Alan Lopez
School of Population Health, University of Queensland, Brisbane, QLD 4006, Australia
a.lopez@sph.uq.edu.au

I was a staff member of WHO from 1980 to 2002.

1 WHO. Financial report and audited financial statements for the financial period 1 January 2006–31 December 2007. Annex: Extra-budgetary resources for programme activities. April 24, 2008. http://www.who.int/gb/ebwha/pdf_files/A61/A61_20Add1-en.pdf (accessed Oct 6, 2008).
2 Stuckler D, King L, Robinson H, McKee M. WHO's budgetary allocations and burden of disease: a comparative analysis. *Lancet* 2008; **372:** 1563–69.
3 Population Division of the Department of Economic and Social Affairs of the UN Secretariat. World population prospects: the 2006 revision. March 13, 2007. http://www.un.org/esa/population/publications/wpp2006/wpp2006.htm (accessed Oct 13, 2008).
4 Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2007; **367:** 1747–57.
5 Snow RW, Craig M, Deichmann U, Marsh KW. Estimating morbidity, mortality and disability due to malaria among Africa's non-pregnant population. *Bull World Health Organ* 1999; **77:** 624–40.
6 WHO. World report on violence and health. Oct 3, 2002. http://www.who.int/violence_injury_prevention/violence/world_report/en/index.html (accessed Oct 13, 2008).
7 Norton R, Hyder AA, Bishai D, Peden M. Unintentional injuries. In: Jamison DT, Breman JG, Measham AR, et al, eds. Disease control priorities in developing countries, 2nd edn. New York: Oxford University Press and the World Bank, 2006: 737–53.
8 Peto R, Lopez AD. Future worldwide health effects of current smoking patterns. In: Koop CE, Pearson CE, Schwarz CE, eds. Critical issues in global health. San Francisco: Jossey-Bass, 2001: 154–61.
9 Ezzati M, vander Hoorn S, Lawes CMM, et al. Rethinking the "diseases of affluence" paradigm: global patterns of nutritional risks in relation to economic development. *PLoS Med* 2005; **2:** e133.
10 Law MR, Frost CD, Wald NJ. By how much does dietary salt reduction lower blood pressure? *BMJ* 1991; **302:** 811–24.
11 Pignone M, Phillips C, Mulrow C. Use of lipid lowering drugs for primary prevention of coronary heart disease: meta-analysis of randomised trials. *BMJ* 2000; **321:** 983–86.
12 James WP, Leach R, Mhurchu CN, et al. Overweight and obesity (high body mass index). In: Ezzati M, Lopez AD, Rodgers A, Murray CJL, eds. Comparative quantification of health risks: vol 1. Geneva: World Health Organization, 2004: 497–596.

What do we mean by rigorous health-systems research?

Health-systems research is recognised worldwide as vital to improving health-system performance.[1] However, such research is commonly perceived to lack rigour. We address the question of how such rigour should be judged, and argue that standards from medical research are largely inappropriate for health-systems research, though we accept that the quality of such research can and must be improved.

Health-systems research is concerned with how health services are financed, delivered, and organised, and how these functions are linked within an overall health system with its associated policies and institutions. Within this broad canvas, we focus on a specific area of health-systems research: the evaluation of changes to components of the health system (such as financing or organisational reforms), referred to as health-system interventions.

The features of such complex interventions pose evaluation challenges that are different from those in medical research. They include the difficulties of precisely defining the intervention and multifactorial causality (panel). Health-system decision makers have to be concerned not only about whether health-system interventions generate better performance, but also about when, why, how, and in what circumstances such interventions work well.[3,4] Most importantly, the impact of a system change on health outcomes is not the only question that matters because good performance is multidimensional and decision makers' interests require evaluation of processes as well as effects.[5] These are then much broader questions than those which usually concern medical research, in which the intervention can be more clearly specified, the desired outcome is largely unidimensional, and the context and processes are less important.

The range of relevant research questions inevitably requires a range of research methods drawn from different disciplinary traditions.[6] The approaches relevant to assessing rigour are thus necessarily more varied than those commonly applied in medical research. Such approaches begin with concern for whether the study design and methods are appropriate to the research question. Appropriate methods must then be applied correctly in terms of the standards of the discipline or tradition from which they are drawn; or, when new methodological approaches are developed, according to emerging consensus. Finally, reporting must always make clear what methods were used and why, how personal and political biases may affect findings, and whether, and in what way, the findings may be relevant in other contexts.

