Sources of Variability in Costing Methods

Implications for Transferability of Cost-Effectiveness Results

Taghreed Adam

Printing of this thesis was realized with financial support of:

The Department of Health Policy and Management, Institute for Medical Technology Assessment, Erasmus MC, University Medical Centre, Rotterdam

Sources of variability in costing methods. Implications for the transferability of cost-effectiveness results.
Taghreed Adam
Thesis Erasmus MC, University Medical Centre Rotterdam,
with summary in English and Dutch.

Front cover illustration: photograph by Pierre St-Jacques, published with permission from the Canadian International Development Agency.
Back cover illustration: photograph by Victoria Darrah, published with the artist's permission.
Cover design: Taghreed Adam
Layout: Taghreed Adam
Printed by: PrintPartners Ipskamp, Enschede

Sources of Variability in Costing Methods

Implications for the transferability of
cost-effectiveness results

Bronnen van variabiliteit in berekeningsmethoden van kosten

Gevolgen voor het overzetten van resultaten van
kosten-effectiviteitsanalyses
van het ene land naar het andere land

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
Rector Magnificus

Prof.dr. S.W.J. Lamberts

and according to the decision of the Doctorate Board.
The public defence shall be held on
Friday 13 January 2006 at 13:30 hrs

by

Taghreed Mohamed Salama Adam
born at Guiza, Egypt.

Doctoral Committee

Promotor:	Prof.dr. F.F.H. Rutten
Other members:	Prof.dr. E.K.A. van Doorslaer
	Prof.dr. J.D.F. Habbema
	Prof.dr. J. Fox-Rushby
Copromotor:	Dr. M.A. Koopmanschap

"In theory, there is no difference between theory and practice. But, in practice, there is."

Jan L.A. van de Snepscheut

To my father, my mother,
my sister and her family

Publications

The chapters in this thesis are based on the following articles[i]:

Chapter 2. Adam T, Evans DB and Koopmanschap MA. Cost-effectiveness Analysis: Can We Reduce Variability in Costing Methods? *International Journal of Technology Assessment in Health Care, 2003 19:2.*

Chapter 3. Adam T, Manzi F, Armstrong Schellenberg J, Mgalula L, de Savigny D and Evans DB. Does the Integrated Management of Childhood Illness Cost more than Routine Care? Results from Tanzania. *WHO Bulletin, May 2005, 83(5) 369-77.*

Chapter 4. Armstrong Schellenberg JRM, Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O, John T, Charles S, Nathan R, Wilczynska K, Mgalula L, Mbuya C, Mswia R, Manzi F, de Savigny D, Schellenberg D and Victora C. Effectiveness and Costs of Facility-based IMCI in Tanzania. *Lancet, 2004 (9445);364:1583-94.*

Chapter 5. Adam T, Evans DB and Murray CJL. Econometric Estimation of Country-Specific Hospital Costs. *Cost Effectiveness and Resource Allocation Journal (CERA) 2003 1:3.*

Chapter 6. Adam T and Evans DB. Determinants of Variation in the Cost of Inpatient Stays versus Outpatient Visits in Hospitals. A Multi-Country Analysis. *Submitted.*

---

[i] Reprinted with permission of Cambridge University Press, WHO Bulletin, Oxford University Press, BMJ and the Lancet.

Chapter 7. Adam T, Evans DB, Ying B and Murray CJL. Understanding Variations in Department-Specific Hospital Costs: An Application of Seemingly Unrelated Regressions to China. *Submitted.*

Chapter 8. Adam T, Amorim DG, Edwards SJ, Amaral JA and Evans DB. Capacity Constraints to the Adoption of New Interventions: Consultation Time and the Integrated Management of Childhood Illness in Brazil. *Health Policy and Planning 2005; 20 (Suppl. 1).*

Chapter 9. Gutierrez JP, Johns B, Adam T, Bertozzi S , Tan-Torres Edejer T, Greener R, Hankins C and Evans DB. Achieving the WHO/UNAIDS Antiretroviral Treatment "3 by 5" Goal: What Will it Cost? *Lancet, Vol 364, July 3, 2004.*

Chapter 10. Evans DB, Tan-Torres Edejer T, Adam T, Lim SS, and the WHO-CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Methods to Assess the Costs and Effects of Interventions for Improving Health in Developing Countries. *BMJ 2005 (In press)*

Chapter 11. Adam T, Lim SS, Mehta SS, Bhutta ZA, Fogstad H, Mathai M, Zupan J and Darmstadt GL. Achieving the Mellinnium Development Goals for Health: Cost-effectiveness of Strategies for Maternal and Neonatal Health in Developing Countries. *BMJ 2005 (In press)*

Chapter 12. Evans DB, Lim SS, Adam T, Tan-Torres Edejer T and the WHO-CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Evaluation of Current Strategies and Future Priorities for Improving Health in Developing Countries. *BMJ 2005 (In press)*

Table of Contents

1

Introduction

Chapter 1. Introduction

1. 1 Background

Information on the costs of health interventions is valuable to health decision-makers for at least two reasons. The first is for budgeting purposes, to identify the resources necessary to undertake, sustain or scale up an intervention[i]. The second is for efficiency assessment, to identify if the benefits outweigh the costs of undertaking an intervention or which of the many interventions that could be undertaken is the best use of scarce health resources.

To ensure that policy makers are provided with consistent evidence, it is important that costing studies use comparable and appropriate methods, regardless of the form of analysis being used. Otherwise it is not possible to compare the efficiency of the various competing alternatives or be sure that the interventions claimed to be cost-effective have been analyzed in an appropriate manner.

A further complication relates to the fact that thorough economic evaluation of health interventions require skilled economists and can be relatively expensive to undertake. In many countries, the necessary financial and human resources are not available and policy makers must draw on the results of studies undertaken in other settings and try to apply them to their own. For this reason, it is not only important to use comparable methods but also to report results in a way that allows policy-makers to assess the replicability and transferability of results between settings.

The need for consistency and standardization of methods has been recognized for some time and has led to the development of several sets of guidelines for economic evaluation and for costs. Despite this, considerable diversity is still apparent in applied studies. Some of this diversity might be defensible, and some might not be.

The overall objective of the thesis is to identify the variability in costing methods used in costing and cost-effectiveness studies and to understand the possible reasons for this; to test the practicality of applying standard methods when conducting costing studies in different settings and in using the results

[i] This is used to describe any programme or activity designed to improve or maintain health, be they promotive, preventive, curative or rehabilitative

to provide policy-relevant information; to understand the determinants of variation in unit costs; and finally to explore valid methods of transferring cost information from one setting to another, when local data do not exist.

The goal is to develop methods and cost data in a way that allows the costs and cost-effectiveness of many different types of interventions to be compared, and improves the methods of transferability and generalizability of the results within and across settings. Only then will the results of costing exercises be of widespread practical value to policy-makers and researchers.

1. 2 Methodological issues

Possible reasons for methodological differences in applied studies

There could be a number of reasons for the observed differences in the methods used in applied studies. They may be related to disagreements on what is appropriate, reflected by varying recommendations in guidelines on specific costing issues. Other possibilities include: areas where guidelines agree on the principles, applied studies follow the recommendations but vary in the detail of how they follow them; areas where there is agreement among guidelines but applied studies do not follow the recommendations; and areas where applied studies use different methods but the subject is not discussed in guidelines.

The importance of assessing the nature of these variations stems from the fact that, to the extent possible, differences in reported costs across interventions and settings should reflect only variations in the quantities of inputs used and/or their prices rather than the methods used to collect or evaluate them. Therefore, highlighting differences in methods that are known not to be best practices or for which recommendations exist but are not being followed is essential. Selected examples of common methodological differences that can be found in applied studies are discussed below.

1.2. 1 Capacity utilization

Differences in the costs and cost-effectiveness of health interventions across and within settings may be due to differences in capacity utilization in the facilities in which the studies were undertaken rather than any intrinsic differences in the actual resources required or the effectiveness of the interventions. The impact on costs of differences in capacity utilization was illustrated in a study of health centre costs and efficiency in Egypt. The cost per visit in 1994 varied from 3 to 60 Egyptian pounds depending on the level of capacity utilization, measured by the number of visits per physician per day (ranging from 3 to 16).[1]

This has been recognized by major costing and cost-effectiveness guidelines.[2,3] Despite this, very few applied studies report the capacity utilization that drives their numbers,

or attempt to adjust their costs to a standard level of capacity utilization before comparing different alternatives or comparing their results to other studies. Complying with the recommendations of guidelines is the obvious solution to increase the usefulness of the results presented in applied studies. In addition, exploring ways to account for the level of capacity utilization when estimating and extrapolating unit costs would be an important contribution to this area of work. This would be particularly useful with the increasing demand for country-specific cost estimates such as the cost of scaling up interventions to meet the Millennium Development Goals, or the WHO 3 by 5 goal of covering three million AIDS patients with antiretroviral (ARV) treatment by 2005.[4]

1.2. 2 Practicality of applying standard methods across different settings, e.g., measurement of staff time

There is very little discussion in guidelines about valid ways of allocating shared costs, particularly staff time, to specific interventions. Accordingly, different methods have been used in the literature. Examples of methods to allocate staff time include personnel interviews, time and motion studies, self-administered time logs, structured questionnaires and interviews with beneficiaries or with staff.[5,6] According to Bratt et al,[7] none of the first three produces estimates that are comparable to the benchmark, the time and motion study. The personnel interview method is particularly weak because it substantially overestimated contact time between staff and clients while underestimating non-productive time. In fact, the time and motion method recorded between six and twelve times as much "non-productive" time as the other three instruments.

Despite this, these methods are still being used in applied studies, which limits their validity and generalizability. Clearly the gold standard method, the time and motion study, is the most expensive option — it would be prohibitive to recommend it for all studies. Alternatively, understanding the determinants of variation in staff time from available time and motion studies may offer an attractive alternative to predict staff time for similar settings (always with a check on the degree of similarity) and interventions, rather than falling back to the less valid alternatives.

1.2. 3 Determinants of variation in unit costs

Studies often do not have the capacity to estimate health facility costs in their settings based on a large, randomly selected, number of facilities. Cost estimates are sometimes based on a small number, sometimes only one observation. It is also common practice to use average unit costs reported in previous studies as the basis for the analysis. A main limitation of these approaches is that average costs based on an unrepresentative sample can be biased and hide important differences between and within health facilities. For example, using more expensive equipment, higher proportions of specialized staff or performing a larger number of diagnostic tests all contribute to differences in average costs between and within hospital departments. Understanding the determinants of variations in unit costs across and within health facilities can provide a means to adjust unit costs so that the results of different studies

are comparable and to add more rigour and external validity to the extrapolation of costs across studies or settings.

1.2. 4 Methods of transferability of cost data and results

Guidelines for cost-effectiveness analysis have warned about making uninformed generalizations of study results to other settings,[8] and analysts have pointed out various reasons why generalizations can be potentially misleading,[9] but there has been no explicit recommendations on how to actually transfer unit cost estimates. The following are the most common recommendations of guidelines for extrapolating cost results to settings where data do not exist: use official exchange rates,[10] for components or subcomponents of costs, to extrapolate the results to other settings; use the same or different factor inputs across settings with prices modified to take into account prices in the new setting; or separate costs into traded and non-traded components and use the official exchange rate to translate the former and purchasing power exchange rates for the latter.[11-13]

Most of these methods, except those involving modification of factor inputs, ignore other determinants of variation in unit costs, apart from the price level. Understanding the determinants of these variations and controlling for them in the extrapolation process is worth exploring.

Another observation related to the transferability of assumptions or cost information from one study to another relates to estimating hospital unit costs. Admittedly, some of the more accurate methods of estimating costs are expensive or difficult to undertake, and budget constraints influence the way a study can be undertaken. Accordingly, some studies have used simple rules-of-thumb, e.g., the cost of an inpatient bed day was assumed to equal that of three outpatient visits, to allocate costs between inpatient and outpatient departments in a hospital instead of using the gold standard step-down allocation method.[14;15] The validity of applying the same ratio to all types of hospitals within a country has been questioned by Lombard et al (1991), who showed that in Cape Town Province of South Africa, the relationship between the cost per inpatient day and outpatient visit varied by type of hospital. It also differed from the commonly used 1:3 ratio – the cost per bed day was found to vary between 1.4 to 2 times the cost per outpatient visit.[16]

This raises the question of whether the countries seeking these data to make decisions about how to allocate resources between interventions should simply wait until full costing studies had been undertaken, or whether there are other alternatives for the short term that provided relatively reliable estimates while awaiting the full costing studies. To date, little of this type of work has been undertaken.

1. 3 Research questions

From the above discussion several questions may be formulated. The following are addressed in this thesis:

1. What are the reasons for variations in the costing methods used in cost and economic evaluation studies?
2. What is the impact of methodological differences on the results and their validity?
3. Short-cuts or rule of thumb approaches to costing; are they valid alternatives to more detailed costing studies?
4. What are the determinants of variation in unit costs across settings and how can they can be used in the transferability of cost estimates across settings?

These questions are all related to the central theme of the thesis:

Sources of variability in methods used in costing studies, implications for the validity and transferability of results.

1. 4 Thesis outline

The subsequent chapters of this thesis are organized as follows. Chapter 2 explores the sources of variations in costing methods used in applied studies, discusses the nature of these variations relative to the recommendations of costing guidelines and the possibility of reducing some of them. Chapters 3 and 4 explore the practicality of applying standard methods when conducting costing studies in multiple settings and in using the results to provide policy-relevant information. A set of empirical studies are then presented. Methods of extrapolating unit costs when local data do not exist are examined in Chapter 5. The validity and generalizability of some short-cut costing approaches such as using "rules-of-thumb" to estimate hospital unit costs for outpatient and inpatient departments are explored in Chapter 6. This is taken a step further in Chapter 7 by exploring the determinants of variations in unit costs across hospital inpatient departments. Moving from hospital to primary facility unit costs, Chapter 8 explores the effect of adopting new interventions on the way staff spend their time and on estimating unit costs of new services provided at primary facilities. Finally, Chapters 9-12 illustrate how a standardized set of costing methods can be used in practical studies to develop important policy conclusions.

2

Cost-effectiveness Analysis: Can We Reduce Variability in Costing Methods?

Reprinted from International Journal of Technology Assessment in Health Care, 19 (2): Adam T, Evans DB and Koopmanschap MA. Cost-effectiveness Analysis: Can We Reduce Variability in Costing Methods? © 2003, with permission from Cambridge University Press.

CHAPTER 2

Chapter 2. Cost-effectiveness analysis: Can we reduce variability in costing methods?

Summary

The need for consistency and standardization of methods for economic appraisals has been recognized for some time and has led to the development of several sets of guidelines for economic evaluations and for costs. Despite this, considerable diversity is still apparent in applied studies. Some of these diversities might be defensible, and some might not. The objectives of this paper are to explore sources of variations in the methods used in applied studies, and to discuss the nature of these variations and the possibility of reducing some of them.

We first use a systematic approach to identify the major sources of variation in costing methods used in applied economic evaluations. We then compare the methods used to the recommendations made in available guidelines.

Four possible sources of variation are identified. The first is where guidelines do not agree in their recommendations so it is not surprising that applied studies use different methods. The second is where guidelines agree in principle but provide little detail on how to comply with their recommendations; and the third is where a particular methodological issue is not discussed in guidelines. The fourth reason is simply lack of compliance with accepted guidelines.

Variability in costing methods used in applied studies raises questions about the validity of their results and makes it difficult to compare the results of different studies. We discuss the implications for the transferability and generalizability of results and suggest ways to minimize the variability in the methods so that the results of costing studies and economic evaluations can be of more value to policy-makers.

CHAPTER 2

2. 1 Introduction

Information on the costs of health interventions is valuable to health decision-makers for two reasons. The first is for budgeting purposes, to identify the resources necessary to undertake or sustain an intervention. The second is for efficiency assessment, to identify if the benefits outweigh the costs of undertaking an intervention or which of the many interventions that could be undertaken is the best use of scarce health resources. This paper is concerned with measuring costs for efficiency analysis.

To ensure that policy makers are provided with consistent evidence, it is important that costing studies use comparable methods, regardless of the form of efficiency analysis being used, e.g., cost-effectiveness, cost-utility or cost-benefit analysis. Otherwise it is not possible to compare the efficiency of competing alternatives or be sure that the interventions claimed to be cost-effective (or good value for the money) have been analyzed in an appropriate manner.

A further complication relates to the fact that thorough economic evaluations of health interventions require skilled economists and can be relatively expensive to undertake. In many countries, the necessary financial and human resources are not available and policy makers must draw on the results of studies undertaken in other settings and try to apply them to their own. For this reason, it is not only important to use comparable methods but also to report results in a way that allows policy-makers to assess the replicability and transferability of results between settings.

The need for consistency and standardization of methods has been recognized for some time and has led to the development of several sets of guidelines covering methods of economic evaluations and costs[2;3;10;13;15;17-30] and GISF Italian Group for Pharmacoeconomic studies (unpublished document, 2001). Despite this, considerable diversity is still apparent in applied work. Studies use different frameworks as the basis of data collection, different methods to collect data, different forms of analysis and a variety of ways to present their results. This is a significant limitation to the comparability and transferability of the results, which in turn restricts the value of such studies to decision-makers.

The purposes of this paper are, therefore, to:
1. summarize the main sources of variation in the way costs are estimated and presented in economic evaluations of health interventions;
2. identify possible causes of these variations;

3. suggest ways of reducing the variability so that the results of costing studies and economic evaluations can be of more value to policy-makers.

2. 2 Methods

Over the last decade, a number of systematic reviews of costing studies undertaken for the purposes of economic evaluation have been published. We do not seek to duplicate these exercises, but draw on their conclusions. To do this we identified recent review articles in which methodological issues surrounding economic evaluations of health interventions were discussed. Articles were identified by searching Medline and Econlit databases from 1990 to the present. Initially we used the following search terms: "Costs and Cost Analysis" AND method* AND (review or compar*). This search retrieved 2655 studies that involved costs in different ways but did not capture reviews of costing and cost-effectiveness studies. We then searched using "Costs and Cost Analysis" as a mesh word which we identified as a major topic. "Methods" was used as a qualifier and "review" was selected as the type of publication. English, French or Arabic language papers were considered, but none were found in French and Arabic. This search identified 87 articles of which only 16 compared costing methods across studies.

To exclude the possibility that some articles could have compared costing studies but were not classified as reviews, we then searched using the same criteria but substituting "compar*" for "review" as publication type. Only one additional review was found resulting in a total of 17 articles,[31-45] covering more than 600 economic evaluations or costing studies. Most focused on interventions against particular disease complexes such as parasitic diseases, tuberculosis, HIV-AIDS, hypertension, cancers or antenatal care, so they covered a wide range of interventions - preventive, curative, palliative and rehabilitative. Because our focus is on costs, issues relating to effectiveness are not reported here.

These reviews identified several types of variations in costing methods used in applied studies. There could be a number of reasons for this. It might be related to disagreement between guidelines. If so, variation in methods used in applied studies could be expected. In order to identify additional reasons for variability, we defined three other possibilities - variations associated with: areas where guidelines agree on the principles, applied studies follow the recommendations but vary in the detail of how they follow them; areas where there is agreement among guidelines but applied studies do not follow the recommendations; and areas where applied studies use different methods but the subject is not discussed in guidelines.

To allocate the observed variations in the applied studies to these four possibilities, we needed to identify the major guidelines. This was done by searching Medline and Econlit for guidelines and reviews of guidelines. Two search strategies were used. In both "Cost and Cost Analysis" was used as the mesh term, identified as a major topic. In the first, "guideline* or method* or standard*" were used as key words. In the second, the key words were "generalis* or generaliz* or transfer*" to incorporate articles discussing generalizability and transferrability of results. We also identified a number of guidelines through contact with regulatory authorities in various countries and included in our review a number of recent studies that compared the recommendations of these guidelines.[46-51]

Review process

The recommendations of the guidelines were summarized based on the main methodological questions related to costs described in Drummond et al.[3] and Gold et al.[2] and taking into account the fact that the main interest of this paper is the question of comparability and transferability of results. Six major methodological areas were identified: the framework of analysis; type of costs included; data collection methods; valuation of identified costs; methods of data analysis; and ways of reporting the results.

The next step was to identify the nature of the recommendations made in the guidelines in each of these methodological areas in order to allow the variations observed in the applied studies to be linked back to the guidelines. Finally, the variation observed in the applied studies in each methodological area was identified. The aim was not to record the frequencies with which different methods were used but to present the main issues where it has been established that there remain significant variations in methods and/or presentation of results.

2. 3 Results

The results are summarized in Table 2. 1. The first column describes the six major methodological areas divided into sub-components. The second summarizes the nature of the recommendations made in the guidelines, while the third reports briefly the nature of the major variations observed in applied work. The reasons for the observed variation in applied studies are related to the following three categories which are discussed in detail below: 1. Disagreement between guidelines; 2. Guidelines agree on principles but do not discuss how to comply in practice; 3. Guidelines agree, but studies do not

follow the recommendations. The fourth category identified earlier - methodological issues that are not discussed in guidelines - was a theoretical possibility but we did not find any examples in our review. Therefore, we do not consider it further.

2.3. 1 Disagreement between guidelines

On a number of issues, the guidelines do not agree or do not provide clear, unambiguous guidance and this is reflected in applied studies. Three inter-related areas of disagreement or ambiguity are the treatment of productivity loses/gains, the incorporation of informal care-givers time, and costs incurred in added years of life gained by an intervention.

Treatment of productivity losses/gains. On the first, guidelines do not agree on whether changes in market production – e.g. reduction in work time necessitated by seeking care or increases in work time resulting from improved health after an intervention - should be included in the numerator of the cost-effectiveness ratio. Gold et al.[2] argue that only losses due to time spent seeking and obtaining care should be included. Other types of productivity changes, such as increased ability to return to work, should not be included because they are already captured in the QALY weights of the denominator. Netting them out of the costs in the numerator would result in double counting.

On the other hand, Drummond et al.[3] argue that productivity changes due to care-seeking are not relevant since patients are already off work because of their condition – the time spent seeking care does not result in any additional productivity loss. The exception is for preventive interventions such as population screening where healthy people using the services and these costs should be included. Further, Drummond et al.[3] argue that productivity changes occurring as a result of the intervention should be included if they are important, but they provide no instruction to the analyst on when and how to do so.

The Australian guidelines take another perspective recommending not to include either type of productivity effect in the base analysis, but that analysts can report them separately if they can show rigorously that they are important.[19;21;22] The Canadian guidelines, similar to Drummond et al.,[3] discuss the different theoretical arguments about incorporating these costs without providing guidance on how and when to include or exclude them.[18;24;52] Other guidelines limit their discussion to patient and family time spent seeking and obtaining care recommending that they should be included in the numerator.[10;30]

The guidelines that recommend including productivity effects do not provide guidance about how to value lost or gained time. Some recommend using the frictional method proposed by Koopmanschap et al.[53;54] which assumes that productivity costs only occur during the time it takes to replace a worker - the friction period - but they do not discuss how to apply with this recommendation.[2;3;18;20;24;25;52] Others recommend using the human capital approach, by applying average wage rates for people who are in paid employment and one of two competing methods for valuing the time of people who are not in paid employment.[30]

Informal care giver time. The second related area of ambiguity is whether to incorporate and how to value informal care, i.e., health care provided by nonprofessionals such as family and friends of a patient, for which they are not financially compensated. Some guidelines recommend imputing a monetary value for this time[10;18;30] others only describe the different theoretical arguments without making any recommendations,[2;3] and some do not mention the topic.[17;20;21] Where guidelines recommend including these costs, guidance is rarely provided on the appropriate shadow wage to use,[2;18] restricting discussion to describing the different ways that have sometimes been used in applied work.[3;10;30]

Costs incurred in added years of life. The third area is whether future health care costs in years of life extended because of an intervention should be included, and if so, which ones. Most guidelines agree that health care costs for diseases or conditions related to the intervention should be included, but there is no agreement about costs incurred for unrelated diseases and non-health care expenditures (e.g., food, clothes, and housing) in added years of life. A number of guidelines simply report the theoretical disagreements without making specific recommendations on whether to incorporate these costs. They leave it to the analyst's discretion, with the recommendation that the sensitivity of the results to the inclusion or exclusion of these costs should be explored.[2;3;18;19;22;24;52]

Practice in applied studies. Not surprisingly, this disagreement and lack of guidance reflects itself in the mixture of methods used in applied studies, where some include productivity costs associated with time lost in seeking care and treatment, while others include the time spent caring for family members.[31;34;40;41;43;55] Others net out from the costs of the intervention the productivity gains expected to emerge from prevention, though this is not common.[35] It is relatively common to net out savings in health costs in added years of life due to prevention and most studies limit attention to costs due to related illnesses. However, examples can be found where costs due to unrelated illness in added years of life, or even additional consumption unrelated to health care, are added to the costs of an intervention.[32;35;40] Where

productivity or time costs were included, different methods of valuation can be found, ranging from valuing time at the average wage of unskilled labour to using sex-specific average earnings for the population as a whole.[41;55]

2.3. 2 Guidelines agree in principle but provide no detail of how to comply

Other sources of variation seem to be related to the fact that while guidelines agree in principle, they provide little detail of how to comply with the principle. Three important examples relate to the allocation of some types of shared costs (e.g., staff time or overhead costs), the question of capacity utilization, and the use of shadow prices.

Allocation of shared costs. Some guidelines simply state that shared overhead costs should be allocated to each specific intervention. Others describe in addition the various ways of doing this without making recommendations about which is preferable.[2] Some provide examples using the different methods of allocating overhead costs using such techniques as direct, step down or simultaneous allocation.[10;26;47] But there is no discussion of the validity of the estimates obtained from these methods, how they might differ and which one is preferable. In particular, there is very little discussion in guidelines about how to allocate shared labour costs to specific interventions in a valid way. Accordingly, applied studies have used different methods to allocate overhead costs.[31] In many other cases, published studies do not clearly show if shared costs have been included or how they were allocated to the intervention, so it is not possible to be confident of their validity.[34;43] Similarly, there are wide variations in the methods used to allocate staff time in the published literature. Examples include personnel interviews, time and motion studies, self-administered time logs, structured questionnaires and interviews with beneficiaries or with staff.[34;39]

Another example is the allocation of programme costs[1], sometimes considered as a component of administrative costs. Most guidelines give examples related only to the allocation of shared/overhead costs incurred at the site of delivery of the intervention - e.g., allocating hospital administration to the level of the ward – rather than higher level shared costs. Some guidelines assume the problem away by arguing that costs that are similar between the interventions compared in a study can be ignored. Apart from a few notable exceptions,[10;13] details of how to take programme costs into account cannot be found in guidelines.

Capacity utilization. Variations in estimated costs and cost-effectiveness might be due to differences in capacity utilization in the setting in which the studies were undertaken rather than to any intrinsic differences in the

efficiency of the interventions. A hospital-based intervention might seem very inefficient in one setting simply because it is located in a hospital that has a 50% occupancy rate rather than because the intervention is inefficient of itself. This is recognized by two of the major guidelines.[2;3] For example, Gold et al.[2] recommend that capacity utilization should be identified and reported when estimating the costs of an intervention or service. If information on capacity use is not available, they recommend using a benchmark rate of 80% under an apparent assumption that in the long run, that is the approximate rate that applies in the US. Despite this, very few applied studies report the capacity utilization that drives their numbers, or attempt to adjust their costs to a standard level of capacity utilization.[56]

Shadow prices. A final example of where guidelines agree but do not provide detail on how to follow the recommendations relates to the use of shadow prices. All guidelines argue that for the social perspective, opportunity costs should be used to value inputs rather than financial expenditures. Most recognize that there are market imperfections and that the use of shadow pricing might be needed. However, apart from discussing the opportunity cost of labour,[2;3;10;13;30] little detail is provided on how to shadow price other inputs or how to value traded and non-traded goods.[26]

Not surprisingly, very few applied studies[41;43] have tried to separate traded goods from non-traded goods, for example. Many studies have used the official exchange rate to present their cost estimates in foreign currency units (e.g. US dollars) with no discussion of whether this is appropriate or whether shadow exchange rates (or conversion factors) should be used instead.[43] It is rare to find examples of the type of rigorous shadow pricing common in economic appraisals undertaken in other sectors.[57;58]

2.3. 3 Guidelines agree but applied studies do not follow the recommendations

We now turn our attention to variations in costing methods which need not occur, particularly related to areas where there is agreement in the guidelines but variation in the applied studies. Examples relate to the perspective of the analysis, the types of costs included and using the ingredient approach to collect and report cost information.

