a minimal package)
  - Primary care package of the Pan American Health Organization (a slightly more extensive package)
  - Colombia's basic benefit package/subsidized regimen or Thai package
  - Catastrophic coverage (unclear just where Pakistan package fits, but probably below Colombia, through public facilities).
  - Drug coverage
  - Medical transportation costs
- Portability of coverage (geographical, employment status)

II. Insurance for formal sector
encourages moving populations from informal to formal sector
- Reduction of the following obstacles to enrolling people in the formal sector:
  - corruption and enforcement of tax requirements, mandatory enrolment
  - worker resistance to enrolment
  - small employer resistance
- Family coverage for enrolled workers
- Drug coverage
- Medical transportation costs
- Producing uniform benefits across all groups of workers
- Integrating various schemes involving those workers

## Benchmark 2: financial barriers to equitable access

Fairness requires reducing financial and nonfinancial barriers to access to needed services. Benchmark 2 recognizes the large "informal," nontaxable employment sector in many developing countries, often including 60–90% of the population. Since workers and their families in the informal sector generally include the poorest part of the population services must be provided in full or in large part through general tax revenues. The larger the informal sector, the larger the need for public financing, but the smaller the tax base to meet it.

Benchmark 2 encourages a long-term strategy aimed at moving as much of the population as possible into the formal sector and then into insurance schemes that can be built on broadly based general tax revenues, social security payments or employer-based contributions.

Benchmark 2 also specifies interim goals in both sectors. Because public resources are scarce in the informal sector, a crucial issue is whether the most important services are available to all. Benchmark 2 encourages reforms to specify a basic package of services that all will receive by a specific target date, then to improve that package over time. For example, the 1995 Mexican reforms, funded by external loans, aim to provide universal access to a very modest package of services. By 1999, over 90% of the population had access to this, and when 100% is reached, the Mexican government is obliged to finance this universal but modest package itself. In Colombia, the 1993 reforms aimed at a more

comprehensive benefit package for the informal sector. The new constitution, however, created legal pressure deriving from a right to life, to expand those benefits. It has not been possible with existing resources in Colombia to deliver that package universally. Neither reform would meet fully the criteria specified in Benchmark 2.

In Thailand, the debate continues about whether to implement a defined minimum benefit package proposed in recent reforms, or whether to continue to rely on a type of public insurance (type B) that allows providers discretion to negotiate what kinds of services will be available to those without any insurance. As a result of scoring Thai proposals using Benchmark 2, a specific question emerged about the levels of unmet need in this population. Research on that issue should improve policy planning.

In Pakistan, the informal sector includes 90% of the population. In *theory*, all people have access to a robust set of services. In *reality*, many services, including the provision of essential drugs, are not available for various reasons (e.g., the existence of shadow providers or inadequate funding), driving people to seek care from private sources. In scoring reforms, attention is paid to the gap between intention and implementation.

Benchmark 2 concentrates on two aims of reform of the formal sector besides increasing the size of the sector: producing uniform and more adequate benefits across all groups of workers and integrating the various schemes that involve these workers. In Thailand, for example, the long-range reform plans call for considerable integration of formal sector insurance plans through district fundholding and regulative controls, and eventual expansion of coverage to all family members. In Pakistan, with only 10% of workers in the formal sector, the team focused on the need to develop a plan that would lead to a well-integrated formal sector scheme, not an unplanned mixture of private plans with little regulation or equity.

Benchmark 3. **Nonfinancial barriers to access**

I. Reduction in geographical maldistribution
   - Facilities and services
   - Personnel (mix and training)
   - Supplies
   - Drugs
   - Clinic hours (appropriate to village routines, work schedules)
   - Transportation for medical purposes

II. Gender
   - Status in family regarding decision-making
   - Mobility
   - Access to resources
   - Reproductive autonomy
   - Gender sensitive provision of services, involvement of community political groups to address gender barriers

III. Cultural
   - Language
   - Attitude and practices relevant to disease and health
   - Uninformed reliance on untrained traditional practitioners (some healers, midwives, dentists, pharmacists)
   - Perception of public sector quality

IV. Discrimination by race, religion, class, sexual orientation, disease, including stigmatization of groups receiving public care

Benchmark 4. **Comprehensiveness of benefits and tiering**

I. All effective and needed services deemed affordable, by all needed providers
   No categorical exclusions

II. Reform reduces tiering and achieves more uniform quality
   Integrates services to the poor and others

## Benchmark 3: nonfinancial barriers to access

The first criterion evaluates reforms according to the measures they take to address the poor distribution of drugs, supplies, facilities and personnel common in all four countries. Where the reform relies on local fund-holding and decentralization, the criteria also examine specific goals and accountability for these (see also Benchmark 8).