Discussion of two specific issues allows more consideration of what rigorous health-systems research entails and how this may depart from

***Panel:* Complex social programmes are:**[2]

- Based on set of theories and assumptions about how an intervention will lead to change
- Achieved through active participation of individuals
- Developed and implemented through long process which may be fallible
- Not necessarily implemented in linear fashion and influenced by respective power of those actors involved in implementation
- Very susceptible to effect of different contexts (eg, policy timing, organisational culture and leadership, resource allocation, staffing levels and capabilities, interpersonal relationships, and competing local priorities and influences)
- Prone to being changed during process of implementation
- Open dynamic systems in themselves, which are able to change the conditions that enable them to be implemented successfully (generating unintended positive and negative effects)

standard criteria for medical research. Two points are important. First, randomisation and control groups. Evaluation of the effect of an intervention requires designation of the counterfactual—ie, what would have happened in the absence of the intervention. For medical interventions, random allocation of individuals (or clusters) to control or intervention groups is the gold standard against which other designs are judged, and alternatives must be defended against the charge that some other factor(s) might be responsible for the measured effect. There has been mixed success in introducing randomisation into the evaluation of health-system interventions. Practical considerations, such as the requirement for a change in legislation or national policy, or political resistance to gradual scale-up, may mean that programmes have to be implemented everywhere at once. There are also concerns about the external validity of randomised designs.[7] Various alternatives have been suggested, including quasiexperimental methods and ex-post adjustments, such as propensity score matching.[8] Plausibility designs, which combine the opportunistic selection of control groups with a systematic effort to rule out confounding factors,[7] are an important way to ensure rigour in health-systems research. Systematic application of a theoretical model of how the intervention is supposed to work, together with collection of data on contextual factors, assumptions, and intermediate pathways, can provide additional assurance that measured effects are likely to be due to the intervention and help compensate for lack of a true control.

Second, understanding processes and contexts. Health-system change is always dynamic. Understanding what influences play out in particular experiences is essential to determine: whether the promise of performance gains in well-resourced pilot studies can be achieved more widely; the replicability of experience across different contexts; the management strategies that can support effective implementation; and why change generates unexpected and unwanted effects. These interconnected sets of issues require study designs and methods that generate indepth understanding of complex realities and processes. Case-study designs, involving the collection of qualitative and quantitative data, are particularly appropriate to such investigation and a well-developed body of work offers guidance.[9] Recent reflection on improving case-study rigour in health-systems research emphasises the need to: select cases carefully so that they are appropriate and offer valuable insights; gather rich and deep information on the contextual features relevant to each case; adopt an explanatory rather than a descriptive focus; consider multiple cases and conduct cross-case comparisons; and use relevant theory to guide inquiry.[10] Appropriate application of generally accepted criteria for qualitative and quantitative data collection and analysis is also clearly necessary.

Examples from the health-financing literature illustrate these points. When user fees were promoted in the 1990s, two studies were widely cited as showing their potentially beneficial effects.[11,12] The quasiexperimental design of the studies was taken to suggest that their findings would be robust. Yet a recent systematic review of the evidence on user fees[13] has highlighted the presence of confounding factors as well as some inappropriate econometric specifications in these studies. It has become clear that a series of questions relating to health-system characteristics should have been explored in greater detail, to assess applicability to other contexts. Institutional features, national power dynamics, and resource constraints are particularly important when scaling up pilot programmes. For user fees, the functioning of exemption schemes, and local financial and drug-procurement systems, have proved

major bottlenecks for successful policy reform. We hope to stimulate debate on rigour in health-systems research. By identifying some of the issues, we aim to encourage better understanding of this field of work among the broader health-research community, as well as better research.

Anne Mills, Lucy Gilson, Kara Hanson, Natasha Palmer, Mylene Lagarde
Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
anne.mills@lshtm.ac.uk

We declare that we have no conflict of interest.