Perspective of the analysis. Despite the fact that all guidelines argue that the social perspective is appropriate (even if they also recommend other perspectives in addition), many applied studies include only health care costs borne by the provider.[31;33;41;43] In fact, in some studies it is not possible to determine what perspective is being taken.[33;36] In selected circumstances it

might be justifiable to exclude some of these other costs, but these circumstances are limited to cases where the other costs are a very small proportion of total costs, or if including them will simply strengthen the results (e.g., increase an already high cost-effectiveness ratio).[3;59]

Costs included. A second example is where components of direct costs are overlooked. Some studies exclude donated items (such as drugs, vaccines, supplies or equipment or the costs of recruitment and invitation of target groups, e.g., in screening programmes.[41;43] Some use charges instead of costs with no comment on how they might differ from actual costs[37] despite the fact that most guidelines provide examples of how to adjust charges to obtain more accurate estimates of costs.[47;60]

Ingredients approach. Virtually all guidelines recommended that costs be estimated and reported using the ingredients approach, in which physical quantities of inputs are reported separately to their unit prices. But it is uncommon to find published economic evaluations providing this information apart from few notable exceptions.[34;37;41;43]

Table 2. 1 [i]: Comparison between recommendations made in guidelines and nature of variations in applied studies.

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
I.) The framework of analysis		
a. Perspective: societal versus provider	– Guidelines agree that the perspective needs to be clearly stated in a study. They generally recommend that the societal perspective is the appropriate one to use but allow that other view points to be included depending on the question to be answered.	– Most studies adopted the provider's perspective (health system), a smaller number used some form of societal perspective. – Other perspectives used were the employer's (private), patient or community. In many studies the perspective is not specified or is difficult to determine.
b. Choice of comparator:	– Most guidelines recommend using an incremental analysis, comparing the new intervention to the best current alternative. – The recently published WHO guidelines argue that in addition to incremental analysis, comparison to doing nothing is always relevant.	– Some compare a new alternative to current practice using incremental analysis. Others compare a new alternative to doing nothing.
II.) Types of costs included		
a. Overhead costs	– Most guidelines agree that overhead costs should be included. But alternative	– Often difficult to understand whether overhead costs were included or not

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
	methods of allocation and their validation are not fully discussed – The main discussion involved methods of allocating overhead costs in health facilities, e.g., hospitals. Methods of allocating shared costs at higher levels, e.g., the district or national-levels, were rarely discussed.	due to aggregate presentation of results. – Where it is clear, many studies apparently overlook part or all of these costs, e.g., administration costs at higher levels of the health system, e.g., planning, lobbying, training of medical staff, supervision etc..
c. Other shared costs including allocation of labour time	– Methods to collect and allocate labour costs were briefly described in guidelines. However, validation of these methods was not discussed, nor recommendations made about which to use.	– In general, methods of allocation of shared costs were not specified. Among the methods used were: staff rosters, time and motion studies, personnel interviews or self-administered time logs kept by clinician.
d. Indirect costs and inclusion of productivity gains into net costs (for valuations of time cost see section IV b.)	– Guidelines do not agree on whether indirect costs due to illness or informal care should be included. – Volunteer time should always be identified and included unless deemed to be minimal	– Most do not include indirect costs or consequences – Some include lost time in seeking care (mainly by patients and to a lesser degree by caregivers). – Costs of lost non-market time, including leisure time, is rarely included. – Very few included volunteer time.
e. Health care costs for unrelated illness in the	– Most agree that future costs associated with current illness should be included but there	– Very few included costs in future life gained when evaluating interventions

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
added years of life	is no agreement whether future costs due to unrelated illness should be incorporated in the analysis.	with long term effects. Of those who included these costs, some included medical costs due to related illness, or medical and/or non-medical of unrelated illness. Some did not specify which future costs were included
III.) Data collection methods		
a. Sources of cost data	– Some guidelines describe the different sources of data collection but issues regarding the validity and the reliability of data collection instruments are not discussed, e.g., recall bias, question formats, answer formats etc.	– Sources of cost data included: expenditure records, case registers, interviews with clinicians or patients, surveys and questionnaire. Others used assumptions made by investigators, projections and modeling or clinician consensus or resources use from trials.
IV.)The valuation step		
a. Bottom-up versus top-down [ii]	– Methods and issues to consider when using either approach were not discussed, particularly related to the determination of capacity utilization when collecting and reporting unit costs [iii]. - The importance of reporting capacity	– Some used a top-down approach, some used a bottom-up, others used a combination of bottom up and unit costs from the literature, while still others used a combination of bottom up to collect patient costs and a step

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
	utilization and the methods used to measure it were not discussed in guidelines.	down costing to allocate overhead costs.
b. Price adjustments: 1. price distortions (e.g., shadow prices, transfer payments) 2. Exchange rates	- Generally agree that prices should theoretically be adjusted for market distortions. Some say that the gains from doing this are not worth the effort. Little detail is provided on how to do it, for example, methods for comparing traded goods with non-traded goods are not discussed. - Most guidelines recommend presenting results in local currency. No discussion of how and when to report in foreign exchange currency (e.g., US $) or international dollars.	– Generally no discussion on whether market prices are a realistic reflection of opportunity costs. – Details are rarely provided about how traded and non-traded goods have been treated. – Where studies report results in foreign currency units, official exchange rates have mostly been used. Rare use of purchasing power parity exchange rates
b. Valuation of time costs (e.g., volunteer time, home production, productivity loss due to illness)	– Guidelines discussed the different valuation methods, e.g., human capital approach or willingness to pay, but the majority did not recommend which one to use. – No discussion on how to collect wage rates in developing countries particularly in agricultural areas or where labour markets do not exist.	– When included, the valuation methods were not systematically mentioned. Among the methods used were market values of professional labour that could substitute this input, or the wage rate which this person could earn in alternative paid work.
c. Capital costs	– Guidelines discuss alternatives, e.g., rental versus annualized costs, but there is no recommendation on how to collect this data	– Sometimes it was difficult to determine whether capital costs were included or not. When they were included, many

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
	in a valid way e.g., appropriateness of using rented equivalents in developing countries, especially rural areas where markets do not exist. Or using replacement versus historical costs when calculating annualized costs. – The recommendations vary from restricting analysts to using a specified discount rate and life spans of capital items to leaving it open to analysts to choose which ones to use within a certain range.	studies did not systematically report how they were treated, i.e., using rental values or annualized costs and in the latter case which discount rate was used as well as the useful life of capital items.
d. Prices or charges	– Guidelines agree that resources that are important for the specific study should be valued based on costs not charges, and if charges are to be used they should be adjusted. – Methods of adjustment of charges to approximate costs were discussed for the US.	– In a large number of studies (mostly done in the US), costs were based on charges with no comment on how they differ from actual costs.
V.) The methods of data analysis		
a. Discounting costs	– All guidelines agree that costs should be discounted and a range of rates is usually given [iv]	– Many studies discounted costs. Some did not when they should, e.g., when they aggregate annual costs of a programme for several years to present

Methodological areas	Nature or recommendations made in guidelines	Nature of variations in applied studies
		value. Studies used different rates ranging from 2.5 to 14%.
b. Capacity utilization	– Guidelines recommend that capacity utilization should be identified when developing cost estimates of an intervention or service. However, none of the guidelines provide guidance on how to measure capacity utilization or how to use the results for policy analysis.	– This issue was not discussed in reviews, and rarely reported in applied studies. – Applied studies rarely report capacity utilization associated with their estimates.
c. Sensitivity analysis and other statistical analysis of variations in units costs	– All guidelines agree that sensitivity analysis should be performed. Only one country specific guideline provided recommendations, favouring multi-way analysis.	– Not systematically done and when done, it was often done inadequately. Often authors did not provide justification of the ranges of values employed in the analysis and many key study parameters were not accounted for. – One-way sensitivity analysis was the most commonly used. Other methods were multi-way analysis and threshold analysis.
VI.) Reporting results		
a. Ingredient approach and transparency in describing methods and results	– General agreement on the usefulness of using the ingredient approach in estimating and reporting results.	– It is rarely possible to identify the physical inputs separately from unit costs in published studies.

Table footnotes:

[i] The classification of this table is based on the main methodological issues surrounding the steps required to estimate costs described in Drummond et al.[3] and Gold et al.[2]

[ii] Bottom-up approach means that the unit cost of each service in each product line is computed as the sum of the resource inputs used in that service's production. The top-down approach uses total costs figures obtained from the financial accounts of the health care provider as the primary sources for determining the unit costs per unit of output.

[iii] The bottom-up approach runs greater risks of forgetting slack time than the top down approach. For example, one of the methods of allocating staff time is based on observing the average time spent in patient encounters. If analysts do not add a proportion of down time to the overall estimate of staff time/patient, this will underestimate labour costs and biases the cost-effectiveness ratio. On the other hand, top down may give biased estimates in case of heterogeneous production (e.g. radiology)

[iv] It is still debated whether or not outcomes of interventions should be discounted. This is not addressed here.

2. 4 Discussion and policy implications

Despite the presence of guidelines on economic evaluation and costing methods, there is still considerable variation in the methods used and the approaches taken in applied costing studies.[32;33;36;37;40;41;44] This paper has identified a number of possible reasons for this by relating the observed variations to the recommendations of the guidelines. The categories identified were: where there is disagreement between guidelines; where guidelines agree on principles but do not discuss how to comply in practice; and where guidelines agree but studies do not follow the recommendations.

The variability associated with the first is likely to remain for the foreseeable future because it is related to fundamental disagreement among economists on questions like the boundaries to be drawn around the analysis. As mentioned above, economists simply do not agree on whether and how to incorporate productivity losses in seeking or obtaining care, or whether productivity gains in added years of life resulting from an intervention should be subtracted out of the costs of the intervention (called net costs).[59;61]

Economists also disagree on the theoretically correct approach to treating time lost in providing informal care,[62;63] or if future unrelated costs in added years of lives due to an intervention should be included in the economic evaluation of health interventions.[64-66] It is not possible for us to resolve these debates here. We simply note that there will continue to be some variations in costing practices observed in applied studies until these theoretical debates are settled.

The second source of some of the variability is linked to the fact that guidelines do not provide details of how to follow their general recommendations. This gives analysts the flexibility to comply in a number of ways. Does this create significant variation in the cost estimates? Bratt et al.[7] recently compared the estimates produced by different methods of allocating shared staff costs to specific interventions - provider interviews (PI), self-administered timesheets (ST), patient flow analysis (PFA) and an observational time and motion technique (TM) taken to be the benchmark. None of the first three produced estimates comparable to the benchmark. The PI method was particularly weak because it substantially overestimated contact time between staff and clients while underestimating non-productive time. In fact, TM recorded between six and twelve times as much "non-productive" time as the other three instruments.

This suggests that it may be important to reach agreement not only on the principles underlying costing studies, but also on the most appropriate ways to collect critical data. While it is necessary to allow some flexibility in

collecting costs, it is important to ensure that the methods used are valid. This is not generally discussed in guidelines, possibly because they have focused more on the big picture rather than on the specifics of how to comply with recommendations.

Admittedly, some of the more accurate methods of estimating costs are expensive or difficult to undertake, and budget constraints influence the way a study can be undertaken. Accordingly, it would be valuable to compare the gold standard methods to simpler, less costly alternatives e.g., using a simple rule-of-thumb to allocate costs between inpatient and outpatient departments in a hospital versus the gold standard step-down allocation method.[15;67] This will identify the tradeoffs that exist between the level of accuracy obtained and the costs of undertaking each method. To our knowledge, little of this type of work has been undertaken.

Similar conclusions apply to the area of capacity utilization which is rarely reported in the results of applied studies.[3;18;19;21;22;24;47;52] Yet the impact of variations in capacity use on costs can be substantial. For example, a study of costs per outpatient visit in 19 health centers in Egypt in 1994 showed that a major determinant of the differences in costs (ranging from three Egyptian pounds per visit to 60) was the variation in the number of visits per physician per day (ranging from 3 to 16).[1] The costs of facility based interventions, perhaps antenatal care, would differ greatly depending on what level of capacity use was assumed. Accordingly, it is important for guidelines to develop more specific recommendations on how to measure, report and adjust estimates taking into account capacity use.

Finally, a considerable reduction in the variability of methods used in studies could be achieved if future studies complied more closely to the guideline recommendations. For example, studies should ensure that all costs are included and that resources are valued in an appropriate manner. This task is facilitated by using an ingredients approach to measuring and reporting costs, as recommended by the guidelines. This will also help the generalizability and credibility of cost estimates. In this way, analysts in a different setting could determine if they would use the same types of inputs – e.g. perhaps they would use nurses rather than doctors to perform a particular task – and adjust the inputs accordingly. They could also adjust the prices used to their local situation, making the results more applicable to their settings. By presenting information on capacity utilization and the methods used to obtain those estimates, it will be possible for decision-makers to estimate the costs of the same programme in their setting after applying the average capacity use observed in their setting using the same measurement methods.

CHAPTER 2

The past few years have shown a considerable increase in the number of applied studies reporting the costs or economic efficiency of health interventions.[39] This reflects increasing interest in providing policy-makers with information to assist them in the efficient allocation of scarce resources. To ensure that they are provided with consistent evidence, it is important that costing studies use consistent and comparable costing methodology. Otherwise it is not possible to compare the efficiency of competing interventions or be sure that the interventions claimed to be cost-effective have been analyzed in an appropriate manner.

This paper suggests that it is possible to reduce some, though not all, of the observed variability in the methods used in applied work. This can be achieved, firstly, if guidelines – or companion volumes to guidelines - provide more detail about how to follow their recommendations. More empirical research is needed to guide this process, for example, on issues related to validation of different data collection and measurement methods. Particularly valuable would be efforts to validate relatively rapid, low cost data collection techniques compared to more expensive gold standard methods. Secondly, it is important that studies which do not comply with established standards are not funded or published. A more careful review process is required from funding agencies and journals. A minimum requirement should be that studies use an ingredients approach to measure and report costs, and that these data are provided to reviewers, to other analysts and to policy-makers as desired. This would reduce the number of studies published despite having apparently omitted important components of costs, or without enough detail to convince readers that their results are valid.

3

Does the Integrated Management of Childhood Illness Cost more than Routine Care?

Results from Tanzania

Based on: Adam T, Manzi F, Armstrong Schellenberg J, Mgalula L, de Savigny D and Evans DB. Does the Integrated Management of Childhood Illness Cost more than Routine Care? Results from Tanzania. WHO Bulletin, 2005; 83(5):369-77 ©. Published with permission from the Bulletin of the World Health Organization.

CHAPTER 3

Chapter 3. Does the Integrated Management of Childhood Illness cost more than routine care? Results from Tanzania

---

Summary

Integrated Management of Childhood Illness (IMCI) is a strategy designed to address the five leading causes of childhood mortality, which together account for 70% of the ten million child deaths in the world each year. Although IMCI is associated with improved quality of care, it is not yet widely adopted, partly because it is assumed to be more expensive than routine care. Here we report the cost of IMCI compared with routine care in four districts in Tanzania.

A plausibility design compared two districts with IMCI with two contiguous comparison districts where routine care continued. Total costs of child care were estimated from the societal perspective and were collected from the household, primary health facility, hospital, district and national levels.

The annual cost per child of caring for under-fives in districts with IMCI was US $ 11.19, 44% lower than in the comparison districts ($16.09) using routine care. The average cost per under-five visit to a health facility was $1.39 in facilities with IMCI, compared to $1.61 in comparison facilities. IMCI facilities spent 30% less per visit on drugs and vaccines than comparison facilities.

We found no evidence that IMCI was associated with higher costs than routine child care in the four study districts in Tanzania. Given the evidence of improved quality of care in the IMCI districts, it is important to take rapid steps to scale up the adoption and coverage of IMCI.

CHAPTER 3

3. 1 Introduction

More than 10 million children die each year before they reach their fifth birthday and 70% of these deaths are due to five conditions: diarrhea, pneumonia, malaria, measles and malnutrition.[68;69] The Integrated Management of Childhood Illness (IMCI) strategy seeks to reduce these deaths through improvement of skills of health staff, health systems and family and community practices. Tanzania is one of over 80 developing countries in which the Integrated Management of Childhood Illness (IMCI) strategy has started to be implemented. The Multi-Country Evaluation (MCE) of the IMCI is a set of studies in five countries using complementary designs that assess the effectiveness, cost and impact of the strategy, see www.who.int/imci-mce. It has already been shown that IMCI is associated with improved quality of care, something that should result in improved outcomes.[70-73] However, concern about the costs of implementing IMCI, especially in terms of human resources, is likely to have prevented some countries from adopting it.[74;75]

In Tanzania the study uses a *'plausibility'* design to document process and outcome indicators in two districts with IMCI (Morogoro Rural and Rufiji) and in two contiguous comparison districts (Kilombero and Ulanga) where routine case management continued. Plausibility statements are derived from evaluations that – despite not being randomized – are aimed at making causal statements using observational designs with a comparison group.[76]

Here we present results from the cost component of the MCE study in Tanzania, the first such study for which detailed cost data are available. The specific objectives of the costing study were to estimate the total economic costs of starting-up and implementing IMCI in a district —i.e., the full cost to society of IMCI-based services to children under five; and the additional economic costs (additional to those previously expended on under-fives) of introducing and running IMCI.

3. 2 Methodology

3.2. 1 Study setting

We compared two rural districts where IMCI has been implemented since late 1997 ("intervention" districts) with two neighbouring rural districts where implementation began in 2002 ("comparison" districts). The four districts had reasonably well-functioning health services, comparable levels of per-capita health expenditure, high utilization rates of government health facilities and

high coverage of selected key interventions (e.g., EPI). Large numbers of non-governmental health actors were also active in the districts, many of which were involved in health worker training and community activities, although their coverage was patchy. The two districts where IMCI was implemented had engaged in activities designed to strengthen district management skills, and had authority for priority setting and control over the health budget decentralized to them. A high level of coverage of IMCI training for health workers was achieved, but there had been no increase in the provision of under-five interventions at the community level, as opposed to the facility level. More detail about the study setting and IMCI as implemented in Tanzania can be found elsewhere.[73;77]

3.2. 2 Data collection and cost analysis

Cost data were collected for the start-up period of implementing IMCI (from 1996 to 1997) and subsequently for maintaining child health services including IMCI during the year of 1999. The start-up period is defined as the time from the national decision to implement IMCI to the time when IMCI was provided to the first under-five child through trained health workers in primary facilities. Annual costs for maintaining IMCI were collected for 1999. Costs are presented in 1999 US$. The start up costs were annualised and inflated to 1999 values using GDP deflators.[78]

Costs were estimated from the societal perspective and were collected for the national, district, hospital, health facility and household levels (Panel 3. 1). Cost at all these levels was summed to obtain the *total cost* to the district of providing care for under-fives. To allow comparison across districts, cost estimates were standardized to a hypothetical district with a population of 50,000 under-fives. This corresponds to a total population of around 300,000, which is roughly the average district population for Tanzania. Estimates of the *additional cost* to the district of implementing IMCI were based on the difference in cost of under-five care between IMCI and comparison districts. The total cost of care for under-fives in a standard comparison district was subtracted from the total cost in a standard IMCI district. The difference is the estimated change in under-five costs attributable to IMCI.[79]

Panel 3. 1 Cost data collected at each level

1. *National*: National costs of start-up and annual post-implementation costs of IMCI, and of other activities related to under-fives such as the Expanded Programme on Immunization (EPI), nutrition and malaria programmes, were collected using interviews and record reviews based on a national-level cost questionnaire.
2. *District:* District-level start-up and post-implementation costs of under-five care were estimated through interviews and record review using a district-level cost questionnaire.
3. *Hospital*: The proportion of under-five children admitted to hospital during the previous year was estimated through interviews with a representative sample of households using a household survey questionnaire.[80] This information was combined with local estimates of costs per bed-day and average length of stay in hospital [81;82] to estimate total costs of providing inpatient care for under-fives in each district.
4. *Primary facility*: Primary health facility costs at government health facilities were estimated through interviews and record reviews using the facility cost questionnaire during a cross-sectional survey of a representative sample of health facilities (75 facilities in total).[77;83;84] During the same survey the proportion of time health workers spent with under-fives and with over fives was collected through observation of health workers using time-and-motion study observation record forms. Primary health care costs at non-government facilities are partly represented as out-of pocket payments made at these facilities, collected at the household-level. It is not in the scope of this analysis however to determine the extent to which these out-of-pocket payments relate to actual cost per visit made at non-government facilities.
5. *Household*: Out-of-pocket payments for services provided at facilities that were not included in the above categories, and time spent in seeking all types of care, were estimated through interviews with a representative sample of households using a household survey questionnaire (around 2000 households in total).[80]

In addition, ordinary least squares multivariate regression analysis was performed to explore whether IMCI had an effect on the costs of providing services at health facilities, independent from the effect of other factors not related to IMCI such as facility size and availability of vehicles. Finally, sensitivity analysis was used to test the sensitivity of the results by using a range of values for the uncertain variables. The variables selected for uncertainty analysis were the useful life ofcapital inputs, district-level cost per child (i.e., the district component of total district cost per child), hospitalization rate and average number of facility visits per child per year.[77] WHO CostIt[85] and STATA software[86] were used for the analysis of cost data. Further detail of the data collection methods and analysis are given elsewhere.[77]

3.2. 3 Quality control and data processing

For national, district and primary health facility data, all forms were checked for completeness and consistency, and follow-up visits were made to re-collect inconsistent or incomplete data. During the household survey, a field supervisor checked all forms, sat in on one or two interviews, and made random re-visits to a sample of households each day.

Two data-entry clerks made double entries of the household survey into a FoxPro database system. The two files were compared, any inconsistency was verified with reference to the original forms, and range and consistency checks were carried out regularly. Excel was used to process data on national, district and primary health facility costs. Quality was checked visually and through range and consistency checks.

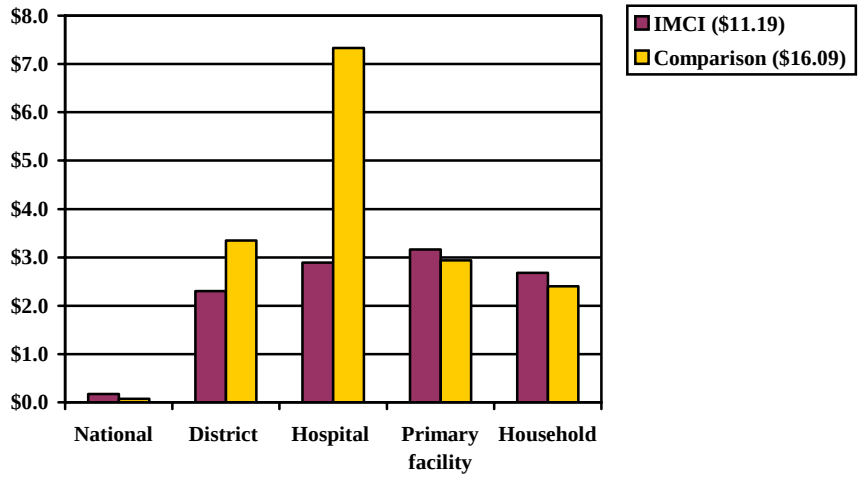
3. 3 Results

For 1999, the cost per child of caring for under-fives in the districts with IMCI was US$11.19, which was 44% lower than in the comparison districts ($16.09).[77] The lower cost per child in IMCI districts was due to lower hospitalization and administrative costs at the district level. There was no difference in costs incurred at primary care facilities and at the household level (see Figure 3. 1).

District-level costs were 50% lower in IMCI districts, linked to less frequent trips for supervision and drug distribution than in comparison districts. It is not possible to know if the lower costs resulted from the introduction of IMCI or were due to external factors. Similar costs of training were observed in comparison and IMCI districts during the study period. This was unexpected

given the emphasis of IMCI on training, but a wide variety of training courses were performed in comparison district for preventive, curative and administrative issues during the study period. These included training for immunization, case management of malaria and use of insecticide-treated bednets and the use of district Health Management Information System forms.

Figure 3. 1 Components of cost of under-five care per child in a standard[1] district in Tanzania (1999 US$)



[1] standard district with 50,000 under-fives.

Hospital-level costs were 250% higher in comparison districts than in districts with IMCI, not because of differences in the cost per under-five admission, but because more under-fives were hospitalized in the year prior to July-August 1999 in these districts relative to IMCI districts (6% in IMCI districts against 15% in comparison districts, $p<0.001$). Because this difference in hospital admissions may or may not have been related to IMCI, we also calculated total costs without the component of hospital costs, which resulted in a 6% lower cost per under-five child in districts with IMCI ($ 8.30) than in districts without IMCI ($ 8.76) (Table 3. 1).

At the primary health facility level, including both government health centres and dispensaries, IMCI facilities had a 16% lower average cost per under-five visit (including vaccination visits) than comparison facilities ($1.39 and $1.61 in IMCI and comparison districts respectively, $p=0.5$). The average number of

visits per child per year was 30% higher in IMCI districts (3.28) compared with comparison districts (2.49). The lower cost per visit together with the higher number of visits per child per year in IMCI districts explains the similar cost per child at government health facilities in the two types of district.

Table 3. 1 Differences between IMCI and comparison districts in Tanzania in the annual cost per child for health care, in 1999 US $.

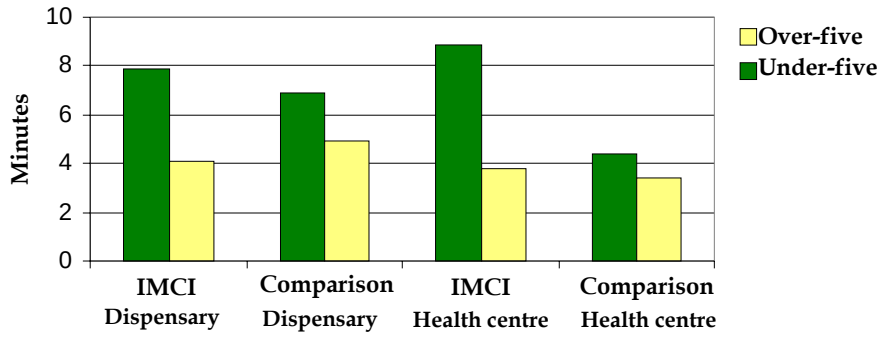
Level	Average cost **IMCI** districts	Average cost **Comparison** districts	**Difference** in cost per child in IMCI districts	**Ratio** Compa-rison : IMCI
National	$0.17	$0.07	$0.10	0.43
District	$2.30	$3.35	$-1.06	1.46
Hospital	$2.89	$7.33	$-4.44	2.54
Primary facility	$3.16	$2.94	$0.22	0.93
Household	$2.68	$2.40	$0.28	0.90
Total per child cost	$11.19	$16.09	$-4.90	1.44
Total *excluding* hospital costs	$8.30	$8.76	$-0.46	1.06

Estimates are adjusted for a standard district with 50,000 under-fives.

Results of the two major components of cost per visit at government facilities, i.e., personnel and drug costs, are presented in turn. With respect to personnel cost per visit, the time-and-motion study, where health workers are observed to record the time they spent on different activities, showed that IMCI health workers spent, on average, almost two more minutes per consultation with each under-five than did those in comparison facilities (8.2 vs 6.3 minutes, p=0.0003, see Figure 3. 2). This difference was largest in health centres, which received only 18% of the total visits by under-fives. It is worth noting, however, that health centre workers did not compensate by spending less time with over-fives (p=0.4). It appears, therefore, that the increase in time spent with children in health centres was due to a shift in the time spent in administrative activities or non-productive time, part of which was used to provide clinical services for under-fives. Because the longer time spent with under-fives was mainly observed in health centres, which receive a smaller proportion of under-five visits than dispensaries, overall, the average

personnel cost per under-five visit was similar between IMCI and comparison districts ($0.49 compared with $0.57 in comparison districts, p=0.41).

Figure 3. 2 Comparison of average time spent per consultation visit, by facility type and age group, in districts with and without IMCI in Tanzania.



With respect to drug costs, IMCI facilities spent an average of $0.29 on drugs and vaccines per visit, 30% less than facilities in comparison districts, although because of the considerable variation in drug costs per visit, the difference was not statistically significant (Table 3. 2). Further analysis of drug expenditure showed no difference in the availability or shortage of particular drugs between facilities with and without IMCI. Two separate analyses of prescription patterns, using different sources of information from both the quality of care[70] and costing components of the MCE study, confirmed a more rational use of antibiotics and injectables in IMCI facilities, suggesting higher efficiency and better health outcomes than in comparison facilities.