The second criterion addresses gender barriers, which are especially important barriers to primary care in Pakistan, for example in the squatter communities of Karachi, where studies of children at high risk of death from diarrhoeal disease and pneumonia suggest that lack of maternal autonomy is a key risk factor. The benchmarks emphasize involving community political groups as an essential way to reduce these barriers, since simply providing services will not overcome them.

## Benchmark 4: comprehensiveness of benefits and tiering

The underlying rationale is that all people, regardless of class or ethnicity or gender, have comparable health needs and there are similar social obligations to meet these. Inequalities in the coverage and quality of care ("tiering"), reduce the fairness of systems. Some kinds of tiering are worse than others. It is less serious if a small but wealthy group does better than others, provided the others do well (e.g., private-sector insurance in the United Kingdom) than if a poor group is worse off than the rest of society (e.g., failing to insure the working poor in the United States, or failing to deliver a minimal benefit package to the whole informal sector while the top 5% of the population has excellent private insurance, as in Colombia). Some tiering is also unavoidable in systems with severe resource constraints and a large informal sector.

All teams focused on extensive differential treatment of people by class within a system, not only between the public and private sectors but within the public sector. Residents of Sultanabad, a squatter community of Karachi, remarked that "the tradesman will do better than the labourer in a public

Benchmark 5. **Equitable financing**

- I. Is financing by ability to pay?
  - If tax based-scheme
    - How progressive (by population subgroup)?
    - How much reliance on cash payments (by subgroup)?
  - If premium-based
    - Is it community-rated (by subgroup)?
    - Reliance on cash payments (by subgroup)?
  - Out-of-pocket payments contribute to both
    - Main source of shifting burdens to the sick

hospital," suggesting a widespread perception of tiering in the system, where the poor commonly wait four to five hours to be seen in a hospital, then get five minutes with the doctor, while well-to-do patients can just walk into private sector services and be seen right away. Tiering exists in the benefit packages available to different subgroups in the formal sector in Colombia, Mexico, Pakistan and Thailand. In Thailand, civil service workers will have better access to haemodialysis than other formal sector workers. In Mexico and Pakistan some multinational employers provide better coverage than the social security schemes, and the military in Pakistan has the best coverage of any group.

## Benchmark 5: equitable financing

This rests on the fundamental idea that financing medical services, as opposed to access, should be according to ability to pay. Three main sources of funding are involved in most systems: tax-based revenues, insurance premiums and out-of-pocket payments. The benchmark distinguishes primarily between tax-based and premium-based parts of the system, noting that in both there are still out-of-pocket payments for care. Tax-based schemes are more equitable if their structure is more progressive. Premium-based schemes are more equitable if they are community-rated, rather than risk-rated. Risk-rating shifts the burden to those at higher risk of illness. The same inequity is involved in out-of-pocket contributions in both tax-based and premium-based systems. A good measure of equity in financing must combine all financing systems (*12*).

The substantial out-of-pocket costs for health care in all four collaborating sites was the main source of regressivity in financing and the main way of shifting burdens to the sick, rather than pooling them across the whole population. There are many pressures on systems to rely on and even increase cash payments for services.

## Benchmark 6: efficacy, efficiency, and quality of care

The rationale for this and the next benchmark is that, other things being equal, a system that gets more value for money in the use of its resources is fairer to those in need. Distributive justice and fairness are issues because resources are always limited. A key criterion in Benchmark 6 is primary health care for community-based delivery. Reforms aimed at improving primary care must assure appropriate training, incentives, resource allocation and community participation in decisions affecting delivery. Emphasis was placed on a population focus and on the need for the integration of different parts of the health system, such as referrals. Community participation ideally involves an interactive relationship that goes beyond mere "outreach."