1 Alliance for Health Policy and Systems Research. Strengthening health systems: the role and promise of policy and systems research. 2004. http://www.who.int/alliance-hpsr/resources/Strengthening_complet.pdf (accessed Oct 17, 2008).
2 Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review—a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy* 2005; **10** (suppl 1)**:** 21–34.
3 Walt G, Gilson L. Reforming the health sector: the central role of policy analysis. *Health Policy Plan* 1994; **9:** 353–70.
4 Kelly MP, Morban A, Bonnefoy J, Butt J, Bergman V. The social determinants of health: developing an evidence base for political action. Final report to the World Health Organisation Commission on the Social Determinants of Health from the Measurement and Evidence Knowledge Network. 2007. http://www.who.int/social_determinants/resources/csdh_media/hskn_final_2007_en.pdf (accessed Oct 17, 2008).
5 Lavis JN, Becerra Pasada F, Haines A, Osei E. Use of research to inform public policymaking. *Lancet* 2004; **364:** 1615–21.
6 Fulop N, Allen P, Clarke A, Black N, eds. Studying the organisation and delivery of health services: research methods. London: Routledge, 2001.
7 Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidemiol* 1999; **28:** 10–18.
8 Ravallion M. The mystery of the vanishing benefits: an introduction to impact evaluation. *World Bank Econ Rev* 2001; **15:** 115–40.
9 George AL, Bennett A. Case studies and theory development in the social sciences. Cambridge, MA: MIT Press, 2005.
10 Walt G, Shiffman J, Schneider H, Murray SE, Brugha R, Gilson L. "Doing" health policy analysis: methodological and conceptual reflections and challenges. *Health Policy Plan* 2008; **23:** 308–17.
11 Diop F, Yazbeck A, Bitran R. The impact of alternative cost recovery schemes on access and equity in Niger. *Health Policy Plan* 1995; **10:** 223–40.
12 Litvack J, Bodart C. User fees plus quality equals improved access to health care: results of a field experiment in Cameroon. *Soc Sci Med* 1993; **37:** 369–83.
13 Lagarde M, Palmer N. The impact of user fees on health services utilization in low- and middle-income countries—how strong is the evidence? *Bull World Health Organ* 2008; published online Oct 31. DOI:10.2471/BLT.07.049197.

Scaling up research and learning for health systems: time to act

See **Editorial** page 1519
See **Health Policy** page 1571

Recent years have seen a growing recognition of the importance of health systems in achieving the health-related Millennium Development Goals (MDG) by 2015, and the constraints related to systems short-falls which have hindered progress. These complex constraints are more easily understood through single-dimension diagnoses, such as a global shortage of over 4 million health workers,[1] or out-of-pocket payments for health care that push 100 million persons annually into poverty.[2]

These systems shortfalls are now being recognised at the highest level in global policy circles, as witnessed by their inclusion in the G8 follow-up Framework for Action, the Accra Agenda for Action,[3] and the MDG High Level Taskforce follow-up.[4] The Global Fund for AIDS, Tuberculosis and Malaria, the GAVI Alliance, and other major global-health initiatives are now encouraging countries to include activities for health-systems strengthening in their applications for support. These efforts coincide with other efforts specifically aimed at encouraging greater private-sector and civil-society involvement, strengthening information systems and health workforces at the country level, and aligning global partners around countries' health plans.[5–8]

This intense activity raises the question of where we stand in systems learning and draws attention to the consensus statement of an earlier Ministerial Summit on Health Research in Mexico in 2004.[9] In today's *Lancet*, Sara Bennett and colleagues[10] report a stocktaking exercise. This analysis notes increases in funding, but points to underinvestment in researchers in low-income countries, fragmentation of global efforts, and a research agenda that lacks focus. These concerns prompted WHO's Director-General to convene a high-level consultation in June, 2008, to explore how research and learning on health systems could be scaled up.[11] In follow-up, a WHO-convened Task Force is preparing a set of recommendations for consideration at the Bamako Ministerial Forum on research for health in November, 2008.

These recommendations aim to raise the global profile of health-systems research but also emphasise several priority areas. One of these relates to better understanding of how health systems and programmes funded by global-health initiatives interact to influence demand, service delivery, health-system functioning, and outcomes. Indeed, a recently launched initiative, "Maximizing Positive Synergies between Health Systems