The results of the regression analysis to explore the relationship between total costs of under-five care at health facilities and factors such as whether the facility had implemented IMCI, facility type (health centre or dispensary) and availability of vehicles are shown in Table 3. 3. By taking into account differences in the other determinants of costs across facilities, the multivariate regression analysis increased the precision of comparison between the two types of management strategies, showing that total costs of under-five care and the cost per under-five visit were around 30% lower in IMCI facilities ($p<0.001$).

Table 3. 2 Comparison of average drug costs (1999 US$) per under-five visit to a primary government facility in districts with and without IMCI in Tanzania.

Category	IMCI district			Comparison district		
	Mean	SD	n	Mean	SD	n
Dispensaries	$0.30	$0.20	33	$0.46	$0.48	29
Health centres	$0.22	$0.19	6	$0.19	$0.11	6
Weighted average	$0.29	$0.20	39	$0.41	$0.45	35

Differences not statistically significant at the 0.05 level.

Table 3. 3 Regression analysis of determinants of primary health facility costs; dependent variable: natural log of total costs of under-five care, in 1999 US$

Adjusted $R^2 = 0.79$ F-statistic = 74.41 p<0.0001 N=73

Variable	Description	β coef	S.E.	t	p
IMCI	Dummy variable, IMCI=1	- 0.34	0.06	-5.22	<0.0001
Dispensary	Dummy variable, Dispensary=1 and health centres =0	- 0.65	0.12	-5.48	<0.0001
Vehicle dummy	Dummy variable, facility has vehicles=1	0.69	0.14	4.94	<0.0001
Log visits	Natural log of total under-five visits (vaccine and curative)	0.19	0.05	3.47	0.001
Constant		14.07	0.48	29.4	<0.0001

Only significant variables are presented.

Sensitivity analysis showed the importance of hospitalization costs in the interpretation of total district costs - the difference between IMCI and comparison districts was not sensitive to variation in parameters other than the assumption about rates of hospitalization.[77] If it is believed that the observed difference in hospital admissions per child was not related to IMCI, it can be concluded that there is no difference in the cost of under-five care in the two types of districts. Otherwise, the costs in IMCI districts are lower than in the comparison districts.

3. 4 Discussion and policy implications

We believe this study to be the first attempt to estimate the actual cost of IMCI implementation in comparison with routine care. Our estimate of the total economic cost of implementing IMCI and the additional cost compared to routine care showed that, for 1999, there is no evidence that IMCI was associated with higher costs. These findings were unexpected, as IMCI has often been assumed to be more expensive than routine care for under-fives.[74;75]

We found that hospital costs were a main determinant of the district cost per child of caring for under-fives. There are two possible explanations: (1) improved quality of care and drug availability for under-fives at IMCI primary facilities reduced the need for referral and subsequent admission to hospital; or (2) factors other than IMCI, such as differences in the quality of the hospitals in the different settings or access to them meant that children in comparison districts were more likely to seek care at hospitals. Even if we assume that this difference was entirely due to other factors, and exclude the hospital component from the analysis, total costs per under-five child in IMCI districts were still lower than in comparison districts (6%).

The US$ 11.19 cost per child of treating children under five using IMCI in Tanzania translates into a per capita cost of $ 1.79 compared to $2.56 for routine care. This is similar to previous estimates of the per capita cost of IMCI in resource poor countries.[87] In addition, the Tanzania evaluation had similar findings with respect to the lower drug costs associated with IMCI as previous studies.[88] However, our study reaches different conclusions regarding staff requirements than earlier work.[89] In Tanzania, health centre staff were able to accommodate IMCI within their usual working hours by reallocating part of their non-clinical or slack time to provide better care for under-fives. No additional staff were required during the period of the study.[90]

The effects of IMCI can be assessed in terms of changes in intermediate outcomes, such as improved quality of care at health facilities, or final outcomes such as changes in under-five mortality or DALYs averted. No information is yet available on the impact of IMCI on mortality. However, some information on the effect in terms of intermediate outcomes is available. In the Tanzania evaluation, a health facility survey was carried out in 2000 to compare the quality of case-management and health systems support in IMCI and comparison districts. The results indicate that children in IMCI districts received better care than children in comparison districts: their health problems were more thoroughly assessed, they were more likely to be diagnosed and treated correctly as determined through a gold-standard re-examination, and the caretakers of the children were more likely to receive

appropriate counselling and reported higher levels of knowledge about how to care for their sick children.[73] In Brazil and Uganda, as well as in Tanzania, there is evidence of better anti-microbial prescribing patterns with IMCI than with routine child care.[70] This suggests that IMCI is a cost-effective intervention in the Tanzanian setting. It is no more costly than standard care, yet achieves a better outcome in terms of quality of care. Given the evidence that improved health worker performance and prescription patterns lead to reductions in child mortality,[68] it is reasonable to claim that IMCI is no more costly than standard care and is more effective at saving lives.[68]

A number of qualifications should be taken into account when interpreting the results. Firstly, utilisation rates for public health facilities in these districts were high relative to those reported from other developing countries – 41% of children reporting an illness in the two weeks prior to the MCE household survey had been taken to a public health facility as the first point of contact for care.[80] Secondly, districts differed in ways that could affect the cost of child care (e.g., the number and proportion of facilities managed by non-government organizations and under-five hospital admissions). Thirdly, in the intervention districts IMCI was implemented concurrently — and because of — measures designed to strengthen district management such as evidence-based planning and expenditure mapping at district level. Our findings, therefore, can be interpreted as the costs of IMCI in the presence of a strengthened health system with adequate managerial capacity. In addition, an ideal study design would have assessed the costs of care over a number of years before IMCI implementation in all four districts. This would have allowed before-and-after analysis to control for any cost changes resulting from district-specific factors or from trends in costs or utilization over time. Because IMCI was already in place, in 1999 when this study commenced, this type of analysis was not possible.

In conclusion, we found no evidence that treating children using IMCI was associated with higher costs than routine care. The costs are either similar or lower in IMCI districts depending on the interpretation of the rates of hospitalization in those areas. Given the evidence of improved quality of care in the IMCI districts, more active steps should be taken to rapidly scale up the adoption and coverage of IMCI.

4

Effectiveness and Costs of Facility-Based IMCI in Tanzania

Based on: Armstrong Schellenberg JRM, Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O, John T, Charles S, Nathan R, Wilczynska K, Mgalula L, Mbuya C, Mswia R, Manzi F, de Savigny D, Schellenberg D and Victora C. Effectiveness of facility-based IMCI in Tanzania. Lancet, 2004 (9445);364: 1583-94.

CHAPTER 4

Chapter 4. Effectiveness and costs of facility-based IMCI in Tanzania

Summary

Integrated Management of Childhood Illness (IMCI) has been adopted by over 80 countries as a strategy for reducing child mortality and improving child health and development. It includes complementary interventions designed to address the major causes of child mortality at community, health facility, and health system levels. The Multi-Country Evaluation of IMCI Effectiveness, Cost and Impact (IMCI-MCE) is a global evaluation to determine the impact of IMCI on health outcomes and its cost-effectiveness. The MCE is coordinated by the Department of Child and Adolescent Health and Development of the World Health Organization. MCE studies are under way in Bangladesh, Brazil, Peru, Tanzania and Uganda.

In Tanzania, the IMCI-MCE study uses an observational design comparing four neighbouring districts, two of which have been implementing IMCI in conjunction with evidence-based planning and expenditure mapping at district level since 1997, and two of which are beginning IMCI implementation in 2002. In these four districts, child health and child survival are documented at household level through cross-sectional, before-and-after surveys and through longitudinal demographic surveillance respectively.

Here we present results of a survey conducted in August 2000 in stratified random samples of government health facilities to compare the quality of case-management and health systems support in IMCI and comparison districts. The results indicate that children in IMCI districts received better care than children in comparison districts: their health problems were more thoroughly assessed, they were more likely to be diagnosed and treated correctly as determined through a gold-standard re-examination, and the caretakers of the children were more likely to receive appropriate counselling and reported higher levels of knowledge about how to care for their sick

children. There were few differences between IMCI and comparison districts in the level of health system supports for child health services at facility level.

This study suggests that IMCI, in the presence of a decentralised health system with practical health system planning tools, is feasible for implementation in resource-poor countries and can lead to rapid gains in the quality of case-management. IMCI is therefore likely to lead to rapid gains in child survival, health and development if adequate coverage levels can be achieved and maintained.

4. 1 Introduction

More than 10 million children continue to die in the world every year.[91] Just 42 developing countries account for 95% of these deaths. Inequities in child survival between the rich and poor countries are unacceptably wide, as are the differences in survival between richer and poorer children within most countries.[76] Although child survival globally has increased over the past 30 years, the rate of increase peaked around 1980 and has virtually stagnated in sub-Saharan Africa. In some countries, child survival has even declined, and HIV/AIDS is only partly to blame.[92] The main causes of child death in the world are neonatal disorders, diarrhoea, pneumonia and malaria, although HIV/AIDS accounts for at least 10% of child deaths in some African countries.[91] Undernutrition is a major underlying cause, and has been estimated to contribute to over half of all child deaths.[91]

Effective interventions are available that could prevent or treat over 60% of all child deaths.[93] Yet mothers and children are not receiving these interventions: coverage levels remain unacceptably low.[94] The Integrated Management of Childhood Illness (IMCI) is a strategy for improving child health and development through the combined delivery of essential child health interventions. Originally, IMCI consisted of case-management guidelines for sick children in peripheral first-level health facilities, to be adapted for each country.[68] Later, the strategy expanded to include guidelines for delivering child survival interventions at household, community and referral levels, with three components: (1) improvements in case-management, (2) improvements in health systems, and (3) improvements in family and community practices. By the end of 2003, the first two components of IMCI were in the early implementation or expansion phase in 108 developing countries, including virtually all African countries south of the Sahara.[95] The Multi-Country Evaluation of IMCI (IMCI-MCE) seeks to generate information on the effectiveness, cost and impact of IMCI that can be used to strengthen the delivery of child health interventions and the implementation of the IMCI

strategy, and includes in-depth studies in Bangladesh, Brazil, Peru, Tanzania and Uganda.[96]

Tanzania has an under-five mortality rate of 147 per 1000 births resulting in almost a quarter of a million child deaths each year.[97] The country has a Gross Domestic Product of US$501 per capita[98] and is undergoing health sector and local government reforms. As districts gain more control over their health budgets, IMCI is one of the strategies recommended by the Ministry of Health to address major child health problems, such as malaria, pneumonia, malnutrition and diarrhoea, which together account for over 83% of post-perinatal under-five deaths.[99] Here we report the effectiveness of facility-based IMCI, by which we mean the first two components of the strategy, on child health and survival in rural Tanzania. We compared child health, household-level child health behaviours, and child survival in two districts with facility-based IMCI and two neighbouring comparison districts without IMCI over the time period from 1997 to 2002.

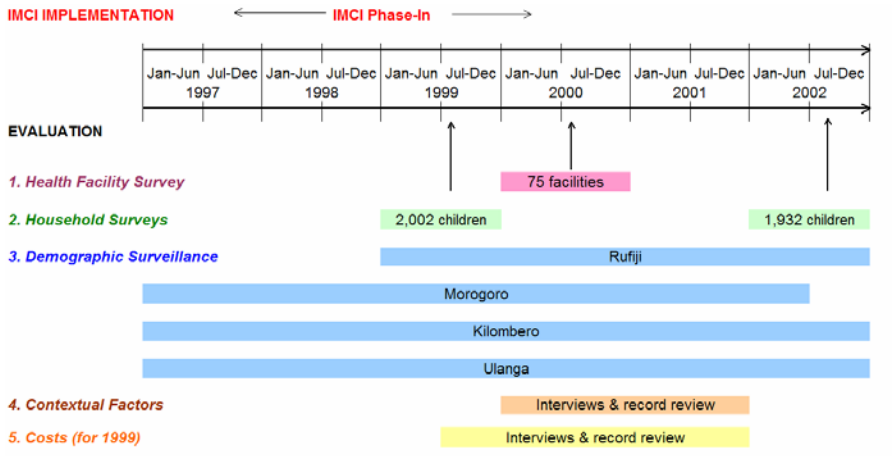
4. 2 Methods

4.2. 1 Design

We used a non-randomised controlled trial, or "plausibility" design,[100] to compare child health and survival in four neighbouring rural districts of Morogoro and Coast Regions, southern Tanzania, in 1999 and 2002. The two 'IMCI districts', Morogoro Rural and Rufiji, started to implement IMCI in 1997-8, while the two 'comparison districts', Kilombero and Ulanga, started implementation in 2002. IMCI and comparison districts are separated by a large uninhabited game reserve, making population movement between intervention and comparison areas minimal. Figure 4. 1 shows the timing of implementation of IMCI and each of the five main study components. Firstly, we assessed the quality of case-management for child illness, drug and vaccine availability, and supervision involving case-management, through a cross-sectional survey in a sample of health facilities from all four districts in August 2000: detailed methods and findings are reported elsewhere, with a summary given here for completeness.[73] Secondly, household surveys were used to assess child health indicators in July-August 1999 in a probability sample of children from all four districts,[101] early in the implementation phase, and again three years later in July-August 2002. Thirdly, child survival in a part of each district was tracked through demographic surveillance[102] throughout the study, with particular emphasis on a pre-defined two-year period from mid-2000, by which time IMCI implementation was thought to have reached high enough coverage for a long enough period for an effect on child survival to be

measurable. Fourth, information on contextual factors – programs and issues other than IMCI that might have affected child health in the four districts over the study period – was gathered through interviews with all health actors in the study districts and desk review of plans, budgets and reports, together with data from the child health surveys. Fifthly, we estimated the economic cost of child health care in IMCI and comparison districts through interviews and record review at national, district, facility and household levels: detailed methods and findings are reported elsewhere, and summarized here for completeness.[77;84]

Figure 4. 1 Timing of implementation of IMCI and the different components of this study



4.2. 2 *Study setting*

Kilombero, Morogoro Rural, Rufiji and Ulanga Districts are in southern Tanzania and have a total population of about 1.2 million people, of whom 200,000 are children under five years. (http://www.tanzania.go.tz/census. Accessed November 16, 2003). Kilombero and Rufiji are low-lying and much of the land is in the fertile flood plain of the Kilombero and Rufiji rivers: Morogoro Rural and Ulanga have mountainous areas as well as low-lying plains. There are two main rainy seasons, October-December and February-May. There is a broad mix of ethnic groups although Swahili, the national language, is widely spoken. The majority of people are subsistence farmers. Most dwellings have wood-framed mud walls with thatched or corrugated

roofs. Most rural roads are unpaved and transport can be difficult in the rainy season. The public health system has a network of hospitals, health centres and dispensaries, with 3,300 to 7,000 people served by each facility. Over 70% of the population live within 5km of a health facility. Utilisation of health facilities is relatively high, with routine Health Management Information System reports suggest 3.0 under-five visits per child per year for curative care in 1999. Malaria, pneumonia and waterborne diseases such as cholera and diarrhoea are the major health problems of the area as reported through the health services and as perceived by local people. For Tanzania as a whole, per capita expenditure on health was USD 11.37 in 1999-2000, including private, out-of-pocket expenses.[103] Monthly total per capita household consumption and expenditure in 2001 was around $10, of which around 70% was for food.[104]

4.2. 3 IMCI implementation

Implementation of IMCI in Tanzania is described elsewhere.[73] Briefly, the Tanzania Ministry of Health began IMCI implementation in 1996, and adapted generic IMCI case-management guidelines to reflect national child health policies (e.g., first- and second-line treatments for malaria and pneumonia) and local terms for illness symptoms and providers. All materials were translated into Swahili and used as the basis for preparing national and district level trainers. The target audience for the 11-day training was all health workers in first-level health facilities who provide case-management to children. The majority of these health workers have a 2-3 year training in clinical medicine following primary education: around a quarter have public health training following primary education. Through local government and health sector reforms, local councils have increased autonomy and control over their own health budgets and plans, and they have access to a limited amount of donor-supported "basket" funding from the health Sector-Wide Approach (SWAp)[105] (The term "council" refers to the local government of both rural districts and urban municipalities). The Council Health Management Team (CHMT) of Morogoro Rural and Rufiji Disticts decided to adopt IMCI, and to give highest priority to its introduction and implementation, based on evidence available to them from a sentinel burden-of-disease information tool and a district health budget mapping tool developed by the Tanzania Essential Health Interventions Project (TEHIP).[106] In addition TEHIP provided financial resources to districts of approx $0.92 per capita per year to simulate sector-wide "basket" funding[105] three years in advance of the actual start of "basket" funding. CHMTs of Morogoro Rural and Rufiji reported that over 80% of health workers managing children in first-

level facilities had been trained in IMCI by mid-2000 based on an 11-day course with approximately 30% of the training time spent in clinical practice.

4.2. 4 Health Facility Survey

Detailed methods are given elsewhere.[73] A representative sample of 75 health facilities was selected from government dispensaries and health-centres. Within chosen health facilities, the first 6 sick children aged 2 to 59 months attending on the day of the survey for an initial visit for any illness, and whose mother consented, were eligible for inclusion. Through observation of case-management, exit interviews with caretakers, re-examination and interviews with health care providers we collected information on 29 indicators relating to assessment, classification and treatment of the child, counselling and communication with the mother, and health systems support.

4.2. 5 Household surveys

Detailed information on the 1999 survey is given elsewhere.[101] Briefly, a representative cluster sample of approximately 2,300 rural households was taken from the four districts in July-August 1999. Thirty rural clusters, each of 20 households, were chosen from three of the four districts and 25 rural clusters of 20 households were chosen from the fourth district, Kilombero, using a modified EPI-type scheme[107] that ensured an equal probability of selection for every household. A modular questionnaire about the health of all children under five years was administered to consenting household heads, generating information on household-level child health indicators as agreed by an inter-agency working group on IMCI including representatives from WHO, UNICEF, the United States Agency for International Development (USAID), the US Centers for Disease Control, and the USAID-funded BASICS Project. Information on proxy markers of household socio-economic status was collected, such as household ownership of a radio, a tin roof, a bicycle, and the education and occupation of the household head. Mothers or carers (here we use the term 'mother' to denote the main carer) of all children under five years were then interviewed about their educational level, whether or not the child was currently breastfed and if so what other food or drink the child had received over the previous 24 hours. Information on routine vaccinations was documented either directly from health cards or, where no health card or other written record was available, according to the mothers' recall. Mothers were asked whether the child had received vitamin A supplementation, and if so, how many months ago. Mothers were then asked about any illness each child had during the 2 weeks prior to the survey, and what action had been taken. For children who had been sick, further modules elicited detailed

information about utilisation of appropriate (non-traditional) health care providers including village health workers, dispensaries, health centres, hospitals, or private doctors. Information on the care the child received at each such provider, and any other treatments the child had taken, was collected. Special attention was given to the care of children with danger signs: those with fast or difficult breathing, fits or convulsions, very sleepy, vomiting everything, or unable to drink/breastfeed.[73] Children were invited to attend a measuring station set up in the middle of the village where they were weighed on digital scales (Seca Vogel & Halke GmbH & Co, Hamburg, Germany) and their height (≥2 years) or length (<2 years old) was measured using purpose-made instruments. A generic version of the questionnaire is available from the authors on request.

In July-August 2002, a similar follow-up survey was done. Households were selected from the same villages (clusters) as in 1999. Within selected villages, the chance of visiting the same household was small: a single sub-village (*kitongoji*) was chosen at random and 20 households selected. Care was taken to ensure that no survey staff visited a village that they had worked in during the earlier survey. The questionnaire was translated to Swahili, back translated, pre-tested, and pilot-tested. Quality control measures in each cluster included supervisors accompanying 1-3 interviews; households reported to be empty were visited by a supervisor; up to 2 mothers bringing their children to be weighed and measured were re-interviewed by a supervisor, with information from the two interviews compared and discrepancies discussed and resolved with the original interviewer; and weight and height measurements were repeated for up to 2 children.

4.2. 6 Analysis of household survey data

Weight for age, height for age and weight for height Z-scores were calculated with reference to the US National Centers for Health Statistics (NCHS) standards using the EPINUT module of EPI-Info v6.0 (CDC Atlanta, Georgia, US). Underweight, stunting and wasting were defined respectively as weight-for-age, height-for-age and weight-for-height z-scores of less than –2, excluding outliers (z-score of <-5 or>3). Because stunting is most prevalent at ages 24-59 months, and wasting at 12-23 months,[108] analyses were carried out for these subgroups. Data was processed in FoxPro (v2.6, Microsoft Corporation, Seattle, WA, USA) and analysed in Stata,[86] following an analytical plan which had previously been agreed by the investigators. All analyses were adjusted for clustering was made using standard STATA commands such as *svymean* and *svylogit*. P-values should be interpreted with caution given the non-randomised study design and the large number of tests. We did significance testing to look for evidence of four types of differences: (1)

between IMCI and comparison districts in each of the two surveys; (2) between 1999 and 2002 for all 4 districts combined; (3) between IMCI and comparison areas, combining 1999 and 2002 data; and (4) whether the difference between IMCI and comparison districts changed on average between 1999 and 2002, ie. for an interaction between IMCI and time. The latter effects are potentially IMCI-attributable, provided there were no contextual factors that explained them.

4.2. 7 Demographic surveillance

As routine registration of births and deaths in Tanzania is uncommon, we used demographic surveillance systems (DSS) to measure child survival in a part of each of the four districts. These systems are described in detail elsewhere[102] a brief overview is given here.

Kilombero and Ulanga districts. The DSS started in 1996 with a baseline census in a population of 52,000, covering 6 contiguous villages of Kilombero district (10% of the district population) and 12 contiguous villages of Ulanga (16%). These areas are not a representative sample of either district. Since January 1997, an interviewer – one of about 40 full-time staff – has visited each household every four months and collected information on pregnancies, births, deaths and migrations, using the household registration system (HRS).[109] Births and deaths are also reported on a continuous basis by key informants based in each *kitongoji*. Extensive quality control measures include repeat interviews in a randomly selected 10% of all households. Data from each week's work is used to update the HRS database before a weekly field meeting. Checking programs are run and queries referred back to the field team for correction within 2 weeks of the original interview.

Rufiji district. Field procedures, quality control and data management are similar to those in Kilombero and Ulanga DSS (above). Rufiji DSS started in 1998 with a baseline census in a population of about 70,000 people covering 32 villages (44% of the district population).

Morogoro Rural district. Demographic and mortality surveillance began in 1992 with an initial census, which has since been repeated annually in about 85,000 people in 50 villages (16% of the district population). Continuous mortality surveillance provides information on numbers and probable causes of death using 'verbal autopsies'. About 70 villagers act as enumerators for the annual census update, and as key informants for reporting deaths, each of which is followed up by one of 4 clinical officers from the CHMT. Data entry uses a tailor-made FoxPro database system. Data quality is assured by checks in the

field, during and after data entry. Supervisors visit a random sample of households to verify entries on the census forms, to check that all households visited have been included in the census, and that no non-existent households have been included. Following each census re-interviews of a sample of households for each enumerator are conducted.

4.2. 8 Analysis of mortality data

Our primary focus was to compare mortality over the two-year period starting in mid-2000, by which point IMCI implementation was thought to have reached sufficient sustained coverage for any effect on child survival to be measurable, as judged by an independent review panel. Following an agreed analytical plan, we compared under-five mortality rates per 1000 children per year between IMCI and comparison areas from mid-2000 to mid-2002, checking for any differences in 1999, which served as a baseline for the mortality analysis. Adjustments for age (0-1 and 1-4 years) and rainfall (estimated from remote sensing data) were made using Poisson regression models, and the between-district differences compared using t-test-based methods of adjusted residuals, as appropriate for clustered data with a small number of clusters.[110] With only four districts, p-values from this approach are likely to be conservative and we also calculated p-values from Poisson regression ignoring between-district variation. Secondary analysis made use of all available DSS data to summarise longer-term trends in child survival in relation to IMCI. We used Poisson regression for the data from each DSS area separately, testing for the size and statistical significance of the trend in mortality rates over time. Due to data completeness problems, analysis was repeated with and without Morogoro data for the year 2000.

4.2. 9 Contextual factors

We summarised factors other than IMCI that may have affected child health in the four districts, with emphasis on those that might have changed over the study period, including geographic, environmental and demographic features, health care infrastucture and activities, and other health-related programs, activities or events (including disasters, famines, etc). These factors were related to quantifiable indicators from the household surveys. In addition to using data from the household surveys, we contacted around 40 health actors in the four study districts, including CHMT's, TEHIP, non-governmental organisations, religious Missions, bilateral and multilateral aid organisations. Information on routine health care delivery and other relevant activities was

systematically collated from written reports and interviews with key informants.

4.2. 10 Economic costs

Detailed methods are given elsewhere.[77;84] Briefly, cost data were collected for the start-up period of implementing IMCI (1996 to 1997) and for maintaining under-five health services including IMCI subsequently. Costs were estimated from the societal perspective and were collected from the national, district, hospital, health facility and household levels. Among those included were drugs and vaccines; training costs attributable to under-five care, including but not restricted to IMCI; the annualized cost of capital items; and the opportunity cost of staff time spent in consultation with under-fives (assessed through a time-and-motion study) and time spent attending meetings and performing supervision visits. Household costs included travel and out-of-pocket expenditures to obtain care for under-fives but did not include a monetary value of time lost in seeking care. Costs at all these levels were summed to obtain the total cost to the district of providing care for under-fives. To allow comparison across districts, cost estimates were standardized to a hypothetical district with a population of 50,000 under-fives. Estimates of the additional cost to the district of implementing IMCI were based on the difference in cost of under-five care between the standardized IMCI and comparison districts.

4. 3 Results

4.3. 1 Quality of care measured through the health facility survey

The introduction of facility-based IMCI was associated with improved quality of care as measured through the health facility survey in August 2000, after the end of the IMCI phase-in (Figure 4. 1). Nearly all children observed in the IMCI facilities were checked for cough, diarrhoea and fever (95% of 231), compared with 36% of 188 children in the comparison districts ($p<0.001$; Table 4. 1).[73] In IMCI districts, 63% of sick children were correctly classified compared with 38% in the comparison districts ($p<0.0001$). Drug availability was reasonably good and comparable in IMCI and comparison areas at the time of the survey (Table 4. 1, $p=0.47$). However, more than twice as many sick children needing oral antibiotics and/or oral antimalarials were prescribed them correctly in the IMCI districts, compared with the comparison districts (73% vs 35%; $p<0.001$). Over 70% of caretakers in the IMCI districts who had

received a prescription for Oral Rehydration Salts (ORS), oral antibiotic or an oral antimalarial, reported correctly at the facility exit how to give the treatment, compared to 56% of those in comparison districts (p=0.02). Supervisory visits that included observation of case-management were more common in IMCI districts than in comparison districts, with over half of IMCI facilities having had such a visit in the 6 months before the survey, compared with 21% of comparison facilities (p=0.007). Follow-up visits after IMCI training were not included in these calculations. IMCI facilities were significantly better than, or comparable to, comparison facilities for all but one indicator out of a total of 29.

Table 4. 1 Selected indicators showing quality of care and health systems support in IMCI and comparison districts in August 2000[73]

	IMCI	**Comparison**	**P-value***
	Morogoro & Rufiji %	**Ulanga & Kilombero**%	
Child checked for the presence of cough, diarrhoea, and fever	95% n=231	36% n=188	<0.001
Child correctly classified	63% n=219	38% n=176	<0.001
Child needing an oral antibiotic and/or an oral antimalarial is prescribed correctly	73% n=219	35% n=178	<0.001
Caretaker of child who received a prescription for an oral medication** reports correctly at facility exit how to give the treatment	72% n=225	56% n=179	0.02
Index of availability of essential oral treatments (mean)	0.93 n=39	0.95 n=35	0.47
Health facility received at least one supervisory visit that included observation of -case management during the previous six months (%)	51% n=37	21% n=34	0.007

* F-tests comparing areas with and without IMCI, unless otherwise stated
**ORS and/or oral antibiotic and/or antimalarial

4.3. 2 Household surveys 1999 and 2002

In 1999, data was available from 2,006 children under 5 years of age living in 1,321 households with children under-five in the 120 rural clusters,

representing 93% of eligible households.[101] In the 2002 survey, 1,932 children were identified living in 1,341 households with children under-five within the same villages, representing 94% of eligible households (residents of 142 (6%) households were away and 9 (0.4%) households refused to take part). The age profile of children involved in the two surveys was similar, with 24% and 21% of children being under one year of age in 1999 and 2002 respectively.