Benchmark 6. **Efficacy, efficiency and quality of health care**

- I. Primary health care (PHC) focus
  - PHC training for community-based delivery
    - Population-based
    - Community participation
    - Integration with rest of system (referrals)
    - Intersectoral integration (social and environmental determinants)
  - Incentives
  - Appropriate allocation of resources to PHC
  - Interactive community participation, including vulnerable subgroups
  - Referral mechanisms
    - Primary health care gatekeepers
    - By-passing primary health care sites
    - Respect for autonomy
- II. Implementation of evidence-based practice
  - Health policies
  - Public health and clinical prevention
  - Therapeutic interventions
    - Incentives for clinical guidelines
    - Evidence-based evaluation of methods for managing utilization of services
  - Information infrastructure and database
    - Evidence-based research on clinical and public health measures
    - Health services research on patterns of care
    - Population health needs and utilization rates, including variation studies (with demographic differentiation)
- III. Measures to improve quality
  - Regular assessment of quality, including satisfaction, with surveys or community group involvement as appropriate
  - Accreditation of plans and hospitals
  - Professional training
    - Curriculum focused on fair design of system
    - Continuing education

The second main concern of Benchmark 6 is promoting evidence-based practice in all areas of services, including preventive, curative and management practices. To advance this, the criteria call for the development of an information infrastructure and database, as well as for health services research to support evidence-based practice. The third main criterion concerns measures to improve quality of services in the system, including professional training, continuing education, accreditation and community participation in evaluating the quality of care.

In all systems we examined, there are problems with referral mechanisms and with the role of primary care gatekeepers. Dissatisfaction with primary care services leads many people to turn to higher level

Benchmark 7. **Administrative efficiency**

I. Minimize administrative overheads
- Appropriate technology acquisition
  - Purchase
  - Maintenance
  - Training
- Reduce excessive marketing costs (hospitals or plans)
- Efficient use of personnel
  - Reduction of excess
  - Appointments and promotions based on competence
- Appropriate economies of scale
  - Adequate risk pools for insurers
- Reduction of duplicate structures, including integration of vertically organized programmes
- Minimize transaction costs
  - Enrolment and disenrolment costs
  - Transfers of personnel or patients
  - Minimize loss of needed personnel from system as a whole (brain drain)
- Oversupply of some services in some areas

II. Cost-reducing purchasing
- Reduce price variation
- Drug cost reduction through large scale purchasing
- Reliance on (quality) generics where possible

III. Minimize cost shifting
- Cost shifting from PC to tertiary
- Cost shifting to patients
- Cost shifting to public sector or insurance from other types
- Cost shifting between schemes

IV. Minimize abuse and fraud and inappropriate incentives
- Shadow providers, partial and total
- Practitioner auto-referral
- Drug sales at profit by rural doctor
- Billing practices
- Unqualified practitioners in rural areas (also a problem in urban areas in Karachi)
- Vehicles and other perks
- Inappropriate promotion of drugs and devices
- Appropriation of public resources for private practice

hospitals for primary care, leading to considerable inefficiency. Similarly, there is no control of efficacy or quality since people will often abandon the public sector primary care services for completely unregulated private sector services. Establishing good referral systems is a critical element in the efficiency of care, but the restrictions to such systems also reduce the kinds of choice or autonomy assessed by Benchmark 9. To justify restrictions on autonomy, there must be qualified practitioners doing the diagnosis and referral, clear, accessible routes to higher levels of care, and general knowledge of the importance of such a system.

## Benchmark 7: administrative efficiency

Benchmark 7 seeks efficiency in the management of the health care system. Addressing these problems, however, also requires greater accountability, including transparency; consequently, Benchmark 8 must be used together with Benchmark 7 if real improvement is to result.

The criteria included in Benchmark 7 were constructed out of consideration of many examples for all four collaborating sites of sources of administrative inefficiency. Key areas of common concern were various sources of administrative overheads (inappropriate technology acquisition, inefficient use of personnel, high transaction costs), costly forms of purchasing, cost shifting and many types of abuse and fraud (shadow providers, drug sales and auto-referrals, inappropriate promotion of drugs and devices).