Appropriate care-seeking, home management, caretaker knowledge, nutrition, and feeding: Just over one-third of all children were reported by their carers as having been ill in the two weeks preceding the survey in 1999 and 2002 respectively (36%, 717/2,006 in 1999 and 35%, 678/1,932 in 2002). Of these, 42% sought care from an appropriate health care provider in 1999, compared with 36% in 2002 (Table 4. 2, p=0.02). There was no evidence of any difference in care-seeking between IMCI and comparison districts (p=0.45), and neither was there any evidence of a differential change in care-seeking between IMCI and comparison districts over time (p=0.36). Care-seeking was generally more common for children reported as having had danger signs: in 1999, 53% (86/162) of such children had reportedly been taken to an appropriate provider in IMCI districts, and 68% (100/147) in comparison districts. In the 2002 survey, care-seeking for children with danger signs had risen slightly in the IMCI districts to an average of 55% (78/142) and had dropped in comparison districts to 43% (49/113) (p=0.006 for the differential change over time). Over 10% (200/1971) of children had been admitted to a health facility in the year before the 1999 survey, with this being more common in the comparison districts than in the IMCI districts (14% 139/976 and 6% 61/995, p=<0.001). Although this difference persisted in the 2002 survey, it was much less marked as hospital admissions had risen by 1% on average in the IMCI districts, and dropped by 3% in the comparison districts (p=0.002 for the differential change between IMCI and comparison areas).

Among children sick with diarrhoea in the 2 weeks before the survey, the use of oral rehydration solution (ORS) showed large variations between the districts, from 6% to 41%. Despite this, ORS was consistently more common in IMCI than comparison areas in both surveys (p=0.03 in 1999 and p=0.01 in 2002).

Appropriate home management of disease, measured in part by the proportion of children sick on the day of the survey who had received increased fluids and continued feeding, was under 10% in all districts and both surveys. Nevertheless there was evidence of differential improvement in the IMCI districts compared to comparison areas (p=0.05), with IMCI districts increasing by 4% on average between 1999 and 2002, and the comparison areas decreasing by 1%.

In 1999, more than one-fifth of carers knew at least two signs for seeking care immediately, and this knowledge increased in all districts by 2002 to around one-third of carers ($p<0.001$ for the average change over time). There was no evidence of a differential change in such knowledge between IMCI and comparison districts ($p=0.93$).

Anthropometric indicators of nutritional status show that underweight is common in the study area, with 28% of all children affected in 1999 (526/1852). The prevalence of underweight had dropped in all districts by 2002, to 22% (394/1826; $p<0.0001$). Underweight affected 4 - 5% more children in the IMCI districts than in the comparison districts in both years ($p=0.03$). Wasting was found in 11% of children aged 12-23 months in 1999 (42/367) and dropped to 6% (23/368) children in 2002 ($p=0.02$), with no evidence of differences between IMCI and comparison areas ($p=0.43$), nor a differential change over time ($p=0.79$). Stunting affected 60% (297/497) of children aged 24-59 months in IMCI districts in 1999, about 10 percentage points more than in comparison areas (51%; 247/483; $p=0.04$). By 2002, however, children in IMCI areas had "caught up", and stunting was comparable in IMCI and comparison areas (43% (191/477) vs 40% (249/585); $p=0.07$ for the differential change over time). When expressed as a mean height-for-age z-score in children aged 24-59 months, the differential change between IMCI and comparison districts reached conventional statistical significance ($p=0.05$, Table 4. 2).

Exclusive breastfeeding in children under 4 months of age was practiced by 23% (41/180) of mothers in 1999, rising to 27% (22/119) by 2002 ($p<0.0001$). For children aged 6-9 months, 93% (141/152) received breast milk and complementary feeding in 1999, rising to 99% (154/156) by 2002 ($p=0.008$). Neither of these breastfeeding indicators showed evidence of a difference between IMCI and comparison districts ($p=0.69$, $p=0.23$), nor of a differential change over time ($p=0.89$, $p=0.26$). Breastfed one-year-old children in IMCI districts received an average of 2.7 meals each day in 1999, compared to 3.0 in comparison areas ($p=0.05$). By 2002, these differences were no longer evident, and children in IMCI districts were receiving an average of 3.1 meals per day, similar to the average number of 3.2 meals given to children in comparison areas ($p=0.03$ for the differential change over time).

Table 4. 2 Indicators of feeding, nutrition, caretaker knowledge, home management of disease, utilisation and care in 1999 and 2002

INDICATOR	YEAR	DISTRICT				P-values			
		Morogoro Rural (IMCI)	Rufiji (IMCI)	Ulanga (No IMCI)	Kilombe-ro (No IMCI)	IMCI vs compa-rison areas	Change over time 1999 to 2002	Difference between IMCI and comparison	Difference between IMCI and comparison changed over time
4.2.1. UTILISATION & CARE									
CS28. Appropriate care-seeking	1999	36% (n=239)	46% (n=273)	38% (n=273)	47% (n=229)	0.81	0.02	0.45	0.36
	2002	32% (n=229)	45% (n=302)	35% (n=276)	31% (n=151)	0.17			
AI21. Care-seeking rate for children with danger signs	1999	51% (n=78)	55% (n=84)	62% (n=74)	74% (n=73)	0.02	0.05	0.54	0.006
	2002	48% (n=64)	60% (n=78)	46% (n=71)	38% (n=42)	0.15			
*AI19. Proportion of children admitted in last year**	1999	7% (n=450)	5% (n=545)	12% (n=570)	17% (n=406)	<0.001	0.45	0.002	0.002
	2002	9% (n=503)	6% (n=547)	11% (n=504)	13% (n=360)	0.06			
AI33. ORS use among children with diarrhoea	1999	13% (n=45)	38% (n=40)	15% (n=48)	6% (n=47)	0.03	0.31	0.08	0.07
	2002	14% (n=42)	41% (n=54)	27% (n=55)	8% (n=26)	0.01			

INDICATOR	YEAR	DISTRICT				P-values			
		Morogoro Rural (IMCI)	Rufiji (IMCI)	Ulanga (No IMCI)	Kilombe-ro (No IMCI)	IMCI vs compa-rison areas	Change over time 1999 to 2002	Difference between IMCI and comparison	Difference between IMCI and comparison changed over time
4.2.2. HOME MANAGEMENT OF DISEASE									
CP15. Sick child (today) receives increased fluids and continued feeding	1999	1% (n=136)	5% (n=125)	9% (n=140)	6% (n=107)	0.02	0.31	0.23	0.05
	2002	8% (n=133)	5% (n=117)	8% (n=149)	4% (n=76)	0.89			
4.2.3. CARETAKER KNOWLEDGE									
CP17. Caretaker knows at least two signs for seeking care immediately	1999	24% (n=316)	19% (n=357)	21% (n=385)	26% (n=286)	0.54	<0.001	0.11	0.93
	2002	33% (n=384)	27% (n=400)	32% (n=379)	38% (n=284)	0.16			
4.2.4. FEEDING & NUTRITION									
CP4. Low weight for age	1999	29% (n=429)	32% (n=518)	28% (n=559)	23% (n=344)	0.10	<0.0001	0.03	0.45
	2002	23% (n=483)	24% (n=518)	22% (n=494)	16% (n=331)	0.07			
CP5. Mean weight for age z-score	1999	-1.29 (n=429)	-1.32 (n=518)	-1.29 n=559)	-1.13 n=344)	0.38	<0.0001	0.04	0.92
	2002	-1.21 (n=483)	-1.22 (n=518)	-1.13 n=494)	-0.91 n=331)	0.007			
CS2-JP. Stunting prevalence in children aged	1999	61% (n=233)	59% (n=264)	54% (n=303)	46% (n=180)	0.04	<0.0001	0.07	0.07

INDICATOR	YEAR	DISTRICT				P-values			
		Morogoro Rural (IMCI)	Rufiji (IMCI)	Ulanga (No IMCI)	Kilombe-ro (No IMCI)	IMCI vs compa-rison areas	Change over time 1999 to 2002	Difference between IMCI and comparison	Difference between IMCI and comparison changed over time
24-59 months	2002	42% (n=289)	44% (n=296)	45% (n=282)	32% (n=195	0.45			
CS4-JP. Mean height for age z- score for children aged 24-59 months	1999	-2.25 (n=233)	-2.27 (n=264)	-2.07 n=303)	-1.87 n=180)		<0.0001	0.02	0.05
	2002	-1.83 (n=289)	-1.84 (n=296)	-1.92 n=282)	-1.47 n=195)				
CS3-JP. Wasting prevalence in children aged 12-23 months	1999	17% (n=72)	9% (n=95	9% (n=115)	13% (n=85)	0.52	0.02	0.43	0.79
	2002	5% (n=91)	8% (n=100	6% (n=110)	4% (n=67)	0.66			
CP1. Child under 4 months of age is exclusively breastfed	1999	25% (n=52)	17% (n=59)	26% (n=42)	26% (n=27)	0.46	<0.0001	0.69	0.89
	2002	23% (n=30)	22% (n=36)	30% (n=30)	35% (n=23)	0.34			
CP2. Child aged 6-9m receives breast milk & complementary feeding	1999	83% (n=30)	94% (n=49	98% (n=41	94% (n=32	0.17	0.008	0.23	0.26
	2002	97% (n=38)	100% (n=43	97% (n=39)	100% (n=36)	0.96			
AI3. Mean meal frequency for breastfed 1-year-olds	1999	2.6 (n=67)	2.8 (n=87	2.8 (n=98)	3.2 (n=82)	0.05	0.01	0.06	0.03
	2002	2.9 (n=81)	3.3 (n=98	3.1 (n=99)	3.2 (n=57)	0.61			

4.3. 3 Survival

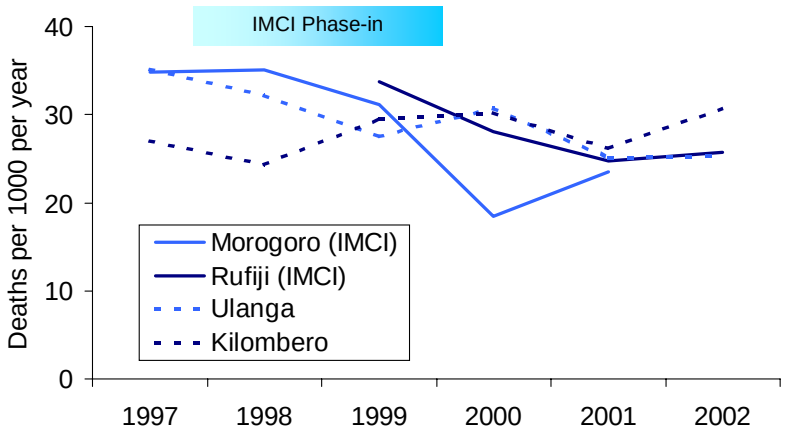
Mortality levels in children under five years, from the area of each district under demographic surveillance, are shown in Table 4. 3. In the period from July 1999 to June 2000, during the phase-in of IMCI, under-five mortality was 27.2 per 1000 per year and 27.0 per 1000 per year in IMCI and comparison districts respectively, giving a rate ratio of 1.01 and a rate difference of 0.2 deaths per 1000 children per year. Over the two following years, from July 2000 to June 2002, mortality levels were 13% lower in the IMCI districts than in the comparison districts (rate ratio = 0.87), corresponding to a rate difference of 3.8 fewer deaths per 1000 children per year associated with IMCI. Adjusting for the difference between the areas in 1999, this corresponds to an adjusted rate ratio of 0.86, almost identical to the unadjusted value. Adjustments for age (0-1 and 1-4 years) and estimated rainfall were made using Poisson regression models, but had no effect on the estimated rate ratio (data not shown). Ignoring between-district variation, the confidence interval for the 13% reduction in mortality associated with IMCI was (5%, 21%)(p=0.004 from likelihood-ratio χ^2 -test). Allowing for variation between districts and using a Normal approximation based on the log rate ratio, the confidence interval was (-7, 30) (p=0.28 using a t-test).

Table 4. 3 Mortality levels from July 2000-June 2002.

Time period	District	Deaths	Child-years	Death Rate (/1000 child-yrs)	Rate Ratio	Rate Difference (/1000/yr)
July 1999-June 2000	Morogoro	252	11,303	22.3		
	Rufiji	387	12,212	31.7		
	IMCI districts	*639*	*23,516*	*27.2*	*1.01*	*0.2*
	Kilombero	146	4,687	31.1		
	Ulanga	96	4,289	22.4		
	Comparison districts	*242*	*8,977*	*27.0*		
July 2000-June 2002	Morogoro	522	23,985	21.8		
	Rufiji	698	25,979	26.9		
	IMCI districts	*1,220*	*49,964*	*24.4*	*0.87*	*3.8*
	Kilombero	362	12,685	28.5		
	Ulanga	257	9,280	27.7		
	Comparison districts	*619*	*21,965*	*28.2*		

Additional mortality data were available from some of the DSS sites from January 1997 until December 2002 (Figure 4. 2). We used all available data from complete years to assess trends over time in each district. This analysis found no evidence of a change over time in under-five mortality in one of the two comparison districts, Kilombero, over the six years from January 1997 to December 2002, where there was an estimated 2% annual increase in under-five mortality rate, 95% CI (-2%, 6%), p=0.23). In the second comparison district, Ulanga, there was an annual decrease in under-five mortality of 6% (95% CI 2,10) , p=0.002) (Figure 4. 2). In districts with IMCI, data was less complete, but analysis of all available data for whole calendar years showed an annual drop of 11% in mortality in Rufiji district between 1999 (February) until December 2002 (95% CI 7, 16; p<0.0001), and an annual drop of 14% from January 1997 until December 2001 in Morogoro Rural district (95% CI 11, 17, p<0.0001). Due to data completeness problems, the Morogoro Rural analysis was repeated without data from the year 2000, and found an 11% annual drop in under-five mortality (95% CI 8, 15, p<0.0001).

Figure 4. 2 All-cause mortality in children under five years of age, from 1997-2002.



4.3. 4 Contextual factors

The desk reviews of plans, budgets, reports, and expenditure, and interviews with health actors suggested a number of factors other than IMCI relevant to child health which differed between IMCI and comparison districts. We were particularly concerned to identify those factors which might account for the observed differential changes in survival between IMCI and comparison districts, ie. factors which changed rapidly, and differentially, over the study period. The factors identified were limited to malaria control efforts, vaccination programmes, and vitamin A supplementation coverage.

For children who had fever in the two weeks before the 1999 household survey, 42% (300/717) had been given an antimalarial drug (Table 4. 5). In 2002, only 28% of such children had received an antimalarial drug ($p<0.0001$), with no evidence of differential change in IMCI and comparison districts ($p=0.35$). In 1999, the majority of these children (85%) received chloroquine, whereas in 2002 the most commonly used drugs were SP (53%), quinine (24%) and amodiaquine (17%). A second malaria control effort that differed between the districts was that of social marketing of treated mosquito nets, and we were also concerned that coverage of untreated nets purchased through the private sector might also have changed differentially. Use of untreated nets the night before the 1999 household survey varied markedly between districts, from 14% (64/449) in Morogoro rural district (IMCI) to 59% (236/399) in Kilombero district (without IMCI). Coverage increased dramatically in all four districts by 2002, reaching 85% (299/352) in Kilombero and 35% (175/500) in Morogoro rural. Although the coverage increase was larger in IMCI than comparison districts when measured as a ratio of 2002 / 1999 coverage (1.95 compared with 1.57; $p=0.03$), coverage remained significantly higher in comparison districts than in IMCI districts in 2002 (66% (567/856) on average in comparison districts compared with 41% (429/1047) in IMCI districts; $p<0.0001$). Use of nets that had been treated in the 6 months before the survey was uncommon in all districts in 1999, at less than 7%, and increased in all districts by 2002, by 4 to 14 percentage points, with no evidence of a differential change between IMCI and comparison areas over time ($p=0.14$). Although vaccine coverage was generally high (over 80%) in all districts and during both surveys, there was evidence of a small drop in coverage of BCG vaccine between the two surveys, from 97% to 96% ($p=0.02$). We also found a differential decline in coverage of DPT vaccine from the 1999 level of 86% (342/397). In 2002, DPT coverage increased to 95% (176/185) in the comparison districts but had dropped to 82% in the IMCI districts ($p=0.03$ for the differential change over time). Vitamin A coverage was only 14% in 1999, and comparable in IMCI and comparison areas. In 2002 coverage had dramatically increased in all districts, to an average of 76% ($p<0.0001$), with no

apparent difference between the districts (p=0.73) nor any evidence of a differential change over time (p=0.98).

For vaccination coverage and malaria control efforts, the higher coverage in comparison districts would have tended to negate any apparent effect of IMCI on child survival. For Vitamin A, levels were similarly low in all four districts in 1999 and rose equally by 2002, and cannot therefore account for the greater drop in mortality in the IMCI areas over the study period.

4.3. 5 Economic costs of under-five care

IMCI was not associated with higher economic costs than conventional care. In fact, the reverse was found: the cost of under-five care per child was estimated at $11.19 in IMCI districts and $16.09 in comparison districts (Table 4. 4).[77;84] The major components of cost were at district, hospital, primary facility and household levels, with the estimate of hospital-level costs in the comparison districts accounting for almost half of total cost of under-five care per child (46%). Since we would not have expected IMCI to have affected hospital-level costs to such an extent, and because access to hospitals was better in comparison than in IMCI districts, we re-calculated the cost of under-five care per child without the hospital component, and found similar total costs per child in IMCI and comparison areas ($8.30 in the IMCI districts, $8.76 in the comparison areas).

Table 4. 4 Cost of under-five care per child, in standardised districts of 50,000 under-fives, with and without IMCI. Costs are in Tanzanian shillings (US$ in brackets) (refer to Chapter 3)

Level	Standardised district with IMCI		%	Standardised district without IMCI		%
	Tsh	(US$)		Tsh	(US$)	
National	129	(0.17)	1	55	(0.07)	0
District	1,784	(2.30)	21	2,605	(3.35)	21
Hospital	2,243	(2.89)	26	5,692	(7.33)	46
Primary-facility	2,455	(3.16)	28	2,283	(2.94)	18
Household	2,083	(2.68)	24	1,867	(2.40)	15
Total	8,695	(11.19)	100	12,503	(16.09)	100
Total excluding hospital costs	6,452	(8.30)		6,810	(8.76)	

Table 4. 5 Indicators reflecting programs and issues other than IMCI that may have affected child health and survival over the study period.

INDICATOR	YEAR	DISTRICT				P-values			
		Morogoro Rural (IMCI)	Rufiji (IMCI)	Ulanga (No IMCI)	Kilombero (No IMCI)	IMCI vs compa-rison areas	Change over time 1999 to 2002	Difference between IMCI and comparison	Difference between IMCI and comparison changed over time
CP16. Child with fever receives antimalarial drug	1999	39% (n=174)	40% (n=199)	32% (n=180)	58% (n=164)	0.23	<0.0001	0.33	0.35
	2002	31% (n=177)	25% (n=217)	24% (n=178)	35% (n=107)	0.86			
AI36. Child sleeps under net	1999	14% (n=448)	27% (n=543)	31% (n=565)	59% (n=397)	<0.001	<0.0001	<0.0001	0.03
	2002	35% n=500)	46% (n=547)	53% (n=504)	85% (n=352)	<0.0001			
CP10. Child sleeps under treated net (treated in the last 6m)	1999	3% (n=448)	3% (n=543)	5% (n=565)	7% (n=397)	0.10	<0.0001	0.22	0.14
	2002	8% (n=500)	16% (n=547)	9% (n=504)	21% (n=352)	0.54			
CP6. Anemia in children >=6m (Hb<11.0 g/dL)	1999	88% (n=368)	93% (n=440)	87% (n=497)	88% (n=331)	0.23	<0.0001	0.07	0.67
	2002	81% (n=434)	86% (n=476)	84% (n=452)	69% (n=309)	0.06			
AI13. BCG vaccine coverage (informed or	1999	100% (n=80)	94% (n=100)	95% (n=118)	99% (n=98)	0.97	0.02	0.94	0.97

INDICATOR	YEAR	DISTRICT				P-values			
		Morogoro Rural (IMCI)	Rufiji (IMCI)	Ulanga (No IMCI)	Kilombero (No IMCI)	IMCI vs compa-rison areas	Change over time 1999 to 2002	Difference between IMCI and comparison	Difference between IMCI and comparison changed over time
registered)	2002	95% (n=91)	96% (n=104)	95% (n=108)	97% (n=69)	0.97			
AI11. DPT vaccine coverage (informed or registered)	1999	84% (n=80)	90% (n=100)	81% (n=118)	91% (n=98)	0.68	0.40	0.15	0.03
	2002	75% (n=72)	88% (n=95)	95% (n=108)	96% (n=68)	0.001			
AI12. Polio vaccine coverage (informed or registered)	1999	81% (n=80)	86% (n=100)	80% (n=118)	88% (n=98)	0.90	0.21	0.11	0.12
	2002	61% (n=57)	85% (n=92)	85% (n=97)	92% (n=65)	0.007			
AI10. Measles vaccine coverage (informed or registered)	1999	88% (n=80)	89% (n=100)	84% (n=118)	96% (n=98)	0.83	0.51	0.33	0.72
	2002	88% (n=84)	89% (n=96)	92% (n=105)	94% (n=67)	0.14			
CP14. Vitamin A supplementation coverage	1999	15% (n=79)	13% (n=99)	14% (n=111)	14% (n=96)	0.71	<0.0001	0.73	0.98
	2002	77% (n=73)	75% (n=81)	78% (n=89)	75% (n=53)	0.84			

4. 4 Discussion

We conducted an effectiveness evaluation to estimate the impact of a program that was selected and implemented by district health staff in rural Tanzania. We found evidence of improved case-management, a 13% lower under-five mortality rate in districts with IMCI than in comparison areas, and that costs of child health were comparable or lower with IMCI than with conventional case-management, suggesting that facility-based IMCI is highly cost-effective. Our findings on mortality levels and trends in the presence of facility-based IMCI, while not supporting the same type of inference as a randomised controlled trial covering a large number of districts, strongly support going to scale with this intervention in the context of health sector reform, basket funding, good facility access and high utilisation of health facilities.

Several design issues must be borne in mind when interpreting the findings of the present evaluation. We set out to do an effectiveness evaluation of an integrated delivery strategy encompassing a number of interventions – such as antibiotics, antimalarials, oral rehydration therapy, etc - whose efficacy is already well documented.[96] Our objective was to estimate the impact of a program that was managed by district health staff under routine circumstances, similar to a Phase IV study of a drug or vaccine. A randomised design would not have been possible because IMCI had already started to be implemented in two of the four districts that had demographic surveillance systems. Therefore, much emphasis was placed on documenting and assessing the effect of contextual factors that might confound the observed results. It is reassuring that the observed distribution of contextual factors would tend, if anything, to underestimate the true impact of IMCI, but in a non-randomized design with a small number of units of analysis it is not possible to rule them out completely. The study was designed to detect a 20% reduction in mortality, which seemed feasible given that IMCI includes several life-saving interventions, and that baseline mortality levels were high. We found a smaller impact than we had expected, and the confidence intervals for this estimate included the value of zero, or no impact. Nevertheless, in large scale public health evaluations the number of available units of analysis is often small, and even if the result had been statistically significant it would be hard to interpret on strict probabilistic terms. This is why we took so much effort in documenting intermediate changes that could be ascribed to IMCI implementation, and also contribute to reducing mortality, thus strengthening the plausibility[111] of a mortality impact. For example, health worker performance – including prescription of life saving drugs – improved markedly; key indicators that could be ascribed to IMCI implementation (such as home management of disease and some care-seeking and feeding practices) showed improvements, although most were still well below what could be

desirable; and mean height for age, an excellent indicator of overall child health that could be a result of improved disease management and feeding practices,[108] was significantly improved. Taken together, these results support the hypothesis that IMCI implementation led to a mortality reduction.

We identified several areas within facility-based IMCI that deserve careful attention, with a view to rethinking parts of this strategy. Results from Uganda suggest that supportive supervision is associated with improved quality of care.[112] We found that although general visits from supervisory teams were frequent, many such visits did not involve case-management. Supportive supervision involving case-management observation was more common in the districts with IMCI than in comparison areas, yet there was still room for improvement, with one-fifth of such staff not having had such a visit in the previous 6 months. Despite many attempts to develop an 'IMCI supervision checklist' this has proved impossible to implement due to the many duties that supervisors are expected to do. An integrated approach may help, and is under consideration in Tanzania. We also found that less than one-fifth of children needing referral were actually referred (17%, confidence interval 0,41, n=12),[73] despite an expectation that IMCI guidelines would lead to a massive increase in referrals.[113] More than 60% of first-level health facilities are over two hours travel time from their closest referral facility (data not shown), and a lack of transport and money together with the need to care for other children often means that mothers are unable to travel such distances with a sick child.[114] Health workers are part of the communities they serve, and it seems likely that they do not refer children who are unlikely to be taken to a referral facility. The IMCI guideline on "Where referral is not possible", which is part of the IMCI training course in Tanzania, goes some way towards supporting health workers to offer these children the best possible treatment.

The potential impact of IMCI is likely to depend largely on the efficacy and availability of appropriate treatment for malaria, which is the leading cause of morbidity and mortality in the study area. Nationally, the first-line antimalarial drug changed in 2000 from chloroquine to SP, due to widespread drug resistance. This change was followed by initial quality control and delivery problems, and these may have led to the drop in treatment of fever with antimalarials, from 42% to 28% in children sick in the two weeks before each survey. Despite increased efficacy, it appears that the newly introduced SP was neither so widely available nor so popular as chloroquine had been in 1999, and this is likely to have reduced the potential effectiveness of IMCI in the time-frame of this study.

From its origins as a case-management strategy, IMCI later developed into three linked components: case-management, health facility support, and household and community support. The third component, often known as

'community IMCI', consists of health education messages and programmes in support of 17 'key family practices', focussed around growth promotion and development, disease prevention, care-seeking and compliance to health workers advice. This package has not been implemented in our study area. Despite this, it should be noted that many of the activities with high coverage in all four districts are compatible with community IMCI, including the use of treated mosquito nets and vitamin A. Tanzania has a long history of community-based activities and programs: for example, all villages in Morogoro Region had two Village Health Workers trained in the 1980s, and in a few villages these people remain active in health promotion, including growth monitoring of children, supplemental feeding of those identified to be at risk, and village health days to promote vaccination. The potential gains from such outreach are clear, but there is currently no mechanism by which such community-based health workers can be motivated, supervised, and supported on a large scale. Given the difficulties in providing support and supervision to peripheral facility-based health workers, it remains unclear how a much larger force of unsalaried lower-level workers could be managed and sustained.

The finding that economic costs of IMCI were similar or less than costs of conventional child health care was unexpected. The explanation is that substantial amounts of money were spent on child health in comparison districts, and that this roughly equalled the cost of IMCI. There is no evidence to suggest that these districts are atypical of rural Tanzania: as is often the case, there was an ongoing donor investment in health programs including from multilateral agencies such as UNICEF, from bilateral agencies such as the Swiss Agency for Development and Co-operation, and from NGOs such as Plan International. Estimated funding for health from major donors such as SDC and Ireland Aid was around $0.8 per capita in the comparison districts, not very different from the TEHIP funding of $0.92 per capita. Furthermore, IMCI was implemented at a financial cost within the basket funding available to all districts through health sector reform. For example, the cost of training 48 health workers would use between 1% and 6% of current basket-funding allocated to each district. Implementation in the comparison districts started in 2002.

More than ten years ago, an early estimate of the likely cost-effectiveness of IMCI reported that "implementation of the integrated cluster of treatments, including hospital services, would cost between $30 and $100 per DALY saved".[87] Further, it was thought that if health services were well-used, child deaths might be reduced by 50 to 70 percent. With the benefit of hindsight, these expectations appear somewhat optimistic. Implementation of IMCI has proved to be far more challenging than was first thought: although virtually every country in Africa has started to implement the strategy,[95] not one

country yet has high enough coverage at national level to achieve a measurable impact on mortality. This lack of high-level coverage, and an intervention perceived to be very expensive by many donors, may have led to a general feeling among international policy-makers that IMCI has no impact on child survival. Our data suggest that high coverage of facility-based IMCI leads to reduced child mortality and that this is achievable within existing health budgets. In our setting, simple, practical planning and management tools for strengthening the capacity of district health systems were the essential first step to achieving this impact.

5

Econometric Estimation of Country-Specific Hospital Costs

Reprinted from: Adam T, Evans DB and Murray CJL. Econometric Estimation of Country-Specific Hospital Costs. Cost Effectiveness and Resource Allocation Journal (CERA) 2003 1:3.

CHAPTER 5

Chapter 5. Econometric estimation of country-specific hospital costs

Summary

Information on the unit cost of inpatient and outpatient care is an essential element for costing, budgeting and economic-evaluation exercises. Many countries lack reliable estimates, however. WHO has recently undertaken an extensive effort to collect and collate data on the unit cost of hospitals and health centres from as many countries as possible; so far, data have been assembled from 49 countries, for various years during the period 1973–2000. The database covers a total of 2173 country-years of observations. Large gaps remain, however, particularly for developing countries. Although the long-term solution is that all countries perform their own costing studies, the question arises whether it is possible to predict unit costs for different countries in a standardized way for short-term use. The purpose of the work described in this paper, a modelling exercise, was to use the data collected across countries to predict unit costs in countries for which data are not yet available, with the appropriate uncertainty intervals.

The model presented here forms part of a series of models used to estimate unit costs for the WHO-CHOICE project. The methods and the results of the model, however, may be used to predict a number of different types of country-specific unit costs, depending on the purpose of the exercise. They may be used, for instance, to estimate the costs per bed-day at different capacity levels; the "hotel" component of cost per bed-day; or unit costs net of particular components such as drugs.

In addition to reporting estimates for selected countries, the paper shows that unit costs of hospitals vary within countries, sometimes by an order of magnitude. Basing cost-effectiveness studies or budgeting exercises on the results of a study of a single facility, or even a small group of facilities, is likely to be misleading.

CHAPTER 5

5. 1 Introduction

Information on hospital unit costs is valuable to health decision-makers and researchers for at least three purposes: budgeting (now receiving more attention with the availability of additional funds for health in poor countries through the Global Fund to Fight AIDS, Tuberculosis and Malaria); the assessment of hospital efficiency; and the assessment, by means of either cost-benefit or cost-effectiveness analysis, of the efficiency of different health interventions. Recognizing the need to make this information available on a country-specific basis, WHO has undertaken an extensive effort to collate all sources of data on unit costs from as many countries as possible.[i116] Large gaps remain, however, particularly for developing countries. Although the long-term solution is that all countries perform their own costing studies, the question arises whether it is possible to predict unit costs for different countries in a standardized way for short-term use. The purpose of the work described in this paper is to use the data collected across countries to predict unit costs in countries for which data are not yet available. Both point estimates and uncertainty intervals are reported.

The model presented here forms part of a series of models that can be used to predict country-specific unit costs for a number of purposes. They may be used, for instance, to estimate: (i) unit costs at different capacity levels for the purposes of efficiency analysis or economic evaluation of health interventions; (ii) the "hotel" component of average cost per bed-day for budgeting exercises; or (iii) unit costs excluding components that might be funded from other sources, such as drugs. The paper first reviews the literature on cost-function estimations. Next, it describes the data sources and methods used. It concludes with a discussion of the results, the application of the models, and continuing work and future directions.

5. 2 Background

Health economics has a long tradition of estimating hospital-cost functions econometrically.[67;117-124] Econometric models explain how total costs change in response to differences in service mix, inputs, input prices, and scale of operations. They allow cost and production functions to be specified with sufficient flexibility that a non-linear relationship can be demonstrated

[i] This work is part of the WHO-CHOICE (**CHO**osing **I**nterventions that are **C**ost-**E**ffective – see www.who.int/evidence) project, which makes use of a standardized set of methods and tools developed to analyze the costs and population health impact of current and possible new interventions in a standardized and generalizable way.[23;115]

between costs and quantity of inputs: total costs can rise at a lower rate than prices.[67]

Previous studies have commonly used microeconomic data to analyse and estimate hospital-cost functions. This literature indicates two main approaches: behavioural cost functions and cost minimization functions.[67;117;123;125] Behavioural cost functions have been used to explain the variations in cost per unit of output among hospitals. They have used as determinants all variables for which a causal relationship to hospital costs is hypothesized and data are available—e.g., bed size, global indicators of hospital activity such as average length of stay and occupancy rate, dummy variables for teaching status, etc. On the other hand, the literature on cost minimization has described the minimum cost of providing a given volume of output as a function of an exogenous vector of input prices and the volume of output. The purpose is to determine whether hospitals are cost-minimizers (profit maximizers).

When testing the hypothesis of cost-minimization, the explanatory variables typically comprise only of output quantities (e.g., number of bed days) and input prices. The remaining variables used in the behavioural cost function specification are not part of the cost minimization question but can be used to explain deviation of observed unit costs from the theoretical minimum functions – e.g., possible reasons for inefficiency.[117]

To our knowledge, all previous studies have used within-country data sets; we know of none that has attempted to estimate hospital-cost functions across countries. Such studies require a large number of observations from as many countries as possible.

The model described here follows the tradition of the behavioural cost function literature because its purpose is to estimate country-specific costs per bed-day, not to test the hypothesis of cost-minimization. The analysis controls for across-country price-level differences by using unit costs adjusted for purchasing-power parity, namely in international dollars; and for differences in quantity and complexity of resource use by using macro-level indicators such as per capita GDP.[126-128]

5. 3 Objectives

The specific objectives of this paper are to:
explain the observed differences in hospital inpatient cost per bed-day across and within countries; and

use the results to predict cost per bed-day for countries for which these data are not yet available.

5. 4 Methods

5.4. 1 Data

The search sources used to obtain the data were: Medline, Econlit, Social Science Citation Index, regional Index Medicus, Eldis (for developing-country data), Commonwealth Agricultural Bureau (CAB), and the British Library for Development Studies Databases. The range of years was set at 1960 to the present. Data covering costs and charges were included.

The search terms used were: "costs and cost analysis" and hospital costs or health centre or the abbreviations HC (health centre) or PHC (primary health centre) or outpatient care**.** The language sources searched were English, French, Spanish and Arabic; no Arabic study was found. In addition, a number of studies were found in the grey literature, from such sources as electronic databases, government regulatory bodies, research institutions, and individual health economists known to the authors.[56;67;129-162] Also included were data from a number of WHO-commissioned studies on unit costs.

A standard template was used for extracting data from all sources. Database variables include: ownership; level of facility (see Annex 5. 1 for a definition of hospital levels); number of beds; number of inpatient and outpatient specialties; cost data (cost per bed-day, outpatient visit, and admission); utilization data (bed-days, outpatient visits, admissions); types of cost included in the cost analysis (capital, drugs, ancillary, food) and whether they were based on costs or charges; capacity utilization (occupancy rate, average length of stay, bed turnover, and average number of visits per doctor per day); reference year for cost data; currency, and methods of allocation of joint costs. The database consists of unit-cost data from 49 countries for various years between 1973–2000, totalling 2173 country-years of observations. Some studies provided information on 100% of the variables described above; at the other extreme, some provided information on less than 15%. The number of observations used in this analysis was 1171 (see Annex 5. 2 for percentage of missing data in the model variables and Annex 5. 3 for the list of countries).

Data cleaning comprised consistency checks and direct derivation of some of the missing variables, when possible, from other variables from the same observation (e.g., occupancy rate was calculated from number of beds and number of bed-days). STATA software was used for data analysis.[163]

Cost data were converted to 1998 International dollars by means of GDP deflators[164] and purchasing-power-parity exchange rates used as the basis for WHO's national health accounts estimates[ii].

5.4. 2 Data Imputation

Most statistical procedures rely on complete-data methods of analysis: computational programs require that all cases contain values for all variables to be analyzed. Thus, as default, most software programs exclude from the analysis observations with missing data on any of the variables (list-wise deletion). This can give rise to two problems: compromised analytical power, and estimation bias. The latter occurs, for example, if the probability that a particular value is missing is correlated with certain determinants. For example, if the complete observation sets tend to be from observations with unit costs that are systematically higher or lower than average, the conclusions for out-of sample estimation drawn from an analysis based on list-wise deletion will be biased upwards or downwards.[165]

There is a growing literature on how to deal with missing data in a way that does not require incomplete observation sets to be deleted, and several software programs have been developed for this purpose. If data are not missing in a systematic way, missing data can be imputed using the observed values for complete sets of observations as covariates for prediction purposes. Multiple imputation is an effective method for general-purpose handling of missing data in multivariate analysis; it allows subsequent analysis to take account of the level of uncertainty surrounding each imputed value, as described below.[166-169] The statistical model used for multiple imputation is the joint multivariate normal distribution. One of its main advantages is that it produces reliable estimates of standard errors: single imputation methods do not allow for the additional error introduced by imputation. In addition, the introduction of random error into the imputation process makes it possible to obtain largely unbiased estimates of all parameters.[166]

In this study, multiple imputation was performed with *Amelia*, a statistical software program designed specifically for multiple imputation of missing data.[165;167;170;171] First, five completed-data sets are created by imputing the unobserved data five times, using five independent draws from an imputation model. The model is constructed to approximate the true distributional relationship between the unobserved data and the available information. This reduces potential bias due to systematic difference between the observed and the unobserved data. Second, five complete-data analyses are performed by

[ii] Purchasing power parity exchange rates used in this analysis are available from the WHO-CHOICE website: www.who.int/evidence/cea

treating each completed-data set as an actual complete-data set; this permits standard complete-data procedures and software to be utilized directly. Third, the results from the five complete-data analyses are combined[172] to obtain the so-called repeated-imputation inference, which takes into account the uncertainty in the imputed values.

5.4. 3 Model specifications

From the tradition of using cost functions to explain observed variations in unit costs, we estimate a long-run cost-function by means of Ordinary Least Squares regression analysis (OLS); the dependent variable is the natural log of cost per bed-day.[67;117;120-122;173] The primary reason for using unit cost rather than total cost as the dependent variable is to avoid the higher error terms due to non-uniform variance (heteroscedasticity) in the estimated regression. This could arise if total cost were used as the dependent variable, as the error term could be correlated with hospital size.[67;117] The reason for using cost per bed-day rather than cost per admission is that "bed-days" are better than "admissions" as a proxy for such hospital services as nursing, accommodation and other "hotel services"[117], permitting more flexibility in the use of estimated unit costs.

As the relationship between unit costs and the explanatory variables are expected to be non-linear, the Cobb-Douglas transformation was used to approximate the normal distribution of the model variables. Natural logs were used. The Cobb-Douglas functional form can be written as follows:

Equation 5. 1

$$Y = \alpha_0 X_1^{\alpha_1} X_2^{\alpha_2} \quad \text{or ,}$$

Equation 5. 2

$$\ln(Y) = \delta + \alpha_1 \ln(X_1) + \alpha_2 \ln(X_2)$$

where $\delta = \ln(\alpha_0)$. This function is non-linear in the variables *Y*, X_1 and X_2, but it is linear in the parameters $\delta, \alpha_1, \alpha_2$, and can be readily estimated using Ordinary Least Squares.[174]

Log transformation has the added advantage that coefficients can be readily interpreted as elasticities.[117;174]

Therefore, the cost function specification of the OLS regression model may be written as:

Equation 5. 3

$$UC_i = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i + e_i$$

Where UC_i is the natural log (ln) of cost per bed-day in 1998 I $ in the *ith* hospital; X_1 is ln of GDP per capita in 1998 I $; X_2 is ln of occupancy rate; $X_{3,4}$ are dummy variables indicating the inclusion of drug or food costs (included =1);$X_{5,6}$ are dummy variables for hospital levels 1–2 (the comparator is level 3 hospital $X_{7,8}$ are dummy variables indicating facility ownership (comparator is private not-for-profit hospitals); X_9 is a dummy variable controlling for USA data (USA=1); and *e* denotes the error term.

The choice of explanatory variables is partly related to economic theory and partly determined by the purpose of the exercise, which is to estimate unit costs for countries where the data are not available. In this case, the chosen explanatory variables must be available in the out-of-sample countries. Country-specific[iii] per capita GDP in international dollars (I $) is used as a proxy for level of technology;[126-128] occupancy rate as a proxy for level of capacity utilization; and hospital level as a proxy for case mix. Unit costs are expected to be correlated positively with GDP per capita and case mix and negatively with capacity utilization.

The inclusion of the seven control variables makes it possible to estimate unit cost for different purposes to suit different types of analysis— for example, cost per bed-day in a primary-level hospital, which does not provide drugs or food; or the cost in a tertiary level hospital, with drugs and food included.

The dummy for the USA was included because all data were charges rather than costs and because there were a large number of observations from that country. Dummies for countries other than the USA with a large number of observations, such as China and the United Kingdom, were also tested as was the use of dummy variables to capture whether the cost estimates included capital or ancillary costs. These variables were not included in the model which best fit the data. Utilization variables, such as number of bed-days or outpatient visits, and hospital indicators, such as average length of stay, were not included as explanatory variables because most out-of-sample countries do not have data on these variables, and prediction of unit costs would, therefore, be impossible.

[iii] Whenever possible, provincial GDP per capita was used — for China for example. The source of these data was the National Statistics report.

5.4. 4 Model-fit

Regression diagnostics were used to judge the goodness-of-fit of the model. They included the tolerance test for multicollinearity, its reciprocal variance inflation factors and estimates of adjusted R square and F statistics of the regression model.

5.4. 5 Predicted values and uncertainty analysis

Two types of uncertainty arise from using statistical models: estimation uncertainty arising from not knowing β and α perfectly – an unavoidable consequence of having a finite number of observations; and fundamental uncertainty represented by the stochastic component as a result of unobservable factors that may influence the dependent variable but are not included in the explanatory variables.[170] To account for both types of uncertainty, statistical simulation was used to compute the quantities of interest, namely average cost per bed-day and the uncertainty around these estimates. Statistical simulation uses the logic of survey sampling to learn about any feature of the probability distribution of the quantities of interest, such as its mean or variance.[170]

It does so in two steps. First, simulated parameter values are obtained by drawing random values from the data set to obtain a new value of the parameter estimate. This is repeated 1000 times. Then the mean, standard deviation, and 95% confidence interval around the parameter estimates are computed. Second, simulated predicted values of ŷ (the quantity of interest) are calculated, as follows: (1) one value is set for each explanatory variable; (2) taking the simulated coefficients from the previous step, the systematic component (g) of the statistical model is estimated, where g= f (X,B); (3) the predicted value is simulated by taking a random draw from the systematic component of the statistical model; (4) these steps are repeated 1000 times to produce 1000 predicted values, thus approximating the entire probability distribution of ŷ. From these simulations, the mean predicted value, standard deviation, and 95% confidence interval around the predicted values are computed. In this way, this analysis accounts for both fundamental and parameter uncertainty.

The predicted log of cost per bed day, ln $\overline{UC}$, can then be calculated from:

Equation 5. 4

$$\ln \overline{UC} = \overline{\alpha_0} + \overline{\alpha_1} \ln X_1 + \sum_{i=1}^{n} \overline{\alpha_i} X_i$$

where $\overline{\alpha_0}$ and $\overline{\alpha_{i..n}}$ are the estimated parameters, and $X_{i..n}$ are the independent variables. If $\overline{\beta_0}$ =anti log ($\overline{\alpha_0}$) and $\overline{\beta_1} = \overline{\alpha_1}$, back-transforming Equation 5. 4 (reduced to 1 independent log-transformed variable for simplicity) gives the power function.

Equation 5. 5

$$\overline{UC}_{biased} = \overline{\beta_0} X^{\overline{\beta_1}}$$

where $\overline{UC}_{biased}$ denotes a biased estimate of the mean cost per bed-day due to back-transformation. This is because one of the implicit assumptions of using log-transformed models is that the least-squares regression residuals in the transformed space are normally distributed. In this case, back-transforming to estimate unit costs gives the median and not the mean. To estimate the mean it is necessary to use a bias correction technique. The smearing method described by Duan (1983) was used to correct for the back-transformation bias.[175] The smearing method is non-parametric, since it does not require the regression errors to have any specified distribution (e.g., normality). If the n residuals in log space are denoted by r_i, and b is the base of logarithm used, the smearing correction factor, $\overline{C}_{bias}$, for the logarithmic transformation is given by:

Equation 5. 6

$$\overline{C}_{bias} = \frac{1}{n}\sum_{i=1}^{n} b^{r_i}$$

Multiplying the right side of Equation 5. 5 by Equation 5. 6 almost removes the bias, so that:

Equation 5. 7

$$\overline{UC} = \overline{C}_{bias}\,\overline{\beta_0} X^{\overline{\beta_1}}$$

The smearing correction factor ($\overline{C}_{bias}$) for our model was 1.25.

5. 5 Results

Table 5. 1 shows the variable names, description, mean and standard error, estimated after combining the results of the five datasets of the multiple imputation estimates. Table 5. 2 presents the results of the best-fit regression model. The adjusted R square of the combined regressions is 0.80, with an F

statistic of 509 ($p<0.0001$), indicating that the model explains a large part of the variation of the cost per bed-day across countries.[176] The signs of the coefficients are consistent with the earlier hypotheses. For example, the GDP per capita is positively correlated with cost per bed-day, while the lower the occupancy rate the higher is the cost per bed-day. Unit costs are lower in level-one hospitals than in those of levels two and three. The coefficients for the two main explanatory variables (GDP per capita and occupancy rate) are highly significant ($p<0.0001$), as well as most of the control dummies, e.g., hospital level. The coefficient for food costs is not significant at the 5% level but was included in the model because it added to its explanatory power.

The tolerance test and its reciprocal variance inflation factors (VIF) showed no evidence for multicollinearity between the model variables (Tolerance ranged between 0.2 and 0.89, mean VIF 1.97)[iv].

Table 5. 1 Descriptive statistics of the multiple imputation estimates
N=1171

Variable	Description	Mean	SE
Ln cost per bed day	Natural log of cost per bed day in 1998 I $	4.98	1.63
Ln GDP per capita	Natural log of GDP per capita in 1998 I $	8.90	1.06
Ln occupancy rate	Natural log of occupancy rate	-0.41	0.61
Drug costs	Dummy variable for inclusion of drug costs. Included =1	0.96	0.18
Food costs	Dummy variable for inclusion of food costs. Included =1	0.93	0.25
Level 1 hospital	Dummy variable for level 1 hospital (1)	0.33	0.47
Level 2 hospital	Dummy variable for level 2 hospital	0.41	0.49
Public	Dummy variable for level public hospitals (2)	0.84	0.36
Private for profit	Dummy variable for level private for profit hospitals	0.08	0.27
USA	Dummy variable for USA. USA =1	0.17	0.37

(1) Dummies for levels of hospital are compared with level 3 hospitals
(2) Dummies for hospital ownership are compared with public not-for-profit hospitals

[iv] Tolerance less than 0.05 and VIF more than 20 indicate presence of multicollinearity.

Table 5. 2 Multiple Imputation regression coefficients and SE

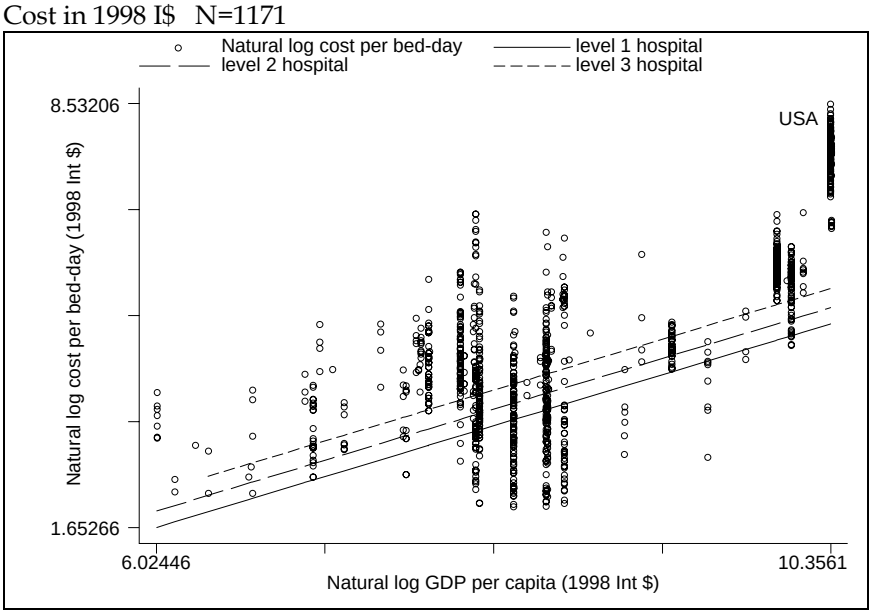
Dependent variable: Natural log of cost per bed-day in 1998 I $
Adjusted R^2= 0.80 F statistic = 509 p of F statistic <0.00001 N: 1171

Variable	B Coef.	SE	t test	P
Ln GDP per capita	0.7624	0.0295	25.813	<0.0001
Ln occupancy rate	-0.2318	0.0474	-4.886	<0.0001
Drug costs	0.6410	0.1769	3.624	<0.0001
Food costs	0.2116	0.1394	1.518	0.152
Level 1 hospital	-0.5777	0.0742	-7.787	<0.0001
Level 2 hospital	-0.3118	0.0594	-5.253	<0.0001
Public	-0.2722	0.1172	-2.323	0.021
Private for profit	0.2444	0.1316	1.857	0.064
USA	1.7471	0.1022	17.104	0.000
Constant	-2.5036	0.3264	-7.672	0.026

Figure 5. 1 shows the three regression lines of levels one, two and three hospitals, respectively, plotted against the log of GDP per capita (the Y-axis is log of cost per bed-day). The regression lines were estimated for public hospitals, with occupancy rate of 80% , including food costs and excluding drugs. Because the original data had a lower average occupancy rate (mean 71% , SD 39%), and most observations included drug costs, it is to be expected that the regression lines will be slightly lower than the actual data points in the database. The regression lines do not pass through the USA data points situated at the upper right side of the graph because they have been calculated for the case where the US dummy was set at zero.

Overall, Figure 5. 1 shows that the regression lines have a good fit with the data used to develop the model. They not only illustrate the relationship between cost per bed-day, hospital level and GDP per capita, but also show that there remains substantial variation in unit costs for any given level of GDP per capita. It would be inadvisable, therefore, to base cost estimates on a single estimate of hospital costs in a particular setting, something that is a common feature of cost-effectiveness studies.

Figure 5. 1 Regression lines for level one, two and three hospitals against the natural log of GDP per capita. (The Y-axis is the dependent variable: natural log of cost per bed day)



To use the equation reported in Table 5. 2 to predict unit costs for a number of in and out-of-sample countries, with the appropriate uncertainty interval, requires consideration of the probability distributions of the predicted unit costs, given a specified level of the model variables. In order to derive these distributions, simulation techniques were used following the steps described in the Methods section. Table 5. 3 presents for selected countries in different regions of the world the average simulated predicted values and 95% uncertainty intervals. The estimates are presented in 2000 I $, based on the 2000 GDP per capita in I $ and assuming that the estimated coefficients will remain constant over a short time period. They are specific to public hospitals, at an occupancy rate of 80%, excluding drug, but including food costs. Regional estimates of cost per bed day, with the same characteristics described above, are available from the WHO-CHOICE website: www.who.int/choice.

Table 5. 3 Predicted cost per bed-day (i) in 2000 I $

Country	GDP per capita (I$)	In or out-of-sample	Hospital level	Cost per bed day			
				Mean (I $)	95% uncertainty interval Low	95% uncertainty interval high	SD
Mali	581	Out	I	7.39	5.46	9.80	1.36
			II	9.64	7.07	12.73	1.74
			III	13.14	9.58	17.40	2.44
Mozambi-que	720	Out	I	8.70	6.45	11.43	1.58
			II	11.35	8.32	15.01	2.03
			III	15.46	11.31	20.42	2.85
Algeria	1,449	Out	I	14.82	10.95	19.45	2.65
			II	19.35	14.26	25.45	3.39
			III	26.34	19.41	34.63	4.72
Indonesia	3,167	Out	I	26.90	19.86	35.19	4.77
			II	35.12	26.05	46.12	6.11
			III	47.80	35.38	63.28	8.44
Ecuador	3,260	In	I	27.50	20.30	35.95	4.88
			II	35.90	26.63	47.17	6.25
			III	48.87	36.17	64.71	8.63
Romania	3,634	Out	I	29.88	22.05	39.11	5.30
			II	39.00	28.86	51.25	6.79
			III	53.09	39.23	70.35	9.37
Greece	6,192	Out	I	44.88	33.02	59.23	8.03
			II	58.58	43.25	76.84	10.28
			III	79.73	59.01	104.62	14.10
Russian Federation	8,035	In	I	54.75	40.17	72.43	9.86
			II	71.48	52.73	94.24	12.62
			III	97.27	71.79	127.92	17.27
Bahrain	14,159	Out	I	84.41	61.48	112.10	15.52
			II	110.19	80.27	146.22	19.85
			III	149.93	110.19	198.97	27.01
United Arab Emirates	20,330	Out	I	111.30	80.27	148.59	20.81
			II	145.28	105.96	191.58	26.62
			III	197.66	144.13	263.01	36.09
United Kingdom	24,348	In	I	127.76	91.83	170.79	24.11
			II	166.77	121.08	221.00	30.85
			III	226.87	164.90	302.91	41.76

Country	GDP per capita (I$)	In or out-of-sample	Hospital level	Cost per bed day			
				Mean (I $)	95% uncertainty interval Low	95% uncertainty interval high	SD
Canada	28,087	Out	I	142.51	102.03	190.22	27.11
			II	186.02	134.55	247.23	34.69
			III	253.05	183.16	337.26	46.89

(i) Cost per bed day is estimated for public hospitals with 80% occupancy rate, excluding drug costs and including food costs.

5. 6 Discussion and Policy Relevance

This paper describes recent work on developing models to predict country-specific hospital unit costs, by level of hospital and ownership, for countries where these data are not available. The main purpose of this work was to feed into estimates of the costs and effects of many types of health interventions in different settings. Estimates are typically available for variables such as the number of days in hospital, or the number of outpatient visits, for certain types of interventions, but unit prices are not available for many countries. The model presented in this paper used all data on unit costs that could be collected after a thorough search to estimate costs for countries where this information does not exist. Data imputation techniques were used to impute missing data, which has the advantage of eliminating the bias introduced by list-wise deletion of observations in cases where information for some of the variables required by the model is missing.

The goodness-of-fit of the model was tested by various regression diagnostic techniques including the tolerance test for multicollinearity, adjusted R square and F statistic. All suggested a good fit of the model with the data and that GDP per capita could be used to capture different levels of technology use across countries. Although this is the first time that costs have been compared across countries, the signs of the coefficients are consistent with results from previous microeconomic studies within countries. For example, these studies have found that occupancy rate was negatively correlated with cost per bed-day while hospital level had the opposite relationship, something also found in the model presented in this paper.[178;179] This adds confidence to the estimated results.

In addition, the estimates produced by this model were sent to health economists and researchers in different countries to check their face validity. Experts from countries in all WHO regions, covering wide differences in GDP per capita and in technologies typically found in hospitals were consulted, including Benin, Canada, Ecuador, Egypt, Kenya, Netherlands and Thailand. They were provided with a description of the estimated unit cost (e.g., which costs were included) and were asked whether they thought they approximated the average cost per bed-day in their countries. All indicated that the results had face validity.

It is of particular note that the model incorporates a more extensive database on unit costs by hospital level and ownership than has previously been available. Increasing the range of observations will increase the validity of extrapolations of cost estimates for countries in which these data are not available. Additional sources of data are being sought for this purpose and to assist countries to develop their own studies. As this body of information grows, the predictive power of unit-cost models will continue to increase.

There are other possible uses of this model such as estimating the possible costs of scaling-up health interventions for the poor, which is receiving increasing attention with the activities of such bodies as the Global Fund to Fight AIDS, Tuberculosis and Malaria. This can be done in many ways, according to the objectives of the analysis. It may be used, for instance, to estimate:
unit costs at different capacity levels for purposes of efficiency analysis or economic evaluation of health interventions;
the "hotel" component of average cost per bed-day;
unit costs, excluding specific items such as drugs or food costs.

Finally, it must be emphasized that there is wide variation in the unit costs estimated from studies within a particular country (Figure 5. 1). These differences are sometimes of an order of magnitude, and cannot always be attributed to different methods. This implies that analysts cannot simply take the cost estimates from a single study in a country to guide their assessment of the cost-effectiveness of interventions, or the costs of scaling-up. In some cases, they could be wrong by an order of magnitude.