Some general points emerged that cut across the local differences. In all countries, public sector practitioners receive very low pay, and this is a reason for many of the forms of abuse that create efficiency and accountability problems (noted in the next benchmark). The failure to have integrated financing schemes means that there are incentives to shift costs from one part of the system to others. In Thailand, where unions are weak, civil service work rules prevent efficiencies of manpower allocation. In Latin America, strong unions and their work rules create the same obstacle to reallocation of personnel.

In the public sector of all the systems, a common set of complaints is articulated: bureaucratic practices and corruption lead to great inefficiencies in the purchase of supplies and equipment, failures to enforce rules about personnel, favouritism and hiring on grounds other than competency, and other highly inefficient practices. In all these contexts, there is talk about "decentralization" as a solution, but decentralization only helps if there is careful planning and regulation to make sure decentralized units are aiming at similar goals.

## Benchmark 8: democratic accountability and empowerment

Benchmark 8 emerged in all four countries as critically important, since without these forms of

Benchmark 8. **Democratic accountability and empowerment**

I. Explicit, public, detailed procedures for evaluating services with full public reports
- Use reports
- Performance reports
- Compliance reports
- Use of adequately qualified consultants

II. Explicit deliberative procedures for resource allocation with transparency and rationales for decisions based on reasons all ''stakeholders'' can agree are relevant

III. Global budgeting

IV. Fair grievance procedures
- Legal procedures (malpractice)
- Non-legal dispute resolution procedures

V. Adequate privacy protection

VI. Measures for enforcement of compliance with rules and laws

VII. Strengthening civil society
- Enabling environment for advocacy groups
- Stimulating public debate, including participation of vulnerable groups

Benchmark 9. **Patient and provider autonomy**

I. Degree of consumer choice
   - Of primary care providers
   - Of specialized care providers
   - Of alternative providers
   - Of procedures

II. Degree of practitioner autonomy

accountability, reforms are unlikely to succeed in any area. The rationale for including accountability is that health systems are responsible for the improvement of population health in an equitable manner, and those affected by decisions and policies that influence well-being in such fundamental ways must have an understanding of and ultimate control over that system. Such control is not exercisable without accountability for reasonableness (*13, 14*) in decision-making about allocation and other matters. Such accountability includes transparency — global budgeting, fair appeals processes, adequate privacy protection, and measures to enforce compliance with rules and laws. None of the criteria is ultimately effective without a strengthening of civil society, so that people understand the problems and are empowered to seek improvements to the health sector.

One important criterion, originally proposed in the Latin American workshops, evaluates reform for its attempt to stimulate the growth of advocacy groups, clearly a matter crossing boundaries between sectors. This criterion is important because of the crucial role such groups play in countries with developed democratic traditions, of pushing public authorities to attack problems in both public and private sectors. In Pakistan and Thailand, this idea was expanded into the criterion "strengthening civil society" which now has two components: establishing an enabling environment for advocacy groups and stimulating public debate about health policy measures. Many aspects of this benchmark go beyond merely holding institutions in the health sector accountable to the public; they actually increase the power of the public to act to remedy problems.

In thinking about scoring reforms, Benchmarks 6–8 should play a key role in helping to think through the content of measures, such as the decentralization of public bureaucracies and the establishment of district or other level budgeting of various revenue flows. The benchmarks aim at avoiding blindness to specifics of reform proposals incurred by fashionable labels or ideas.

## Benchmark 9: patient and provider autonomy

This is the benchmark that most directly addresses a culturally variable issue. How important is autonomy or choice? In some market-based approaches, informed choice is necessary if quality is to be improved and true preferences met. But how much choice, and what kinds of choices? Similarly, provider autonomy is much sought by professionals, but that is often seen by planners as an obstacle to efficient use of services, since professionals and provider institutions are influenced by incentives to utilize what they can supply.

For these reasons, it is important to emphasize how Benchmark 9 may conflict with other benchmarks and that people in the same or different cultures may disagree about weightings. Consequently there may be no one fairest system, but many fair designs. Benchmarks allow for cultural and other variations, but encourage discussion about grounds for designs that value some benchmarks over others.

A clear example of the conflict between Benchmark 9 and other benchmarks involves referral systems and the restrictions on patients they involve. Benchmark 6, for example, may approve of restrictions on autonomy in order to achieve a primary care focus and the efficiency that results from letting primary care physicians filter access to other levels of care, but Benchmark 9 is concerned with loss of choice. Similarly, choice of alternative providers will undermine efficiency and quality if there is no adequate evidence-based assessment of credentials or alternative forms of treatment. Practitioner autonomy may be essential if the practitioner is to address the health care problems of individual patients, but this presupposes high levels of competency and knowledge of appropriate practices.