Annex 5. 1 Definition of facility types as coded in the unit cost database

Facility type	Description
Primary-level hospital	Has few specialities, mainly internal medicine, obstetrics-gynecology, pediatrics, and general surgery, or only general practice; limited laboratory services are available for general but not for specialized pathological analysis; bed capacity ranges from 30 to 200 beds; often referred to as a district hospital or first-level referral.
Secondary-level hospital	Highly differentiated by function with five to ten clinical specialities; bed capacity ranging from 200-800 beds; often referred to as provincial hospital.
Tertiary-level hospital	Highly specialized staff and technical equipment, e.g., cardiology, ICU and specialized imaging units; clinical services are highly differentiated by function; may have teaching activities; bed capacity ranges from 300 to1,500 beds; often referred to as central, regional or tertiary-level hospital.

Definition of hospital levels (adapted from Barnum and Kutzin 1993 [67])

Annex 5. 2 Percentage of missing data in the model variables prior to data imputation

Variable name	Description	% missing
Ln GDP per capita	Natural log of GDP per capita in 1998 I $	0
Ln occupancy rate	Natural log of occupancy rate	48
Drug costs	Dummy variable for inclusion of drug costs. Included =1	3
Food costs	Dummy variable for inclusion of food costs. Included =1	19
Level 1 hospital	Dummy variable for level 1 hospital	16
Level 2 hospital	Dummy variable for level 2 hospital	16
Public	Dummy variable for level public hospitals	1
Private for profit	Dummy variable for level private for profit hospitals	1
USA	Dummy variable for USA to control for charges data. USA=1	0

Annex 5. 3 Countries included in the model

Country	N	Country	N
Australia	64	Nepal	3
Bangladesh	21	Netherlands	1
Benin	1	New Zealand	4
Bolivia	1	Niger	2
Cambodia	1	Norway	6
China	367	Papua New Guinea	8
Colombia	1	Poland	4
Ecuador	70	Republic of Korea	32
Egypt	5	Russian Federation	22
Ethiopia	1	Rwanda	4
Ghana	2	Saint Lucia	1
Indonesia	5	Sri Lanka	93
Italy	2	Thailand	41
Jamaica	3	Turkey	1
Kenya	7	United Kingdom	176
Lebanon	4	United Republic of Tanzania	7
Malawi	2	United States of America	203
Mexico	2	Zimbabwe	2
Namibia	2	Total	1171

6

Determinants of Variation in the Cost of Inpatient Stays versus Outpatient Visits in Hospitals.

A Multi-Country Analysis

CHAPTER 6

Chapter 6. Determinants of variation in the cost of inpatient stays versus outpatient visits in hospitals. A Multi-country analysis

---

Summary

Information on hospital costs is key to many types of economic analyses yet many countries lack reliable estimates due partly to the time and resource requirements to undertake detailed costing studies. Accordingly, some analysts have used simple rules of thumb to estimate hospital unit costs, e.g., total hospital costs are allocated between departments assuming that the cost of an inpatient day equals a fixed number of outpatient visits. This paper first explores the extent to which these simple rules apply within and across countries. It then identifies determinants of variation in the relationship between the cost of outpatient visits and inpatient days, then uses the estimated relationship to calculate average costs of inpatient and outpatient stays for countries where data are not yet available. Cost information from 832 hospitals in 28 countries are used. Simple rules of thumb do not prove to be an accurate basis for cost estimates. The ratio of inpatient to outpatient unit costs varies with GDP per capita, hospital size, ownership, and occupancy rate. We show how the estimated relationship can be used to calculate a mean cost of inpatient stays and outpatient visits, taking into account differences in the levels of key determinants, and argue that this method can be used to estimate costs in settings where data are unavailable. Moreover, we suggest that the observed great variation in unit costs for similar hospitals in the same country means that this method might well be preferable to basing policy advice on the results of costing studies that cover only one, or a few hospitals, which might well be outliers.

6. 1 Introduction

The increasing recognition of the importance of evidence-based decision making in health policy, both at the national and international level, has highlighted several gaps in unit cost information. For example, to inform resource allocation decisions, good quality information on the costs and outcomes associated with different ways of using these resources is necessary. In many developing countries, such data are not available at all, or are based on only a few observations. Moreover, variations in methods used to estimate costs makes it difficult to be sure that differences reflect true differences in resource use and/or their prices rather than differences in the estimation methods.[180]

Information on hospital costs is an example. It is a key requirement for many types of policy decisions and is used, for example, as an input to assessing the relative efficiency of various types of treatment, and of treatment compared to prevention. It is also essential for budgeting and planning exercises, to identify the resources necessary to undertake or sustain interventions or to scale up coverage of current interventions.[67;148;181] Many countries lack reliable estimates, and when they are available they are often available for a small number of hospitals, sometimes only one, or limited to the costs of hospitalization for specific diseases or conditions.[139;148;149;181-184]

Admittedly, some of the more accurate methods of estimating costs, such as step-down costing - considered to be the gold standard - are expensive or difficult to undertake, and budget constraints often mean they cannot be used. Accordingly, some analysts have used simple rules-of-thumb, where the cost of an inpatient bed day is assumed to equal the cost of a specific number of outpatient visits, generally three or four. These ratios are then used to allocate total hospital costs between inpatient and outpatient departments in a hospital.[16;67;185;186]

This approach is relatively straightforward as it requires only information on total hospital costs, and the numbers of in- and out-patients. However, the validity of applying the same ratio to all types of hospitals within a country has been questioned by Lombard et al (1991), who showed that in Cape Town Province of South Africa, the relationship between the cost per inpatient day and outpatient visit varied somewhat by type of hospital - the average cost per bed day varied between 1.4 to 2 times the cost per outpatient visit,[16] compared to the three or four assumed in earlier studies.[16;67;185;186]

Recognizing the need to make unit cost information available on a country-specific and hospital-level basis, WHO collated data on unit costs from as many countries and hospitals as possible, as part of its WHO-CHOICE

project.[116] No information could be obtained for many countries. This raised the question of whether the countries seeking this data to make decisions about how to allocate resources between interventions should simply wait until full costing studies had been undertaken, or whether there was a short term method that provided relatively reliable results while awaiting the full costing studies. This study therefore had three objectives. The first was to determine if the simple rules of thumb that have been suggested are relatively accurate for the countries and hospitals for which data were available. The second was to understand possible determinants of any observed variations in the ratio of outpatient to inpatient unit costs, while the third was to determine if these determinants could be used predict costs in countries where data did not exist - in the event that the simple rules of thumb proved to be inappropriate.

The paper first describes the data sources and methods. It then presents the results and considers their implications for policy decisions in countries with scanty data.

6. 2 Methods

6.2. 1 Data

The sources of primary hospital cost data were derived from indexed search engines such as Medline and Econlit, the grey literature, electronic hospital unit cost databases, government regulatory bodies, research institutions, and individual health economists known to the authors. Data from a number of WHO-commissioned studies on unit costs were also included. Inclusion criteria included whether sufficient detail was provided on: the methods used for costing, e.g., step-down or direct allocation of indirect costs; costs versus charges data; and which costs were included, e.g., drugs, diagnostics, capital costs etc. The availability of this information permits controlling for key sources of variation in unit cost estimates.

A standard template was used for extracting the data on costs and possible explanators. Database variables included: ownership; level of hospital (see Table 6. 1 for a definition of hospital levels); number of beds; number of inpatient and outpatient specialties; cost data (cost per bed-day, outpatient visit, and admission); utilization data (bed-days, outpatient visits, admissions); types of cost included in the original study (capital, drugs, ancillary, food); whether reported data were costs or charges; capacity utilization (occupancy rate, average length of stay, bed turnover, and average number of visits per

doctor per day); reference year for cost data; currency; and methods of allocation of joint costs.

The main database used for this analysis consists of unit-cost data from 72 countries for various years between 1980 and 2000, totalling 2415 hospital-years of observations. Some studies provided information on 100% of the variables described above; at the other extreme, some provided information on less than 15%. Only 832 data points from 28 countries provided sufficient data to enable the ratio of outpatient to inpatient unit costs to be calculated. Annex 6. 3 reports the list of countries.

Data cleaning comprised consistency checks and direct derivation of some of the missing variables, when possible, from other variables from the same observation (e.g., occupancy rate was calculated from number of beds and number of bed-days). STATA software was used for data analysis.[86]

Cost data were converted to 1998 international dollars by means of GDP deflators[164] and purchasing-power-parity exchange rates taken from WHO's national health accounts estimates which are published annually in the Annex tables of the World Health Report.[187] The PPP exchange rates are available on www.who.int/choice).

6.2. 2 Data Imputation

Before model selection, missing data were explored to determine the nature of missing variables and if there was any systematic pattern of missing data. The number of hospital beds, occupancy rate, average length of stay and the types of costs included in the estimates were the main variables with missing data. Annex 6. 2 shows the percentage of missing data for the variables used in the final model. Because the variables appeared to be missing at random, imputation of missing data was performed. The observed values for the variables that were available were used to predict a distribution of likely values for the unobserved data.

Multiple imputation is an effective method for doing this and it allows subsequent analysis to take account of the level of uncertainty surrounding each imputed value, as described below.[166-169] The statistical model used for imputation is the joint multivariate normal distribution. One of its main advantages is that it produces reliable estimates of standard errors while single imputation methods do not allow for the additional error introduced by imputation. In addition, the introduction of random error into the imputation process makes it possible to obtain largely unbiased estimates of all parameters.[166]

Amelia, a statistical software program designed specifically for multiple imputation of missing data and used widely for this purpose, was used.[165;167;170;171] The first step is to create five completed-data sets by imputing the unobserved data five times, using five independent draws from an imputation model constructed to approximate the true distributional relationship between the unobserved data and the available information. This reduces potential bias due to systematic differences between the observed and the unobserved data. The second step is to perform five separate complete-data analyses. This permits standard data analysis procedures and software to be used. The third step is to combine the results from the five complete-data analyses to obtain the "repeated-imputation inference" which takes into account the uncertainty in the imputed values.[172]

6.2. 3 Model specification

Following Lombard et al (1991), we used multiple regression to estimate the determinants of the ratio of the costs of an outpatient visit to an inpatient day – hereafter called "the ratio".[16] Natural logarithmic transformation of the dependent variable (the ratio) was used to overcome problems of heteroskedasticity arising from the variability in hospital size and the country's overall level of income and, therefore, the level of resources used in inpatient and outpatient departments.[16] The continuous explanatory variables explored in this model were not normally distributed in natural units, so the natural logarithmic transformation was used - it best approximated a normal distribution for these variables (the explanatory variables explored in the model are described subsequently). Log transformation has the added advantage that coefficients can be readily interpreted as elasticities.[117;174]

The functional specification of the Ordinary Least Squares (OLS) regression model may be written as:

Equation 6. 1

$$\ln\left(\frac{UCOP}{UCBD}\right)_i = \alpha_0 + \sum_{i=1}^{n} \alpha_i \mathrm{X}_i + e_i \qquad , i = 1\ldots n$$

where $\ln\left(\frac{UCOP}{UCBD}\right)_i$ is the natural log (ln) of the ratio of the cost per outpatient visit to the cost per bed-day in 1998 international dollars(I$) in the *ith* hospital; α_0 and $\alpha_{1\ldots n}$ are the estimated parameters; e_i denotes the error term ; and the independent variables X 1-11 are as explained in Table 6. 1.

Table 6. 1 Variable names and description

Variable	Description
Ln GDP per capita	Natural log of GDP per capita in 1998 I $
Ln occupancy rate	Natural log of occupancy rate
Ln hospital beds	Natural log of hospital beds
Public	Dummy variable for level public hospitals (public = 1, private =0)
Food costs	Dummy variable for inclusion of food costs. Included =1
Sri Lanka	Dummy variable for Sri Lanka. Sri Lanka=1
Thailand	Dummy variable for Thailand. Thailand=1
China	Dummy variable for China. China=1
Ecuador	Dummy variable for Ecuador. Ecuador=1

The choice of explanatory variables is partly related to economic theory and partly determined by one of the eventual uses of the exercise, which is to estimate unit costs for countries where the data are not currently available. In this case, the chosen explanatory variables must be available in the out-of-sample countries. Country-specific per capita GDP in international dollars is used as a proxy for the available level of technology;[126-128] occupancy rate to capture the level of capacity utilization; and hospital beds as a proxy for hospital level and the complexity of cases treated. Whenever possible, sub-national GDP per capita was used — in China, for example, provincial GDP per capita was used.

Dummy variables were introduced for all countries with relatively large data sets. In addition, unit costs in a minority of countries included food and ancillary services. It was not possible to extract these from the total costs because of lack of access to the raw data, so dummies were used for observations including food or ancillary services. Various interaction terms were also tested, such as the interaction between hospital beds and country dummies. Other types of possible determinants, such as number of bed-days or outpatient visits, and average length of stay, were not included partly because they were not available in many of the studies, and partly because most out-of-sample countries do not have data on these variables readily at hand, and prediction of unit costs would, therefore, be impossible.

6.2. 4 Model-fit

Regression diagnostics were used to judge the goodness-of-fit of the model. They included the tolerance test for multicollinearity, its reciprocal variance inflation factor, the plot of the residual versus fitted values and residuals versus each of the independent variables, and estimates of adjusted R square and F statistics of the regression model.

6.2. 5 Predicted values and uncertainty analysis

The expected values were computed using statistical simulation to account for the uncertainty around the estimated parameters of the model. Statistical simulation uses the logic of survey sampling to estimate any feature of the probability distribution of the quantity of interest, such as its mean and variance.[170] This is done as follows. First, simulated parameter values ($\overline{\alpha}, \overline{\beta}$) are obtained by drawing random values from the parameter estimates and the variance-covariance matrix of the estimated model. This is repeated 1000 times. Second, the predicted values of $\hat{y}$ (the quantity of interest) are computed for each value of the simulated parameters and a set value for the explanatory variables. In this way, 1000 predicted values are estimated, thus approximating the entire probability distribution of $\hat{y}$. From these estimates of $\hat{y}$, the mean predicted value, standard deviation, and 95% confidence interval around the predicted values are computed.

To estimate the expected ratio of cost per outpatient visit to cost per bed day the dependent variable is transformed to the normal scale using the Duan smearing factor. This factor corrects for the back-transformation bias arising from the fact that while the error term is normally distributed in the log scale, this might not be the case in the normal scale.[175] The smearing method is non-parametric, since it does not require the regression errors to have any specified distribution (e.g., normality). If the n residuals in log space are denoted by r_i, and b is the base of logarithm used, the smearing correction factor, $\overline{C}_{bias}$, for the logarithmic transformation is given by:

Equation 6. 2

$$\overline{C}_{bias} = \frac{1}{n}\sum_{i=1}^{n} b^{r_i}$$

The expected ratio of cost per outpatient visit to cost per bed day can then be calculated by multiplying the anti log of the product of Equation 6. 1 by Equation 6. 2. This gives:

Equation 6. 3

$$\overline{\left(\frac{UCOP}{UCBD}\right)} = \overline{C}_{bias}\ \overline{\beta_0} X^{\overline{\beta_1}}$$

where $\overline{C}_{bias}$ is the Duan smearing correction factor, $\overline{\beta_0}$ =anti log $\overline{\alpha_0}$ and $\overline{\beta}_i$ =anti log $\overline{\alpha_i}$ (i=1…n).

The smearing correction factor ($\overline{C}_{bias}$) for our model was 1.13.

6.2. 6 Practical application of the model

As a practical application of the model, the average cost per bed-day or outpatient visit can be derived in at least one of two ways. First, from the ratios computed in Equation 6. 3, the unit cost per outpatient visit can be calculated as:

Equation 6. 4

$$UCV_i = TC_i * \overline{\left(\frac{UCOP}{UCBD}\right)} / V_i \quad ,$$

where UCV_i is the unit cost per outpatient visit of hospital i, TC_i is total hospital cost, $\overline{\left(\frac{UCOP}{UCBD}\right)}$ is the ratio of cost of the outpatient departments to cost of inpatient departments (assuming an average ratio is used but the specific ratio for hospital i could be used instead) and V_{ij} the number of outpatient visits of the *ith* hospital.

Similarly, the cost per bed day can be calculated as:

Equation 6. 5

$$CBD_i = TC_i * \overline{\left(1 - \frac{UCOP}{UCBD}\right)} / BD_i$$

where CBD_i is the unit cost per bed day of hospital i, BD_i is the number of bed days of the *ith* hospital and all other variables as in Equation 6. 4.

The second alternative to estimating hospital unit costs from the equations presented in this paper is to multiply the estimated ratio with available estimates of unit cost per bed day (e.g., from Adam T et al 2003)[188] to derive the unit cost per outpatient visit. This alternative is particularly useful when information on total hospital costs are not readily available. This is illustrated in Table 6. 3 below for a selected number of countries.

6.2. 7 Face Validity

To validate the results of the model, the estimated cost per outpatient visits described in the second alternative above, and presented in Table 6. 3, were compared to the literature and shared with economists in several countries to determine their face validity.

6. 3 Results

Table 6. 2 shows the results of the best-fit regression model (the description of variable names can be found in Table 6. 1). The adjusted R-square of the combined regressions is 0.63, with an F statistic of 85 ($p<0.0001$). The tolerance test and its reciprocal variance inflation factors (VIF) showed no evidence of multicollinearity between the model variables. Tolerance fell between 0.05 and 0.81 and the mean VIF was 4.91 (tolerance less than 0.05 and VIF more than 20 indicate presence of multicollinearity). Residual plots showed a uniform distribution of the model residuals with a mean of zero and no specific pattern of distribution.

The positive signs of the coefficients of the inpatient occupancy rate and the number of hospital beds are consistent with *a priori* expectations. Other things being equal, unit costs of inpatient care should fall (so the ratio would rise) the greater is the occupancy rate and the greater the size of the hospital – assuming no diseconomies of scale.

For GDP per capita, a positive correlation was found. Several explanations are possible. Higher income countries may be more likely to spend additional resources on outpatient infrastructure or to provide a greater range of outpatient services than lower income countries. They might also be able to spend more time per outpatient than in poorer countries because of less stringent capacity constraints in terms of the availability of trained personnel.

Four country dummies proved to be significant - for China, Ecuador, Sri Lanka and Thailand but the interactions with other explanatory variables did not

(Table 6. 3). This shows that the value of the ratio differed compared to countries with similar levels of GDP per capita. No other interactions proved to be consistently significant across the various models that were tried.

The negative coefficient for hospital ownership (public = 1, private = 0) suggests that public hospitals are likely to spend a higher proportion of their resources on an inpatient day than an outpatient visit. We are not sure of the explanation. Perhaps public hospitals have fewer physical amenities in outpatient wards than private hospitals that must compete for paying patients. Finally, the negative coefficient of the variable denoting the inclusion of food costs (food included =1; no food costs included =0) simply reflects the effect on the ratio when the unit cost per bed-day, the denominator, does not include food. The dummy for ancillary costs was not significant.

Table 6. 2 Ordinary Least Squares Regression coefficients and SE
Dependent variable: Natural log ratio of cost per outpatient visit to cost per bed-day in 1998 I $

Adjusted R^2= 0.61	F statistic = 101	p of F statistic <0.00001		N: 832
Variable	β Coef	SE	T	P
Ln GDP per capita	0.1303	0.0237	5.497	<0.0001
Ln occupancy rate	0.1683	0.0636	2.645	0.011
Ln hospital beds	0.0884	0.0221	4.007	<0.0001
Public	-0.4890	0.1564	-3.125	0.004
Food costs	-0.1985	0.0765	-2.595	0.013
Sri Lanka	-1.2401	0.0828	-14.979	<0.0001
Thailand	-0.2998	0.0778	-3.855	<0.0001
China	0.3449	0.0510	6.763	<0.0001
Ecuador	-0.8187	0.1040	-7.872	<0.0001
Constant	-2.2698	0.2567	-8.843	<0.0001

Figure 6. 1 shows the three regression lines for level one, two and three hospitals, respectively, plotted against the log of GDP per capita (the Y-axis is log of cost per bed-day). The ratio is estimated for public hospitals with 80% occupancy rate, including food costs. The number of beds for a level one hospital was set at the sample average - 106, at 273 for level 2, and at 673 for level 3 hospitals. These values are consistent with the definition of hospital levels in Barnum and Kutzin (1993),[67] see Annex 6. 1. The regression lines are slightly higher than the mid point of the raw data points from Ecuador and Sri Lanka and lower than the mid point of the data from some of the Chinese

provinces. This is because the graphs are based on the country dummies being set at zero.

Figure 6. 1 shows that, holding constant all variables except GDP per capita and hospital level, the estimated ratio gets higher as hospital complexity increases, although the absolute difference between levels is relatively small. This is probably because tertiary, and in some cases, secondary hospitals have specialist outpatient services which are not provided at primary hospitals. The ratio is also inelastic with respect to changes in GDP per capita.

The regression lines hold other determinants constant. The raw data plotted on the Figure illustrate that the unadjusted ratio varies substantially across hospitals, even within a given country. Examination only of the countries with a relatively large number of observations shows a minimum five fold variation in the ratio for Thailand, and a maximum of a 52 fold variation in Ecuador.

Figure 6. 1 Regression lines for level one, two and three hospitals against the natural log of GDP per capita. (The Y-axis is the dependent variable: natural log ratio of cost per visit to cost per bed day) N=832

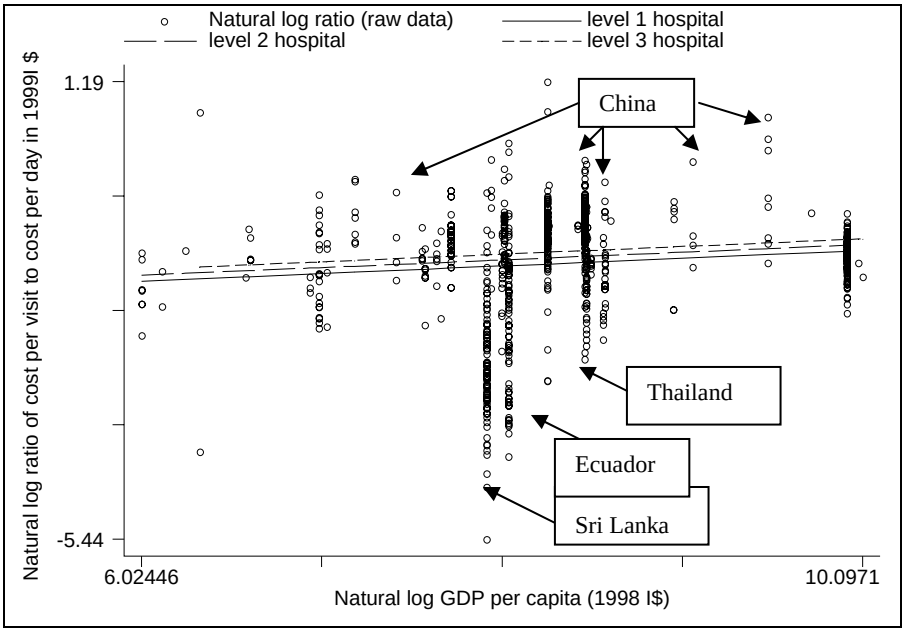


Table 6. 3 presents the mean predicted ratios of inpatient unit costs to inpatient unit costs, and their uncertainty intervals, for a selected number of in and out-of-sample countries. Estimates are based on the equation reported in Table 6. 2 and the methods described earlier. Like Figure 6. 1, the ratio is estimated for public hospitals with 80% occupancy rate, 106 beds for level 1 hospital, 273 beds for level 2 hospital, and 673 beds for level 3 hospital, excluding drugs and ancillary costs (laboratory and other diagnostic procedures) and including food costs. Similarly, cost per bed day is estimated for public hospitals with 80% occupancy rate, excluding drug costs and including food costs based on Adam et al 2003.[188]

The selected countries represent different levels of income and different regions of the world. The four countries for which dummies were used in the model are also presented. The table shows that holding everything constant except GDP per capita, the ratio differs across countries. For example, the minimum ratio was 0.08 for level 1 hospitals in Sri Lanka, reflecting a ratio of the cost per inpatient day to an outpatient visit of 12.5. The maximum was 0.46 in level 3 hospitals in China, where the cost per inpatient day was 2.1 times the cost of an outpatient visit.

Table 6. 3 shows only the variation associated with differences in GDP per capita and hospital level. The actual mean cost for any given country will also depend on the other determinants in Table 6. 2. We tested the calculations of Table 6. 3 on experts from countries covering a wide range of GDPs per capita and hospital technologies. The countries included Benin, Canada, Ecuador, Egypt, Kenya, Netherlands and Thailand. They were provided with a description of the estimated unit cost (e.g., which costs were included) and were asked whether they thought they approximated the average cost per hospital visit in their countries. All indicated that the results had face validity.

6. 4 Discussion and Policy Relevance

This paper describes recent work on developing models to explain the observed variation in the relationship between the cost per outpatient visit to cost per inpatient day across hospitals and countries. Results are based on a much larger data set that had previously been available. The first objective was to determine the extent to which this ratio varied across hospitals. This would show if simple rules of thumb can be used to allocate hospital costs between inpatient and outpatient departments in settings where detailed step-down costing studies have not yet been undertaken.

Table 6. 3 Mean predicted ratio of cost per outpatient visit to cost per bed day and estimated unit costs for a selected number of countries in terms of 2000 I$

Country	GDP per capita	Hospital level	Predicted ratio of cost per visit to cost per bed day [i]				Mean cost per bed day using Adam T et al 2003[188iv]	Mean cost per outpatient visit
			Mean Ratio (a)	Low 95% CI	High 95% CI	SD	(b)	(a*b)
Mali [ii]	581	I	0.20	0.17	0.22	0.02	7.37	1.45
		II	0.21	0.19	0.24	0.02	9.61	2.06
		III	0.23	0.20	0.26	0.02	13.12	3.04
Mozambique [ii]	720	I	0.20	0.18	0.23	0.01	8.68	1.76
		II	0.22	0.20	0.25	0.02	11.31	2.49
		III	0.24	0.21	0.27	0.02	15.45	3.68
Algeria [ii]	1,449	I	0.22	0.20	0.24	0.01	14.80	3.28
		II	0.24	0.22	0.26	0.01	19.28	4.64
		III	0.26	0.23	0.29	0.02	26.33	6.87
Indonesia [ii]	3,167	I	0.24	0.23	0.26	0.01	26.87	6.58
		II	0.27	0.25	0.29	0.01	35.02	9.32
		III	0.29	0.26	0.31	0.01	47.80	13.79
Ecuador [iii]	3,260	I	0.11	0.09	0.13	0.01	27.47	2.99
		II	0.12	0.10	0.14	0.01	35.80	4.24
		III	0.13	0.11	0.15	0.01	48.87	6.28
Sri Lanka [iii]	3,292	I	0.07	0.06	0.08	0.01	25.13	1.79
		II	0.08	0.07	0.09	0.01	32.75	2.53
		III	0.08	0.07	0.10	0.01	44.70	3.74
Romania [ii]	3,634	I	0.25	0.23	0.27	0.01	29.85	7.44
		II	0.27	0.25	0.29	0.01	38.90	10.54
		III	0.29	0.27	0.32	0.01	53.09	15.59
China [iii]	3,727	I	0.35	0.33	0.38	0.01	41.03	14.45
		II	0.38	0.36	0.40	0.01	53.47	20.47
		III	0.41	0.39	0.44	0.01	72.96	30.26
Greece (ii)	6,192	I	0.27	0.25	0.28	0.01	44.85	11.97
		II	0.29	0.27	0.31	0.01	58.45	16.96
		III	0.31	0.29	0.34	0.01	79.76	25.08

Country	GDP per capita	Hospital level	Predicted ratio of cost per visit to cost per bed day [i]				Mean cost per bed day using Adam T et al 2003[188iv]	Mean cost per outpatient visit
			Mean Ratio (a)	Low 95% CI	High 95% CI	SD	(b)	(a*b)
Thailand [iii]	6,626	I	0.20	0.18	0.22	0.01	47.13	9.40
		II	0.22	0.19	0.24	0.02	61.37	13.31
		III	0.24	0.21	0.27	0.02	83.71	19.68
Russian Federation	8,035	I	0.28	0.26	0.29	0.01	54.74	15.11
		II	0.30	0.28	0.32	0.01	71.32	21.41
		III	0.33	0.30	0.35	0.01	97.32	31.65
Bahrain [ii]	14,159	I	0.30	0.28	0.31	0.01	84.42	25.07
		II	0.32	0.31	0.34	0.01	109.99	35.52
		III	0.35	0.32	0.38	0.02	150.06	52.51
United Arab Emirates [ii]	20,330	I	0.31	0.29	0.33	0.01	111.34	34.65
		II	0.34	0.32	0.36	0.01	145.06	49.09
		III	0.37	0.34	0.40	0.02	197.88	72.57
United Kingdom	24,348	I	0.32	0.30	0.34	0.01	127.81	40.72
		II	0.35	0.32	0.37	0.01	166.53	57.69
		III	0.38	0.35	0.41	0.02	227.15	85.28
Canada [ii]	28,087	I	0.32	0.31	0.35	0.01	142.59	46.27
		II	0.35	0.33	0.38	0.01	185.77	65.56
		III	0.38	0.35	0.42	0.02	253.38	96.91

[i] The ratio is estimated for public hospitals with 80% occupancy rate, 106 beds for level 1 hospital, 273 beds for level 2 hospital, and 673 beds for level 3 hospital, excluding ancillary costs (laboratory and other diagnostic procedures) and including food costs.