## Scoring and uses of the benchmarks

We have adapted the benchmarks for use in evaluating competing reform proposals within a country and the discussion focuses on that use. It may be possible to use the benchmarks to make some international comparisons of fairness across systems (*1, 14*), but we have ignored such an application in the work reported here. In evaluating reforms, progress is made if people can agree on what they think the current limitation of the system is and then agree about how much a specific reform would improve or worsen that aspect of the system. Disagreements about scoring will improve the discussion about merits of reforms, which is the ultimate goal of the benchmarks. It is crucial to understand this purpose of scoring in order to see why we have adopted a particular approach to it.

In the original use of the benchmarks to evaluate competing reforms in the United States, a scoring system was adopted that took the status quo as a "0," assigned a maximal positive outcome a "5," and maximal regression from the status quo a "–5." Since numbers invite confusion, our Latin American teams used symbols ("pluses" or "minuses") to show that scoring was primarily aimed at a clear presentation of the underlying principles. The Asian teams were comfortable with the convenience of numbers,

carefully explaining that they were used for ranking. All agreed on the primary point: the scoring exercise is aimed at generating clear fundamentals, and we agreed to leave the choice of symbols to country teams using the tool.

The point of this method of scoring is to see how well particular reform proposals fare on the many aspects of fairness covered by the benchmarks. Some proposals will be stronger on some dimensions than others. Where these do not represent true tradeoffs, it may be possible to formulate policies that are true improvements overall. Where tradeoffs are being made the framework stimulates discussion of the competing values underlying the alternatives.

## Some preliminary findings

Our scoring exercises in Colombia, Mexico, Pakistan, and Thailand showed that the adapted benchmarks could reveal:

- places where proposed reforms were insufficiently detailed or vague about mechanisms to reveal their effects;
- problematic assumptions about how goals of reform would be achieved;
- empirical issues that would have to be resolved in order to determine the likely success of implementing a reform.

As a result, we were able in both Pakistan and Thailand to construct practical lists of such issues to be brought before groups considering the proposals for implementation. We also noted the large gap that often loomed between the intention and the results of reforms. We were able to show this gap by scoring both the intention and implementation of a proposal, where we had evidence about implementation.

In Thailand, the benchmarks were "field tested" by asking people to use them for evaluation of national reforms under consideration (and partly implemented) as well as for changes at the provincial level over a two-year period. Results are reported fully elsewhere (*10*). This exercise has led to the proposal that the benchmarks be deployed for use in evaluating current national proposals for system reform and for use in evaluating plans made by provincial health officers, who will have more autonomy under proposed reforms. In Pakistan, plans exist to incorporate the benchmarks into training programmes for provincial and district health officers, as well as into medical school curricula. Plans also exist to have a public health network of academic centres promote the use of the benchmarks at national and provincial levels. There are more ambitious plans to involve regional WHO organizations in the broader adaptation and dissemination of the tool.

All participants at the four sites agreed that a useful format for presenting the final product will be an interactive computer program that allows policy analysts and broader community groups to draw on a database of similar reforms and their outcomes. Properly designed, such a tool would allow concentration on selected benchmarks with the option of ignoring those less relevant. This flexibility would allow it to be used at different levels within a system — not just for comprehensive national reforms, but for more specific reforms at the provincial or district level. The Pan American Health Organization has already expressed interest in posting such a tool on its web site as one item in a policy "toolbox". (We note that such a program would have a much more specific function than the program Policy Maker; Policy Maker provides analytical techniques for evaluating the ability to implement any kind of policy, but it lacks the detailed framework for assessing fairness that is included in the benchmarks.)

## The benchmarks versus other measures of equity and health system performance

In conclusion, we emphasize that the benchmarks supplement or complement, rather than compete with, various other efforts to monitor equity in health systems or to index health system performance across countries. Consider, for example, WHO-sponsored efforts to develop measures for monitoring health inequities across demographic groups and for setting goals and targets for reducing these inequities (*4*). Some new approaches to measuring health inequalities may better highlight subgroup differences (*16, 17*). Some of these measures could be incorporated into the benchmark approach; in addition, since setting targets requires evaluating how reforms will affect a system, the benchmarks will prove a useful supplement to such an approach.

It is commonly noted that strategies that aim to reduce the aggregate burden of disease in a population sometimes conflict with strategies aimed at reducing inequalities in health status. On the assumption that reasonable people may disagree about how to resolve such conflicts, the benchmarks refrain from making an overall judgement on this issue and instead insist on fair procedures for making them within a country. Specifically, Benchmark 8 requires that reforms put into place procedures for making resource allocation decisions in a way that is transparent and publicly accountable.

As noted earlier, the benchmarks attempt no uniform scaling of fairness across systems. Instead, we adopt the status quo as a baseline for purposes of evaluating intracountry reforms. Suppose, however, that WHO develops an index that measures health system performance across countries that includes health status, responsiveness, and fairness in financing (*18*). Such an index would help focus attention on areas where reform was clearly needed. At the same time, such an index would only supplement, not replace, the use of the benchmarks in evaluating intracountry reforms, for any country stimulated by

the cross-country index to undertake further reforms would still benefit from the multidimensional tool for evaluating reforms that we propose.

To date, the benchmarks have been used in a preliminary way to evaluate reform proposals or recent reforms. Pannarunothai & Srithamrongsawat (*10*) report on field tests in which the benchmarks were used to evaluate proposed national reforms and recent provincial reforms that establish the benchmarks can be used for scoring reforms in Thailand. Teams in Colombia, Mexico, and Pakistan have also carried out scoring exercises to test the usefulness of the specific criteria. In each setting, the benchmarks have shown that they can stimulate thinking about the mechanisms of reforms, force greater specification of reform measures and help to pose research questions that can bring evidence to bear on choices. A fuller evaluation of the approach must await a wider testing of the benchmarks. ■

## Acknowledgements

This project was funded by The Rockefeller Foundation and its Global Health Equity Initiative. We thank the National Institute of Public Health in Mexico, the National Institute of Health in Thailand, and the Community Health Sciences Department of Aga Khan University, Karachi, Pakistan, for hosting our workshops. We also wish to thank the members of the collaborating country teams and our hosts at the various site visits we made. N. Daniels was also supported by a Robert Wood Johnson Investigator Award.

## Résumé

### Points de référence de l'équité pour la réforme des soins de santé : un outil pour l'élaboration des politiques dans les pays en développement

Des équipes de collaborateurs de la Colombie, du Mexique, du Pakistan et de la Thaïlande ont adapté un outil initialement mis au point pour évaluer les réformes de l'assurance-maladie aux Etats-Unis d'Amérique afin d'en tirer des « points de référence de l'équité » permettant d'évaluer les réformes des systèmes de santé dans les pays en développement.

L'équité est une notion présentant de nombreux aspects et qui envisage notamment les résultats sanitaires, l'exposition aux facteurs de risque, l'accès à toutes les formes de soins et le financement. Elle s'intéresse aussi à l'efficacité de la gestion et de la répartition des ressources, de façon à pouvoir répondre aux besoins lorsque celles-ci sont limitées, en garantissant la transparence et l'autonomie du malade et du dispensateur de soins. Pour la mise au point initiale des points de référence aux Etats-Unis, la justification éthique se fondait sur une théorie de la justice et des soins sociaux, l'idée de départ étant que la maladie et l'incapacité réduisent l'éventail des chances individuelles et qu'un principe régissant l'égalité des chances offre une base pour réglementer un système de soins de santé. Cette justification fondamentale n'a joué aucun rôle dans notre processus d'adaptation, qui a visé à dégager un consensus sur les composantes, les outils et les procédés assurant l'équité d'un système. Une caractéristique frappante de nos résultats est la convergence de ces points de référence malgré les importantes différences historiques, culturelles et politiques entre les sites concernés.