[ii] Country out of sample.

[iii] Country dummy included in the model.

[iv] Cost per bed day is estimated for public hospitals with 80% occupancy rate, excluding drug costs and including food costs.

CI= confidence interval.

The results suggest that it would be misleading to use any simple rule of thumb, something that has been done a number of times in the recent literature, because inpatient costs - even holding determinants except GDP per capita and hospital level constant - could be as low as two, and as high as 12 times the cost of an outpatient visit. While this might not be totally surprising, another implication is that there was enormous variation in this ratio even for hospitals of the same type in the same setting. This means that even where some step down costing studies have been undertaken, analysts would be unwise to base policy advice on the results of only one or two studies. The hospital or hospitals for which data are available may well be atypical of other hospitals of the same type in that country. The implication is that a sufficiently large random sample of hospitals is really required to provide accurate policy advice, further increasing the costs of doing such research. The second conclusion is that a high proportion of the observed variation in the ratio observed across hospitals can be explained. All regression diagnostics suggested a good fit with the data and the signs of the coefficients have face validity. This suggests that the equations reported here could provide a viable alternative for estimating unit costs than rule of thumb if only a small number of observations from hospital costing studies is available. Knowledge of a few, readily obtainable explanatory variables (i.e. those in Table 6. 1) could be used to estimate the ratio, and hence, the costs of inpatient stays and outpatient visits.

Moreover, we argue that the approach is more likely to provide a reliable estimate of the average costs than relying on a single, or a few, cost studies in a given setting, even if those studies use appropriate methods. This is because of the great variation in unit costs observed for hospitals in the same country, described earlier. This is the approach used in the WHO-CHOICE (www.who.int/choice) project on cost-effectiveness, and subsequently in the Disease Control Priorities in Developing Countries exercise that will be published shortly (http://www.fic.nih.gov/dcpp/).

This work is the first attempt to compare unit cost ratios across countries taking into account hospital and country characteristics. The model incorporates a much more extensive database on unit costs by hospital level and ownership than has previously been available. Increasing the range of observations and including other possible explanatory variables will increase the explanatory power of the model and the validity of extrapolating the results to countries where full step-down studies across a range of hospitals have not yet been undertaken. Although it would be preferable for analysts to estimate hospital costs using the step-down procedure if time and financial resources permit, this paper shows that econometric analysis of existing data can provide useful estimates in the interim.

Annex 6. 1 Definition of facility types as coded in the unit cost database

Facility type	Description
Primary-level hospital	Has few specialities, mainly internal medicine, obstetrics-gynecology, pediatrics, and general surgery, or only general practice; limited laboratory services are available for general but not for specialized pathological analysis; bed capacity ranges from 30 to 200 beds; often referred to as a district hospital or first-level referral.
Secondary-level hospital	Highly differentiated by function with five to ten clinical specialities; bed capacity ranging from 200-800 beds; often referred to as provincial hospital .
Tertiary-level hospital	Highly specialized staff and technical equipment, e.g., cardiology, ICU and specialized imaging units; clinical services are highly differentiated by function; may have teaching activities; bed capacity ranges from 300 to1,500 beds; often referred to as central, regional or tertiary-level hospital.

Definition of hospital levels adapted from Barnum and Kutzin 1993 [67]

Annex 6. 2 Percentage of missing data of the model variables prior to data imputation

Variable name	Description	% missing
Ln GDP per capita	Natural log of GDP per capita in 1998 I $	0
Ln occupancy rate	Natural log of occupancy rate	32
Ln hospital beds	Natural log of hospital beds	32
Public	Dummy variable for level public hospitals (1)	0
Ancillary costs	Dummy variable for inclusion of ancillary (2) costs. Included =1	34
Food costs	Dummy variable for inclusion of food costs. Included =1	25
Costs or charge	Whether observation is cost or	0

Variable name	Description	% missing
	charge data. Costs =1	
Sri Lanka	Dummy variable for Sri Lanka. Sri Lanka=1	0
Thailand	Dummy variable for Thailand. Thailand=1	0
China	Dummy variable for China. China=1	0
Ecuador	Dummy variable for Ecuador. Ecuador=1	0

Annex 6. 3 Countries included in the model

Country	N	Country	N
Australia	1	Namibia	2
Bangladesh	20	Nepal	3
Benin	1	Netherlands	1
Cambodia	1	New Zealand	1
China	358	Papua New Guinea	8
Ecuador	68	Poland	4
Egypt	5	Russian Federation	17
Ethiopia	1	Rwanda	4
Indonesia	5	Saint Lucia	1
Jamaica	3	Sri Lanka	90
Kenya	7	Thailand	41
Lebanon	4	United Kingdom	173
Malawi	2	United Republic of Tanzania	7
Mexico	2	Zimbabwe	2
		Total	832

7

Understanding Variations in Department-Specific Hospital Costs:

An Application of Seemingly Unrelated Regressions to China

CHAPTER 7

Chapter 7. Understanding variations in department-specific hospital costs: An application of seemingly unrelated regressions to China

Summary

This paper tests whether there are simple but reliable methods to estimate hospital unit costs without having to undertake a full costing study. Statistical methods for compositional data are used to estimate and predict department-specific hospital costs using seemingly unrelated regressions. The model explains most of the variation in the department-specific ratio of inpatient to outpatient unit costs, which is shown to be close to the simple rules of thumb (e.g. 3:1 and 4:1) suggested by earlier studies. However, it varies substantially by inpatient department and a number of other explanatory variables. The ratio can reach 6:1.

7. 1 Introduction

Information on hospital costs is valuable to researchers and health decision-makers for at least two reasons. The first is for efficiency assessment, to identify if the benefits of undertaking a hospital-based intervention outweigh the costs, or which of the many possible interventions involving hospitalization is the best use of scarce health resources. The second is for budgeting purposes, to identify the resources necessary to undertake or sustain hospital-based interventions. The recommended approach involves a step-down costing process where overhead and general service costs are distributed across the different patient departments on the basis of a set of allocation rules determined by the main cost drivers of these costs.[15;67] Many countries lack this type of data because they are costly to collect and the analysis requires high levels of technical expertise. Where such studies have

been undertaken, they are often limited to a small number of health facilities that may well be atypical of other settings.

The total costs of an institution is, however, more easy to obtain than the disaggregated costs of inpatient and outpatient care. Some studies have, therefore, used simple rules of thumb to allocate hospital costs between inpatient and outpatient departments as a quick method to estimate the costs of an inpatient stay or an outpatient visit, e.g., assuming that each bed-day costs either three or four times that of an outpatient visit.[16;67] The ratio of inpatient to outpatient unit costs has, however, been shown to differ between hospitals within a country, varying by type (size) and degree of specialization. In Cape Town province, South Africa, it varied between 2:1 and 1.4:1.[16]

We hypothesize that the ratio will also differ by inpatient department according to variation in the intensity of input use across departments, but to our knowledge, no previous study has tested this. The purpose of this paper is to explore the extent to which the ratio of inpatient to outpatient costs varies by inpatient department, whether there is a consistent pattern in this variation across hospitals and if so, what are the determinants of variation. This will reveal whether simple rules of thumb can be derived for allocating total hospital costs between the different inpatient and outpatient departments for use at times when it is not possible to undertake full step-down costing studies in a large number of hospitals.

Previous studies had estimated the relationship between inpatient and outpatient unit costs using multivariate regression analysis.[16] As our aim is to distinguish between the different types of inpatient departments, an alternative to the use of independent regression models for each department is offered by compositional models, with applications found in a range of disciplines including economics, political science and epidemiology.[189-194] This paper describes the adaptation of compositional models to the estimation of department-specific hospital costs in which we estimate the ratios of department-specific inpatient costs to the cost of outpatient departments at different types of hospital, with appropriate uncertainty intervals.

Seemingly unrelated regressions as developed by Zellner [195] are ideal for this purpose as they simultaneously estimate a system of equations that appear unrelated but where the errors are potentially correlated across equations. This leads to significant improvement in efficiency of the estimation model.[193;195;196] In this case the errors of the equations are likely to be correlated because of factors common to all departments in a particular hospital, such as its location, the hospital type, and the availability of medical specialties. One of the advantages of this system is that fewer observations are required to obtain

reliable parameter estimates than if each of the equations is estimated separately.

The paper begins by defining the characteristics of modelling compositional data using the seemingly unrelated linear regression model. Next, it describes the data sources and methods. It concludes with a discussion of the results, the application of the models, and future directions.

7. 2 Methods

7.2. 1 Data

Forty-one hospitals from 12 provinces of China were included in the empirical analysis (see Annex 7. 1). The data were for the years 1997 and 2000. Based on the frequency of their existence in the data-set, the following inpatient departments were selected for analysis: internal medicine, obstetrics and gynaecology, surgery and paediatrics. All other inpatient departments were grouped into one category called "others". The number of departments included in this category varied between one and seven depending on the hospital, and included ear nose and throat, dermatology, ophthalmology, physiotherapy and Chinese traditional medicine. The outpatient department was the final category.

The database variables included total and unit cost per bed day and per case for each department estimated using full step-down allocation methods; hospital and department-specific indicators such as type of hospital, number of beds, occupancy rate, average length of stay and utilization rates (e.g., number of bed-days). Details of standard step-down allocation methods are available elsewhere.[15;67] Other variables included the year to which the cost data referred, province name and provincial GDP per capita. Costs were converted to year 2000 Chinese Yuan by means of the overall GDP deflator from the World Bank as no health-specific deflators were available for China.[197]

7.2. 2 Statistical model

Compositional data are vectors of proportions describing the relative contributions of each of *J* categories to unity. Seemingly unrelated regressions are commonly used for estimation purposes of compositional models, where the model typically contains several equations and an additional identity.[189;198;199] This identity implies that the *j* dependent variables sum to a

fixed known value. The implicit adding-up condition causes the seemingly unrelated dependent variables to be correlated. Haupt and Oberhofer[199] show that in regression analysis the residuals inherit this correlation from the dependent variables, resulting in a singular covariance matrix. As noted by Bewley, deleting one equation before the estimation procedure is one possible way of solving this problem as the *J - 1* linear independent equations contain the entire statistical information required for estimation.[200] The choice of the equation that should be dropped does not affect the estimated parameters, so one arbitrary equation can be dropped and the remaining *J - 1* equations estimated.[189;199]

Estimating the parameters of a set of regression equations can be more efficient if the independent variables in different equations are not highly correlated and if the disturbance terms are highly correlated.[195;196] We return to this later.

In this paper, we used seemingly unrelated regression methods for compositional data to estimate the ratio of department-specific costs to total hospital costs. Clearly, the department-specific costs must sum to the total hospital costs. In order to model compositional data with *J* different hospital departments, we first define a vector of cost-proportions. The description below uses the same notation as a general statistical model for compositional data that has been presented in an application to multiparty electoral data.[191]

$$P_{ij} = DC_{ij} / TC_i \ , \qquad (1)$$

where P_{ij} ($P_{ij},...,P_{iJ}$) denotes the proportion of a particular department's cost for each department *j* (*j=1,…, J)* to total hospital cost, for each hospital *i (i =1,…,N)*. Six proportions were calculated for each hospital using the departments described earlier.

The second step follows where the compositional data is modelled using the additive logistic normal distribution.[190] First, a (*J*-1) vector Y_i is generated by calculating the log ratios of each department fraction relative to the dropped fraction *J* (outpatient department*):*

$$Y_{ij} = \ln\left(\frac{P_{ij}}{P_{iJ}}\right). \qquad i=1, \dots, N \quad j=1, \dots, J-1 \qquad (2)$$

The vector $Y_{ij} = (Y_{i1},...,Y_{i(J-1)})$ is assumed to be multivariate normal with mean μ and variance matrix Σ. The expectation of each log-ratio is assumed to be a linear function of the explanatory variables in the model.

7.2. 3 Model-fit

The objective was to identify variables that would be likely to have a strong relationship to department-specific costs, but also variables for which estimates would be available in other areas of the country in the event that the model could be used for predictive purposes. Of further interest was the possibility to compare the results of this model with data from other countries in future work. The variables that were selected based on these criteria were occupancy rate reflecting capacity utilization; the ratio of department-specific bed days to total hospital bed-days as a measure of the proportion of overhead costs attributed to each department; and number of staff per department as a measure of department size. In addition, various dummy variables were used to explore if hospitals from the different provinces in China behaved differently. Only the dummy for Henan province – the province where 40% of the data come from - proved significant, so it is the variable reported here. All except the last explanatory variable vary by department. The final version of the model reported here included all variables in the natural log form. This transformation ensured that the distributions were approximately normal. It also provided the best fit.

A number of other explanatory variables were explored including: provincial GDP per capita, dummy variables for hospital type (primary, secondary, tertiary) and department-specific variables such as the average length of stay, number of beds, beds as a proportion of hospital beds and number of bed-days. The equation with the best fit did not include these variables.

Given the specification described above, the multivariate normal model for the log ratios may be written as a system of equations as follows:

$$y_{i1} = \alpha_0 + \alpha_1 W_{i1} + \alpha_2 Z_{i1} + \alpha_3 S_{i1} + \alpha_4 G_i + \varepsilon_{i1}, \quad i=1, \ldots, N \qquad (3a)$$

$$y_{i2} = \gamma_0 + \gamma_1 W_{i2} + \gamma_2 Z_{i2} + \gamma_3 S_{i2} + \gamma_4 G_i + \varepsilon_{i2}, \quad i=1, \ldots, N \qquad (3b)$$

$$y_{i3} = \nu_0 + \nu_1 W_{i3} + \nu_2 Z_{i3} + \nu_3 S_{i3} + \nu_4 G_i + \varepsilon_{i3}, \quad i=1, \ldots, N \qquad (3c)$$

$$y_{i4} = \kappa_0 + \kappa_1 W_{i4} + \kappa Z_{2i4} + \kappa_3 S_{i4} + \kappa_4 G_i + \varepsilon_{i4}, \quad i=1, \ldots, N \qquad (3d)$$

$$y_{i5} = \lambda_0 + \lambda_1 W_{i5} + \lambda_2 lZ_{i5} + \lambda_3 S_{i5} + \lambda_4 G_i + \varepsilon_{i5}, \quad i=1, \ldots, N \qquad (3e)$$

where y_{i1} to y_{i5} are the log-ratios as defined in Eq. (2); W_1 to W_5 are department-specific natural log of occupancy rate; Z_1 to Z_5 are department-specific log ratios of department bed-days over total hospital bed-days; S_1 to S_5 are department-specific log of number of staff; G_i is a dummy variable for Henan province; and ε_{i1} to ε_{i5} are residual terms which have expectations of 0, variances of $\sigma^2{}_1$ to $\sigma^2{}_5$, respectively, and correlations of ρ_s. The variable names, description and mean values are shown in Table 7. 1.

The asymptotically efficient, feasible generalized least-squares algorithm were computed using STATA 8 software.[86;201]

7.2. 4 Goodness of fit

The percentage of variance explained by the explanatory variables was summarized by the adjusted R-squared. F statistics for testing the hypothesis of equal parameter vectors were calculated for each equation.[196] To test the hypothesis of independent equations, the correlation matrix of residuals and the Breusch-Pagan test for independent equations was performed.[86]

Table 7. 1 Variable names, description and mean values

Variable name	Description	Variable notation[vii]	Mean in natural units	SE
Lnp**g**	**Dependent variable:** natural log-ratio of cost of *Ob/Gyn* to cost of outpatient departments		0.04	0.004
ln**g**_occupancy	Natural log of occupancy rate in *Ob/Gyn* department	W_1	0.86	0.066
ln**g**_bday_tbd	Natural log of ratio of bed days in *Ob/Gyn* department to total bed days	Z_1	0.07	0.009
ln**g**_staff	Natural log of number of staff in *Ob/Gyn* department	S_1	44.68	3.235
Henan	Dummy variable for Henan Province. Henan=1	Gi		
Lnp**m**	**Dependent variable:** natural log-ratio of cost of *Internal Medicine* to cost of outpatient departments		0.15	0.015
Ln**m**_occupancy	Natural log of occupancy rate in *Internal Medicine* department	W_2	0.86	0.054
Ln**m**_bday_tbd	Natural log of ratio of bed days in *Internal Medicine* department to total bed days	Z_2	0.22	0.024
Ln**m**_staff	Natural log of number of staff in *Internal Medicine* department	S_2	134.83	13.70
Henan	Dummy variable for Henan Province. Henan=1	Gi		
Lnp**s**	**Dependent variable**: natural log-ratio of cost of *Surgery* to cost of outpatient departments		0.13	0.014
ln**s**_occupancy	Natural log of occupancy rate in *Surgery* department	W_3	0.92	0.074
ln**s**_bday_tbd	Natural log of ratio of bed days in *Surgery* department to total bed days	Z_3	0.22	0.026

[vii] See Equation 3(a) to 3(e)

Variable name	Description	Variable notation[vii]	Mean in natural units	SE
ln**s**_staff	Natural log of number of staff in *Surgery* department	S_3	114.33	9.885
Henan	Dummy variable for Henan Province. Henan=1	*Gi*		
Lnp**p**	**Dependent variable**: natural log-ratio of cost of *Pediatrics* to cost of outpatient departments		0.02	0.002
ln**p**_occupancy	Natural log of occupancy rate in *Pediatrics* department	W_4	0.74	0.06
ln**p**_bday_tbd	Natural log of ratio of bed days in *Pediatrics* department to total bed days	Z_4	0.04	0.005
ln**p**_staff	Natural log of number of staff in *Pediatrics* department	S_4	27.30	1.998
Henan	Dummy variable for Henan Province. Henan=1	*Gi*		
Lnp**oth**	**Dependent variable**: natural log-ratio of cost of *other* departments to cost of outpatient department		0.11	0.015
ln**oth**_occupancy	Natural log of occupancy rate in *other* departments	W_5	0.71	0.063
ln**oth**_bday_tbd	Natural log of ratio of bed days in *other* departments to total bed days	Z_5	0.12	0.018
ln**oth**_staff	Natural log of number of staff in *other* departments	S_5	35.27	6.731
Henan	Dummy variable for Henan Province. Henan=1	*Gi*		

7.2. 5 Computing quantities of interest and uncertainty intervals

We used the method described in Katz to derive the initial fraction vectors $\overline{P_j}$ – the ratio of department-specific cost to total hospital cost[8].[191] This required several steps. First multiple simulation methods were used where, in each of 1000 iterations, a random draw was taken from a multivariate normal distribution around the estimators, with a mean vector consisting of the maximum likelihood estimates of the coefficients and the variance-covariance matrix derived from the regression results. The expected values were then computed for the log-ratios P1 to P5, done 1000 times for each of the original values for the explanatory variables, using the variance matrix constructed from the regression estimates of $\sigma^2{}_1$ to $\sigma^2{}_5$ and ρ_s.

Second, the simulated expected values were back-transformed by multiplying their exponent by the smearing correction factor to correct the bias introduced by the fact that while the expected value of the residuals in the log space has a mean of zero, it might not be the case in the natural scale.[175] The smearing correction was calculated as follows: If the *n* residuals in log space are denoted by r_i, and *b* is the base of logarithm used, the smearing correction factor, $\overline{C}_{bias}$, for the logarithmic transformation is given by:

$$\overline{C}_{bias} = \frac{1}{n}\sum_{i=1}^{n} b^{r_i} \quad (4)$$

Multiplying the anti-log of the expected value of the log-ratios μ_j by Eq. (4) gives:

$$\overline{Y_j}(unbiased) = \overline{C}_{bias} * \exp \mu_j \quad (5)$$

The smearing correction factor ($\overline{C}_{bias}$) of the best-fit model was 1.12.

Third, the product from Eq. (6) was transformed into the initial proportions of department costs to total hospital costs (see Eq. (1)) using the multivariate logistic transformation:

[8] Note that the model estimates the log-ratios of the department-specific costs to outpatient department costs.

$$\overline{P_{j-1}} = \frac{(\overline{Y}_j)}{1+\sum_{j=1}^{J-1}(\overline{Y_j})}, \quad (6)$$

with $\overline{P_6}$ calculated as $1 - \overline{P_1} - \overline{P}_2 - \overline{P}_3 - \overline{P}_4 - \overline{P_5}$.

Thus, in generating 1000 simulations of the fraction vector $\overline{P_j}$, we can summarize the probability distribution of the predicted proportions of the six departments given values of the explanatory variables. In this way, this analysis accounted for both types of uncertainty arising from using statistical models: estimation uncertainty arising from not knowing the estimated coefficients perfectly, and fundamental uncertainty represented by the stochastic component as a result of unobservable factors that may influence the dependent variable but are not included in the explanatory variables.[170]

As a practical application of the model, two quantities of interest can be computed. The first is the department-specific average unit cost ratio - e.g. the cost of an inpatient day in that department to the cost of an outpatient visit. The second is the department-specific average cost per bed-day or outpatient visit. The derivation of these quantities is explained below.

7.2.5. 1 Estimating department-specific average unit cost ratios

Using the estimates from Eq. (6) and information on the number of bed days and outpatient visits we can derive the average department-specific unit cost ratios as follows.

$$(ucw_j / ucopv_i) = (\overline{P_j} * V_i)/(\overline{P_J} * Bd_j), \quad (7)$$

where $(ucw_j / ucopv_i)$ is the ratio of unit cost per bed day of the *jth* department to the unit cost per outpatient visit in the outpatient department, $\overline{P_j}$ is the ratio of cost of the *jth* department to total hospital cost estimated from Eq. (6), V_i is total number of outpatient visits of hospital *i*, $\overline{P_J}$ is the proportion that was omitted from the model (ratio of cost of outpatient department to total cost) which corresponds to $\overline{P_6}$ in Eq. (6) , and Bd_j is the number of bed-days of the *jth* department.

7.2.5. 2 Estimating department-specific unit costs

From the ratios computed in Eq. (6), the unit cost of a particular department can be calculated as:

$$UC_{ij} = TC_i * \overline{P_j} / Bd_{ij} \quad , \qquad (8)$$

where UC_{ij} is the unit cost of the *jth* department of hospital *i*, TC_i is total hospital cost, $\overline{P_j}$ is the ratio of cost of the *jth* department to total hospital cost (assuming an average ratio is used but the specific ratio for hospital *i* could be used instead) and Bd_{ij} the number of bed days of the *jth* department.

7. 3 Results

Table 7. 2 shows the results of the feasible generalized least squares estimates of the system of equations described in Eq. (3a-3e). The model variables explained a large proportion of the variation in department-specific ratios, the R-squared varied between 0.83 and 0.92. The F statistics for each equation are highly significant, varying between 112 and 372. The null hypothesis of a zero correlation between residuals was rejected (Breusch-Pagan test, chi2(10) = 297, $p < 0.00001$), which confirms the appropriateness of using seemingly unrelated regression for this analysis rather than estimating each equation separately, see Table 7. 3.

The signs of the estimated parameters conformed with microeconomic theory and the expected determinants of cost. For example, the higher the number of staff and bed days in the inpatient department, the higher the ratio of the costs of that department to the costs of the outpatient department. Given that the number of bed-days is included in the equation, the occupancy rate can be considered to be an efficiency variable so that holding bed-days constant, a higher occupancy rate reduces the ratio because of the increased efficiency.

As mentioned earlier, 40% of the data come from Henan district. The raw data from this province showed systematically lower proportions for all departments than the rest of the provinces, which is the reason for the statistical significance of the dummy variable for this province. The sign of the regression coefficient suggests also that the ratios were systematically lower in this province, holding other explanators constant. Possible reasons for this finding are discussed in the conclusions.

Table 7. 2 Results of the seemingly unrelated regression using iterated feasible generalized least squares estimates.

Seemingly unrelated regression, iterated — Observations = 41

Equation	RMSE	R-squared	F-Statistic	P>F
lnp**g**	0.5229	0.8579	306.17	<0.0001
lnp**m**	0.4402	0.9095	238.45	<0.0001
lnp**s**	0.4384	0.9201	372.83	<0.0001
lnp**p**	0.5951	0.8942	112.91	<0.0001
lnp**oth**	0.5418	0.8309	134.94	<0.0001

	B Coef	Std. Err.	t	P>t	95% Conf.	Interval
Lnp**g**						
Ln**g**_occupancy	-0.2411	0.0468	-5.16	<0.0001	-0.3334	-0.1489
Ln**g**_bday_tbd	0.6620	0.0321	20.61	<0.0001	0.5986	0.7254
Ln**g**_staff	0.2692	0.0301	8.94	<0.0001	0.2098	0.3287
Henan	-0.5654	0.1651	-3.42	<0.0001	-0.8912	-0.2397
constant	-1.3783	0.1917	-7.19	<0.0001	-1.7565	-1.0002
Lnp**m**						
Ln**m**_occupancy	-0.0024	0.0613	-0.04	0.9690	-0.1236	0.1187
Ln**m**_bday_tbd	0.7797	0.0349	22.32	<0.0001	0.7109	0.8488
Ln**m**_staff	0.2755	0.0353	7.80	<0.0001	0.2058	0.3452
Henan	-0.4916	0.1514	-3.25	0.0010	-0.7904	-0.1928
constant	-0.8996	0.1958	-4.60	<0.0001	-1.2859	-0.5133
Lnp**s**						
Ln**s**_occupancy	-0.1143	0.0287	-3.98	<0.0001	-0.1709	-0.0577
Ln**s**_bday_tbd	0.7813	0.0289	27.05	<0.0001	0.7243	0.8383
Ln**s**_staff	0.2914	0.0336	8.66	<0.0001	0.2250	0.3578
Henan	-0.4528	0.1469	-3.08	0.0020	-0.7427	-0.1629
constant	-1.1119	0.1883	-5.90	<0.0001	-1.4835	-0.7403
Lnp**p**						
Ln**p**_occupancy	-0.3607	0.1310	-2.75	0.0060	-0.6191	-0.1023
Ln**p**_bday_tbd	0.9024	0.0621	14.54	<0.0001	0.7799	1.0248
Ln**p**_staff	0.3459	0.0795	4.35	<0.0001	0.1889	0.5027
Henan	-0.6970	0.2392	-2.91	0.0040	-1.1691	-0.2249
constant	-0.3166	0.2759	-1.15	0.2530	-0.8609	0.2277
Lnp**oth**						
ln**oth**_occupancy	-0.2559	0.0532	-4.81	<0.0001	-0.3609	-0.1509
ln**oth**_bday_tbd	0.6355	0.0356	17.86	<0.0001	0.5652	0.7057
ln**oth**_staff	0.2346	0.0559	4.19	<0.0001	0.1242	0.3449
Henan	-0.6787	0.1729	-3.93	<0.0001	-1.0197	-0.3376
constant	-1.6446	0.2572	-6.39	<0.0001	-2.1523	-1.1369

Table 7. 3 Correlation matrix of residuals:
Breusch-Pagan test of independence: chi2 (10) = 297 P < 0.00001

	lnp**g**	lnp**m**	lnp**s**	lnp**p**	lnp**oth**
lnp**g**	1.00				
lnp**m**	0.94	1.00			
lnp**s**	0.97	0.94	1.00		
lnp**p**	0.97	0.93	0.94	1.00	
lnp**oth**		0.67	0.72	0.60	1.00

This model can be used to estimate two quantities of interest for hospitals with any set of characteristics. The first is the department-specific average unit cost ratio and the second is department-specific average cost per bed day and per outpatient visit.

Table 7. 4 shows the average unit cost ratios for the four main inpatient departments, which varied between 3.14:1 to 4.64:1. These are based on the expected proportions calculated from the model (see Eq. (7)) and the observed values of the explanatory variables from the data set. All department-specific unit cost ratios are significantly different to each other except the ratios of the unit cost per bed-day at the obstetrics/gynaecology and surgical departments to the unit cost per outpatient visit (paired t test, $p<0.0001$).