On distingue neuf points de référence, dont chacun contient différents critères pour évaluer les aspects spécifiques de l'équité des propositions de réforme. Le point 1 envisage les questions de l'ensemble des secteurs de la santé publique qui affectent la santé avant la fourniture de soins médicaux. Les points 2 à 4 visent différents aspects de l'équité d'accès aux services médicaux, notamment les obstacles financiers et non financiers et les niveaux de prestations, ou les prestations inégales dont peuvent se prévaloir différents sous-groupes. Le point 5 concerne l'équité en matière de financement, en allant au-delà de la question des obstacles financiers visée par le point 2. Les points 6 et 7 envisagent comment optimiser les ressources dans les services cliniques et l'administration du système. Le point 8 couvre la question critique de la responsabilité des décideurs, des administrateurs et des dispensateurs dans le cadre du système, ainsi que les moyens d'action conférés aux communautés et le renforcement de la société civile. Enfin, le point 9 concerne la question du choix ou de l'autonomie des malades et des dispensateurs de soins.

Les points de référence relient les aspects généraux de l'équité à différents critères spécifiques couvrant les composantes, les mécanismes et les procédés contribuant à l'équité. Les réformes sont évaluées par un score indiquant le degré d'amélioration apporté à la situation, par exemple sur une échelle de – 5 à +5 où zéro représente le statu quo. Une justification de départ est fournie pour que le score et l'évaluation aient une base objective.

L'objet de l'approche des points de référence consiste à stimuler la réflexion sur l'équité en dépassant les divisions entre les disciplines qui empêchent les analystes et le grand public de comprendre comment les arbitrages entre ces éléments de la réforme du système affectent l'équité de la réforme dans son ensemble. Les scores obtenus dans les quatre sites montrent que les points de référence peuvent révéler le caractère trop vague et inadéquat des spécifications de détail concernant les propositions de réforme ou des problèmes au niveau des hypothèses de départ et de points empiriques, auxquels il faudra s'efforcer de remédier. Les points de référence peuvent être utilisés au niveau tant national que local, et nous décrivons des plans d'utilisation dans les sites participant à la collaboration. L'outil vient moins concurrencer que compléter les autres approches concernant la mesure des inégalités sanitaires et les résultats du système de santé.

## Resumen

### Criterios de equidad de la reforma de la atención sanitaria: un instrumento para el análisis de políticas en los países en desarrollo

Equipos de colaboradores procedentes de Colombia, México, el Pakistán y Tailandia han adaptado un instrumento originariamente elaborado para evaluar la reforma del seguro médico en los Estados Unidos al objeto de extraer de él «criterios de equidad» para evaluar la reforma del sistema de salud en países en desarrollo.

En este contexto la equidad es un concepto complejo que abarca la equidad respecto de los resultados sanitarios, de la exposición a factores de riesgo, del acceso a toda forma de atención y de la financiación. Abarca asimismo la eficacia de la administración y de la asignación de recursos, que posibilita la satisfacción de las necesidades con unos recursos limitados, una gestión responsable y la autonomía del paciente y del dispensador del servicio. Estos criterios se elaboraron originariamente en los Estados Unidos, donde la justificación ética de este concepto de la equidad se basaba en una teoría de la justicia y la asistencia sanitaria. Según esta teoría la enfermedad y la discapacidad reducen las oportunidades de los individuos, mientras que el principio que rige la igualdad de oportunidades ofrece una base para regular el sistema de atención sanitaria. Pero en nuestro proceso de adaptación no recurrimos a este fundamento teórico, sino que nos concentramos en llegar a un consenso acerca de cuáles son los componentes, mecanismos y procesos de un sistema que hacen que éste sea equitativo. Una característica notable de nuestros resultados es la coincidencia acerca de estos criterios, pese a las grandes diferencias históricas, culturales y políticas existentes entre los sitios colaboradores.

Hay nueve criterios, cada uno de los cuales abarca varios subcriterios que permiten evaluar aspectos específicos de la equidad de las propuestas de reforma. El criterio 1 se refiere a cuestiones relacionadas con diversos ámbitos de la salud pública que afectan a la salud antes de la prestación de servicios médicos. Los criterios 2 a 4 se refieren a diversos aspectos de la equidad de acceso a los servicios médicos, incluidos obstáculos financieros y no financieros y el grado de diferenciación, esto es, de desigualdad de los beneficios a que tienen acceso diferentes subpoblaciones. El criterio 5 se refiere a la equidad de la financiación más allá de los obstáculos financieros considerados al aplicar el criterio 2. Los criterios 6 y 7 se refieren al buen aprovechamiento de los servicios clínicos y a la administración del sistema. El criterio 8 abarca el problema crítico de la responsabilización de los decisores, administradores y proveedores de servicios del sistema, así como el empoderamiento de las comunidades y el fortalecimiento de la sociedad civil. Por último, el criterio 9 se refiere a la posibilidad de elección o la autonomía de los pacientes y los prestadores de los servicios.

Los criterios enlazan aspectos generales de la equidad con varios subcriterios específicos que abarcan componentes, mecanismos y procesos que contribuyen a la equidad. Las reformas se evalúan asignando un puntaje según el grado en que mejoran la situación, por ejemplo en una escala de – 5 a 5 en la cual el cero representa el *statu quo*. Se expone la base objetiva racional de los puntajes y de la evaluación.

El objetivo de estos criterios es fomentar la reflexión sobre la equidad trascendiendo todas las divisiones entre disciplinas que impiden a los analistas políticos y al público entender la manera en que las ventajas y desventajas de estas características del sistema de reforma afectan a la equidad general de la reforma. Los ejercicios de asignación de puntajes en los cuatro sitios mostraron que los criterios podían revelar las vaguedades y la falta de precisión de las propuestas de reforma, suposiciones cuestionables y problemas empíricos que era necesario abordar. Los criterios se pueden utilizar a nivel tanto nacional como local y describimos planes para su utilización en los sitios colaboradores. Este instrumento complementa otros métodos utilizados para determinar cuantitativamente la inequidad en materia de salud y el desempeño de los sistemas de salud, pero no compite con aquéllos.


## References


1. **Daniels N, Light D, Caplan R.** *Benchmarks of fairness for health care reform.* New York, Oxford University Press, 1996.
2. **Whitehead M.** The concepts and principles of equity and health. *International Journal of Health Services,* 1992, **22**: 429–445.
3. **Dahlgren G, Whitehead M.** *Policies and strategies to promote social equity in health.* Stockholm, Institute of Future Studies, 1992.
4. **Braveman P.** *Monitoring equity in health: a policy-oriented approach in low- and middle-income countries.* Geneva, World Health Organization, 1998 (unpublished document WHO/CHS/HSS/98.1).
5. **Daniels N.** *Just health care.* New York, Cambridge University Press, 1985.
6. **Daniels N, Kennedy B, Kawachi I.** Why justice is good for our health: the social determinants of health inequalities. *Daedalus,* 1999, **128**: 215–252.
7. **Brock D, Daniels N.** Ethical foundations of the Clinton administration's proposed health care system. *Journal of the American Medical Association,* 1994, **271**: 1189–1196.
8. **Daniels N.** *Seeking fair treatment: from the AIDS epidemic to national health care reform.* New York, Oxford University Press, 1995.
9. **Daniels N, Bryant J.** Parametros de justicia y monitoreo de la equidad: apoyo a un programa de la OMS *Salud y Gerencia*, 1998, **16 (**1): 7–12 (in Spanish).
10. **Pannarunothai S, Srithamrongsawat S.** Benchmarks of fairness for health system reform: the tool for national and provincial health development in Thailand. *Human Resources for Health Development Journal,* 2000, **4** (1) (in press).
11. **Sen A.** Health in development. *Bulletin of the World Health Organization*, 1999, **77**: 619–623.

12. **van Doorslaer E, Wagstaff A, Rutten F.** *Equity in the finance and delivery of health care. An international perspective*. Oxford, Oxford University Press, 1993.
13. **Daniels N, Sabin J.** Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philosophy and Public Affairs*, 1997, **26**: 303–350.
14. **Daniels N, Sabin J.** The ethics of accountability in managed care reform. *Health Affairs*, 1998, **17**: 50–64.
15. **Caplan RL, Light DW, Daniels N.** Benchmarks of fairness: a moral framework for assessing equity. *International Journal of Health Services*, 1999, **29**: 853–869.
16. **Gakidou EE, Murray CJL, Frenk J.** Defining and measuring health inequality. *Bulletin of the World Health Organization*, 2000, **78**: 42–54.
17. **Gwatkin DR.** Health inequalities and the health of the poor. *Bulletin of the World Health Organization*, 2000, **78**: 3–18.
18. *The world health report 2000 – Health systems: improving performance*. Geneva, World Heallth Organization, 2000.