The fact that the coefficients in the equations of Table 7. 2 are statistically significant does not necessarily imply that each explanatory variable substantially effects the ratio of inpatient to outpatient unit costs. These effects are illustrated in Figure 7. 1 for the two explanatory variables with the smallest absolute value of the coefficients, occupancy rate and number of staff. Except in the case of occupancy rate on the internal medicine department, the effect on the cost ratio of varying the explanatory variables can, in fact, be substantial.

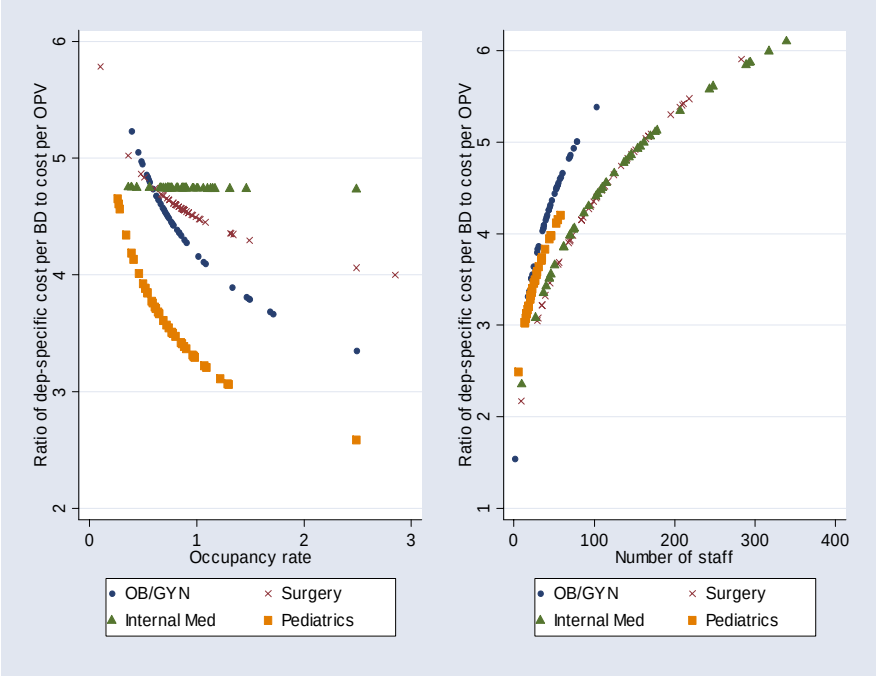
Occupancy rate has a non-linear negative relationship to the ratio of inpatient to outpatient unit costs (except for the internal medicine department). For example, the variation in the occupancy rate observed across the sample causes the ratio to vary from 4:1 to 6:1 for the surgery department. The ratio for the pediatrics department varies between 3:1 and 5:1. Changing the number of staff has the opposite effect on all four departments where a higher number of staff increases the unit cost ratio – from a minimum of 2:1 to a maximum of 6:1 for the surgery department, for example.

Table 7. 4 Ratio of department-specific cost per bed-day to the cost per outpatient visit Observations = 41

Variable	Mean	Std. Err.	[95% Conf.	Interval]
ratio_**go**	4.00	0.28	3.43	4.57
ratio_**mo**	4.64	0.33	3.98	5.31
ratio_**so**	4.12	0.29	3.53	4.70
ratio_**po**	3.14	0.24	5.33	7.45

*ratio_go = ratio of unit cost of Ob/Gyn to outpatient department, mo = Internal medicine to outpatient, so = surgery to outpatient, po = pediatric to outpatient.

Figure 7. 1 The relationship between the department-specific ratio of cost per bed day to cost per outpatient visit (Y axis) on department-specific determinants (X axis).



BD: bed day, OPV: outpatient visit, OB/GYN. Obstetrics and gynaecology.

Finally, Table 7. 5 presents the second quantity of interest that can be calculated from the model, the department-specific cost per bed-day or visit. The average unit costs are estimated as shown in Eq. (8) using the observed

values of explanatory variables in the data set. The actual estimates from the detailed step-down costing procedure are also presented for comparative purposes. The way this information can be used is discussed in the next section.

Table 7. 5 Model estimates of department-specific costs per bed-day in 2000 Chinese yuan (CY) (US$ in parenthesis: $1 = 8.3CY) Observations = 41

Variable*	Mean (observed)	Mean (Model)	Std. Err. (Model)	[95% Conf. Interval] (Model)	
uc**g**	259.88 (31)	306.41 (37)	28.89	248.02	364.80
uc**m**	300.79 (36)	355.50 (43)	34.51	285.74	425.26
uc**s**	199.32 (32)	237.20 (38)	22.34	192.05	282.34
uc**p**	269.04 (24)	313.34 (29)	28.70	255.34	371.34
uc**opv**	80.29 (10)	78.18 (9)	4.95	68.174	88.18

*uc = cost per bed-day, g= Ob/Gyn, m= internal medicine, s= surgery, p= pediatric, opv = outpatient visit

7. 4 Conclusion

In this paper we have described a model to examine the determinants of department-specific ratios of inpatient to outpatient costs. We then showed how the model could be used to estimate the cost of an inpatient bed-day or an outpatient visit for a hospital with any given set of characteristics. The statistical basis relies on the adaptation of models for compositional data that were previously developed for the estimation of demand shares and electoral votes, among others. Compositional models take account of the key features of this type of data, namely that the share attributable to each fraction is bounded by zero and one, and that all of the fractions must sum to one. As demonstrated by Smith, using seemingly unrelated regressions leads to significant improvement in efficiency of the estimation of compositional data through modelling, rather than ignoring, the correlated errors between equations. This is most valuable when most of the independent variables vary across equations, as in this case.[193]

The results showed that the ratio of inpatient to outpatient unit costs in this sample of Chinese hospitals is, on average, close to the simple rules of thumb (e.g. 3:1 and 4:1) suggested by earlier studies. However, it varies substantially by type of inpatient department according to a number of other explanatory variables, so that it is not uncommon for the ratio to be 5:1 or even 6:1. Much of the variation was due to differences in the number of staff and the

proportion of department-specific bed-days to total bed-days, although some variability was also due to differences in occupancy rates.

All cost ratios from Henan province were significantly lower than the proportions estimated from the other 11 provinces in the data set. Consultation with the Henan Provincial Department of Public Health clarified that during the period 1988 to 1998, the global public health policy of the provincial government and the ministry of health declared investment on the infrastructure (e.g., building) of public hospitals as a priority. In China, public hospitals are usually designed to consist of two main buildings, one for outpatient services and the second for inpatients. Each building includes general and specialty clinics and wards, but the outpatient building also includes the laboratory, radiology and imaging (e.g., Computerized Tomography, Magnetic Resonance Imaging) services. As a result, the outpatient building absorbed most of the capital investment during this 10-year period. It is not surprising, therefore, to find lower cost ratios (due to higher costs of outpatient departments, the denominator of the ratio) in Henan compared with the other provinces.

The most accurate way of developing cost estimates is to undertake a full step-down costing study. Where this information does not exist for a given hospital, the analysis in this paper suggests that equations such as those reported above could be used to estimate unit costs if explanatory variables such as occupancy rate, staff and the number of bed days are known. Uncertainty intervals can also be estimated. This information can then be used for planning, monitoring and forecasting purposes such as in the allocation of budgets between departments. The model can also be used in analyzing the expected changes in total costs or cost per bed-day if certain parameters were varied, e.g., the impact of a reduction in the number of staff or a higher occupancy rate.

The results also suggest that caution should be taken when using the result of a single hospital costing study as the basis of subsequent cost-effectiveness studies or in estimates of the costs of increasing coverage of particular interventions.[202] These studies might well use unit costs that are unrepresentative of the country as a whole if they do not assess whether the levels of capacity utilization, staff and bed days, shown here to influence hospital costs, are representative.[180]

This work is the first to attempt to explain the determinants of variation of hospital costs across departments, applied to Chinese hospitals. Further innovations may be possible in future development of this work, including development of compositional models designed specifically for panel data or using a multi-country analysis to explain variation in determinants of

department-specific hospital costs across countries. If there is a standard relationship of the determinants of hospital costs across countries, this would be particularly valuable for cost-effectiveness studies or national scaling up estimates in countries where full step-down costing studies do not exist.[4;203] Although it would be preferable for analysts to estimate department-specific costs using the step-down procedure if time and financial resources permit, this paper shows that econometric analysis of existing data can provide useful estimates in the interim.

Annex 7. 1 List of provinces included in the data-set

Province	Frequency	Percent	Cumulative percent
Beijing	1	2.44	2.44
Fujian	1	2.44	4.88
Gansu	1	2.44	7.32
Guangdong	3	7.32	14.63
Guizhou	2	4.88	19.51
Hebei	3	7.32	26.83
Henan	17	41.46	68.29
Hubei	3	7.32	75.61
Jiangxi	3	7.32	82.93
Shandong	1	2.44	85.37
Sichuan	2	4.88	90.24
Zhejiang	4	9.76	100.00
Total	41	100	

8

Capacity Constraints to the Adoption of New Interventions:

Consultation Time and the Integrated Management of Childhood Illness in Brazil

Based on: Adam T, Amorim DG, Edwards SJ, Amaral JA and Evans DB. Capacity Constraints to the Adoption of New Interventions: Consultation Time and the Integrated Management of Childhood Illness in Brazil. *Health Policy and Planning 2005; 20 (Suppl. 1)* ©. Published with permission from Oxford University Press.

CHAPTER 8

Chapter 8. Capacity constraints to the adoption of new interventions: Consultation time and the Integrated Management of Childhood Illness in Brazil

---

Summary

Information on how health workers spend their time can help programme managers determine whether it is possible to add new services or activities to their schedules and at what cost. One intervention with the potential to reduce under-five mortality is the Integrated Management of Childhood Illness (IMCI). Although it has been shown that IMCI is associated with improved quality of care, it is important to determine if it also requires additional consultation time. To investigate the amount of time required to provide clinical care to children under-five based on IMCI compared with routine care, a time and motion study was conducted in Northeast Brazil.

IMCI-trained providers spent 1 minute and 26 seconds longer per consultation with under-fives than untrained providers, holding confounding factors constant. The difference was greater when patient load was low, and decreased as the number of patients a provider saw per day increased. This has three implications. Firstly, the ability of the system to absorb new technologies depends on current capacity utilization. Second, the cost of treating a child also depends on the level of capacity utilization, at least in terms of provider time. Thirdly, where patient loads are high it is important to determine if the quality of care required for IMCI can be maintained.

CHAPTER 8

8. 1 Introduction

The decision to provide a new intervention, or to modify an old one, poses important questions regarding the resources required, one of which is staff time. At the margin, if health workers are currently fully occupied, it would not be possible to incorporate new activities that require additional time inputs unless new staff were employed or existing activities were eliminated or reduced. Information on how health workers currently spend their time can help programme managers determine whether it is possible to add new services within existing capacity constraints on health worker time.

The World Health Organization has estimated that almost 50% of global childhood deaths are due to pneumonia, diarrhoea, measles or malaria, in combination with malnutrition, all of which are preventable or treatable.[204] One of the interventions that has been shown to improve the quality of child care and has the potential to reduce under-five mortality is the Integrated Management of Childhood Illness (IMCI).[70;73;91;205;206] The strategy includes three components: improving case-management skills of health workers, improving health system support and improving family and community practices.[68;207] It started to be introduced in Brazil in 1996 and is moving ahead in several states, particularly in the northeast and northern regions that exhibit the country's poorest socioeconomic and health indicators.[208] The strategy adopted by Brazil focused on care at the level of the primary health facility, and no community component specific to IMCI has yet been implemented.[209]

At the time IMCI was developed, it was expected to have a positive impact on health outcomes,[68] but it was also expected to be more costly than routine care partly because it involves more clinical tasks which would take more time for a provider to deliver.[75] To investigate the amount of time required to provide clinical care to under-fives based on IMCI in a setting where it has been incorporated into primary health care facilities, as opposed to a trial setting, a time and motion study was conducted in Northeast Brazil, one of five sites participating in the Multi-Country Evaluation of the effectiveness, cost and impact of IMCI (MCE). See www.who.int/imci-mce for more details on IMCI and the MCE. In order to isolate the effect of IMCI from other causal factors, the study also sought to identify other determinants of variations in consultation time across providers.

IMCI in Brazil has been implemented in the context of a Family Health Program (FHP), supported by the World Bank and the Ministry of Health (MoH). The FHP teams are based in first-level government facilities (known as Family Health Program facilities) and a given facility may have one or more

family health teams depending on the size of its catchment area. Each FHP team includes a family physician, a registered nurse, two health auxiliaries, and 4-6 community health workers (CHW). For more information about IMCI implementation in Brazil see Amaral et al 2004.[205]

In this paper we report the results of the time and motion study conducted early in 2003. The main objectives were to identify the length of time spent by providers in consultations with under-fives, to isolate the effect of IMCI training on the length of consultations from the impact of other possible determinants, and to determine if any additional time IMCI providers devoted to under-fives was related to constraints on the amount of time they had available. Other types of capacity constraints, such as the availability of consulting rooms or equipment, were not explored.

8. 2 Methods

8.2. 1 Methods of the time and motion study

Two methods have been used in time and motion studies in health services research: continuous observation CO, and work sampling WS.[5] In CO, an observer measures on a continuous basis the time consumed by the observed person in carrying out the different activities of the day. This method requires a constant physical presence of the observer with the person being observed. In the WS method, one or more than one health workers can be observed at a time. The observer records what each person is doing at a certain point of time either at fixed intervals - perhaps each 5 or 10 minutes - or on a random basis. Typically, an inference is made about the portion of overall work time spent on an activity, based on the percent of observations that relate to that activity. The observer does not need to follow the member of personnel but can be located at a presumably unobtrusive observation point.[6]

We used the CO method as it has the advantage of providing an exact estimate of the time spent with each patient encounter, which is the main purpose of this analysis. In addition, the WS methods is more appropriate when health workers are in a circumscribed area such as a hospital ward setting,[5] while in the primary health facilities we visited there were no obvious places where observer could have good visual access to most of the health workers at the same point of time.

One of the main limitations of time and motion studies is the bias introduced via the Hawthorne effect, where observed health workers, being conscious of

the fact they are being observed, may change their usual working patterns. This is an inevitable consequence of any time and motion study, perhaps more in the CO method due to the direct observation.[5] However, this does not change the fact that time and motion studies are considered to be the most accurate method of measuring staff time compared with the alternatives of personnel interviews, self administered time sheets and patient flow analysis methods.[7] In addition, it can be argued that busy health workers will not be able to change their normal practice patterns for a long period of time, so after the initial observation period they soon return to their normal working patterns. To allow for a possible short term Hawthorne effect, in the analysis we tested whether the results from the first day of observation differed significantly from those of subsequent days.

8.2. 2 Selection of the study facilities and providers

The larger evaluation of IMCI in Brazil (i.e. the MCE) has a mixed retrospective-prospective design, since IMCI was already well implemented in many municipalities at the time the study was designed. Four states in Northeast Brazil - Bahia, Ceará, Paraíba and Pernambuco - were included based on the fact that IMCI implementation was reported to be strong in selected municipalities in those states. Health workers would, therefore, be more likely to have developed long-term practice patterns based on IMCI. The initial IMCI training had lasted eight days in Bahia and Pernambuco and six days in the two other states. For more detail on sampling methods and the rationale see Amaral et al 2004.[205]

The time and motion study was performed in only three of the four states included in the MCE, Ceará, Paraíba and Pernambuco. Preliminary information from Bahia suggested that IMCI had not yet been widely enough implemented to warrant the comparison. Data were collected from a sub-sample of facilities included in the Health Facility Survey (HFS) conducted in 2002 for the evaluation of IMCI. The sample size was determined using the mean and standard deviation estimates of consultation times for IMCI and non-IMCI trained providers from the MCE study in Tanzania, the only other site which had undertaken such a study, with an alpha of 5% and acceptable error of 2 minutes. Data were collected from a total of 32 facilities, of which half were practising IMCI. In each state, facilities were selected at random from those included in the HFS, stratified by the availability or not of an IMCI-trained provider. At the time of data collection, the observer selected one provider at random from those available at the first day of observation and who reportedly examined under-five children. Where the selected provider

was absent on one of the subsequent days of observation, the observer was required to randomly select another provider, who would then be observed for the remaining days of observation.
Health facilities in Northeast Brazil have two modes of service delivery: *integrated*, where the health worker offers all types of health services every day of the working week; and *vertical*, where a different kind of health service is offered each day – for example, prenatal care on Monday; sick children on Tuesday; etc. Therefore, two methods of observation were used, one for vertical and another for integrated-service facilities.

In integrated-service facilities, data were collected for two consecutive days, selected at random. This was to ensure that some days where there might be high levels of utilization and some where there could be low levels, were included. The sample of integrated-service facilities included 16 facilities (8 IMCI and 8 comparison) with a total of 32 observation days.

In vertical-service facilities, data were collected for one working week (5 days) in each of the sampled facilities. Facilities were selected at random, and the order of data collection weeks was selected on the basis of logistic considerations. The one-week data collection period was to explore whether sick children sought care on non-children days and to explore variation in the use of staff time on the different days of the week, and the effect of this form of service delivery on the availability and duration of "down" time. The sample included 16 vertical-service facilities (8 IMCI and 8 comparison), with a total of 80 observation days. As described above, only the provider seeing patients on the day of the visit was observed. This resulted in 47 providers observed in all types of facilities, 34 physicians and 13 nurses. Only two providers had to be replaced due to absence on one of the observation days. The analysis reported here focuses on physician providers, the main providers of curative care, because the number of observations for nurses was too low to carry out the analysis.

8.2. 3 Data collection

The study took place in April-May 2003. The data collection tool was developed and pre-tested in English by the MCE team. A full field test of the survey instrument, which was translated into Portuguese, was carried out during the one week training of the survey teams. Three surveyors collected data, one in each state, to avoid problems of inter-rater variability within states. All of them were qualified nurses and had spent part of their working career in health facilities. Each surveyor recorded the time at which every activity, including all breaks and pauses, began and ended, and what activity was being carried out. Before the data collection started, the observer used a

standard, field tested, introduction to explain the purpose of the study to the person in charge of the facility and to the randomly selected provider. The introduction emphasized that the purpose of the observation was to record the usual types of activities performed by the health worker during a routine working day; that it was not concerned with the quality of work involved; that he/she was selected at random to represent their facilities. Finally, their consent to participating in the study was obtained.

8.2. 4 Quality control

Supervision visits were made to all three states at least twice during the two month data collection period to check that the surveyors were carrying out the observation and were filling in the forms correctly, and for solving data-related problems. The supervisors checked all the forms for completeness and consistency during the supervisory visits and when data were being entered. Through this mechanism we sought to reduce the inter-rater variation across states.

8.2. 5 Methods of Analysis

As described above, the aim of the time and motion study was to record the time health workers who examined under-fives spent on the different activities in primary health facilities. Following Bratt et al. (1999),[7] the main activities are classified into three categories: contact time, non-contact productive time, and non-productive time. Contact time, or time spent with patients (or healthy people for preventive services such as immunization) is further divided into time spent caring for over-fives and for under-fives.

This paper focuses on consultation time with under-five children. Consultation time was estimated in two ways. The first computes the average consultation time for IMCI and non-IMCI providers, by state, without controlling for any possible confounders. The second uses regression analysis to explore and control for other possible determinants of consultation time. STATA software was used for the analysis.[86]

8.2.5. 1 Unadjusted average consultation time

The average consultation time was estimated as the total minutes a provider spent with children under five divided by the number of children seen.

Because there were instances where more than one person presented for a given consultation or the health worker was interrupted to perform another activity, (e.g. to examine another emergency case or resolve urgent administrative issues), the average time was estimated separately for consultations with and without these factors. This information is used in the regressions.

8.2.5. 2 Regression analysis

8.2.5.2. 1 Model specification

The possible determinants of consultation time for which information could be collected are included in Table 8. 1. The choice of explanatory variables was partly related to economic theory and partly determined by the nature and purpose of the exercise. For example, the number of consultations per provider per day is used as a measure of workload or the constraints on provider time; the total hours a provider works per day is also a measure of time availability, in particular as an indicator of whether the provider works part or full time, something that might modify behaviour. We also controlled for factors such as morning versus afternoon shifts to explore if patterns of behaviour changed during the day; the sex of provider; where multiple people presented for a particular consultation; and whether the consultation was interrupted for any reason.

Table 8. 1 Description of variables explored in the regression analysis[i]

Variable name	Definition
IMCI	IMCI-trained =1, otherwise = 0
Single visit	Single or multiple persons presenting at the same visit: single visit =1, otherwise = 0
Interruption	Whether the consultation was interrupted by another consultation or activity: interrupted=1, otherwise = 0
Home visit	Consultation took place at the patient's home: home visit =1, otherwise = 0
Outreach visit	Indicates a consultation at an outreach (usually a small health post in remote areas) site, outreach =1, otherwise = 0
Day one	First day of observation (to assess a possible Hawthorne effect): Day one=1, otherwise=0
Morning-afternoon	Whether consultation occurred during the morning or the afternoon shift. Morning=1, afternoon = 0

Variable name	Definition
Facility type	Indicate facility size. 1=small facility, 0= health centre, serving more than one facility
Sex	Sex of provider: male =1, female = 0
Ceará	Observation is from Ceara state, Ceara=1, otherwise = 0
Paraíba	Indicate observation is from Paraíba state, Paraíba=1, otherwise= 0
Ln visits per provider per day	Natural log of consultations per provider per day
Ln hours per provider per day	Natural log of working hours per provider per day
Ceará_lnvisits	Joint effect of Ceará with natural log of consultations per provider per day
Paraíba_lnvisits	Joint effect of Paraíba with natural log of consultations per provider per day
Ceará_lnhours	Joint effect of Ceará with natural log of working hours per provider per day
Paraíba_lnhours	Joint effect of Paraíba with natural log of working hours per provider per day

[i] Note: Only those variables included in the final model are shown in the Results Section.

The inclusion of state-specific variables makes it possible to control for differences in state characteristics that might affect the behaviour of health providers, such as the duration of IMCI training, degree of financial and political support to child health services and frequency of supervision. It was not possible to obtain information on these variables specific to each facility and provider, so the state variable was used instead.

Finally, we explored the joint effect of key variables on consultation time.[201] For example, if the relationship between number of consultation per provider per day varies by state. Only those which were included in the final model are presented here.

Double log transformation was used to normalize the dependent variable and to linearize the regression model.[210] Log transformation has the added advantage that coefficients can be readily interpreted as elasticities.[211] Natural logs were used. Finally, robust estimation methods was used to control for

clustering associated with having multiple observations per provider. The functional form can be written as:

$$T_{u5} = \alpha_0 + \sum_{i=1}^{n} \alpha_i \mathrm{X}_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} \mathrm{X}_i \mathrm{X}_j \qquad n=13 \qquad (1)$$

where T_{u5} is the natural log of time per consultation with an under-five child (in minutes); α_0 is the intercept; α_i are the estimated parameters for the n explanatory variables X_i; $\alpha_{ij} (\alpha_{ij} = \alpha_{ji}, j \neq i)$ are the estimated parameters for the joint effect of selected X variables, as described in Table 8. 1. Only those variables included in the final model are shown in the Results Section.

8.2.5.2. 2 Model fit

Regression diagnostics were used to judge the goodness-of-fit of the model. They included the tolerance test for multi-collinearity, its reciprocal variance inflation factors and estimates of adjusted R square and F statistics of the regression model.

8. 3 Results

8.3. 1 Univariate analysis

Results of the univariate analysis are shown in Table 8. 2. IMCI-trained providers spent on average 7.27, 8.57 and 13.44 minutes in consultation with under-five children in Ceará, Paraíba and Pernambuco respectively. The range was much smaller in the comparison group where they spent 7.56, 6.16 and 6.29 respectively. Although IMCI-trained providers spent less time per consultation on average than the comparison group in Ceará, the difference is small and not statistically significant. In addition, these results did not hold after controlling for confounders, as described below. In the two other states, Paraíba and Pernambuco, IMCI-trained providers spent significantly longer than their counterparts in the comparison group, the difference being greatest in Pernambuco at almost 7 minutes per consultation. Taken together, and without adjusting for confounders, IMCI-trained providers spent almost 4 more minutes per average consultation than providers in the comparison group ($p<0.0001$).

These results raise the question of whether IMCI-trained providers in Paraíba and Pernambuco compensated for spending more time with under-fives by

spending less time with over-fives. Table 8. 2 shows that this was not the case. IMCI-trained providers in those states also spent significantly more time with patients over five years than their counterparts who were not trained in IMCI.

Table 8. 2 Average time spent (in minutes) by physicians in consultations at health facilities —presented separately for under-five and over-five years of age.

	Under-five			Over-five		
	IMCI	Compa-rison	P (t)	IMCI	Compa-rison	P (t)
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
Ceará	7.27 (3.20)	7.56 (3.52)	0.53 (0.62)	5.58 (3.23)	8.34 (6.34)	<0.0001 (5.43)
Paraíba	8.57 (5.01)	6.16 (3.36)	0.002 (-3.11)	7.17 (5.14)	5.35 (3.13)	<0.0001 (-4.10)
Pernambuco	13.44 (6.22)	6.29 (3.21)	<0.0001 (-11.27)	13.12 (7.29)	8.03 (4.36)	<0.0001 (-8.46)
Average	11.08 (6.15)	6.46 (3.34)	<0.0001 (-9.04)	7.55 (5.51)	7.56 (5.09)	0.12 (-1.56)

This was not found in Ceará, however, where IMCI-trained providers spent significantly less time with over-fives compared with their counterparts (5.58 and 8.34 minutes respectively, $p<0.0001$). As in the case of the time spent in consultation with under-fives, it would be important to determine if this is due to IMCI or other possible determinants of consultation time. The purpose of this paper is to focus on children under-five, so that analysis is not undertaken here.

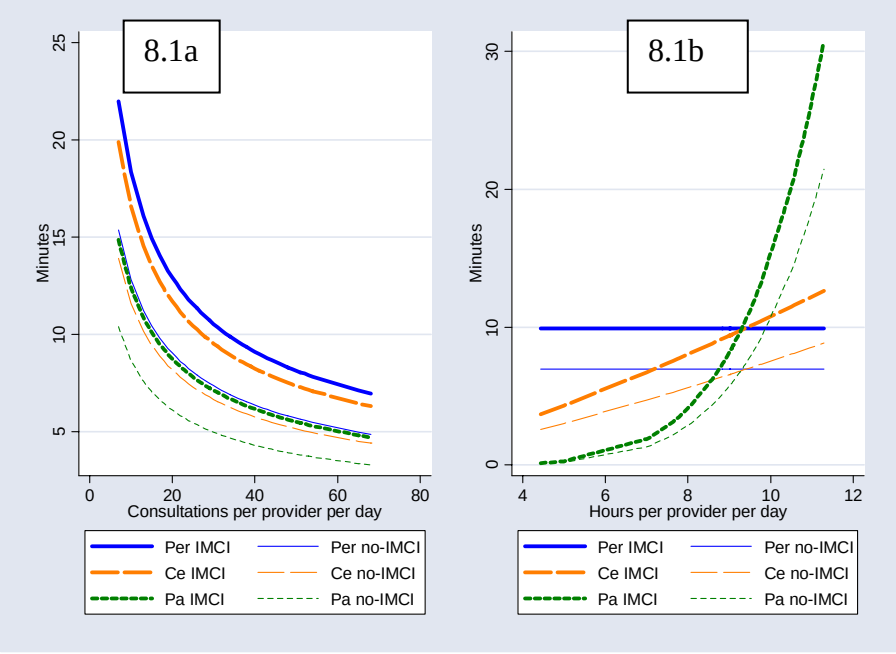
8.3. 2 Regression analysis

The results of the best fit model to explore whether consultation time remains correlated with the presence of IMCI even after controlling for potential confounders are presented in Table 8. 3, see Table 8. 1 for description of variable names. The adjusted R-square was 0.32, with an F statistic of 38.92 ($p<0.0001$).. The tolerance fell between 0.43 and 0.97 and the mean variance

inflation factors (VIF) was 1.48, so no evidence of multi-collinearity was detected— tolerance less than 0.05 and VIF more than 20 indicate presence of multi-collinearity. Residual plots showed a uniform distribution of the model residuals with a mean of zero and no specific pattern of distribution.

The positive signs of the coefficient of the variable denoting IMCI-trained providers and interrupted consultations; and the negative signs of the coefficients of number of consultations per provider per day and single visit are consistent with *a priori* expectations. The results show that, holding all other variables constant, IMCI trained providers spent 1 minute and 26 seconds more on average in consultation with under-fives than untrained providers ($p<0.0001$). Consultation time was negatively correlated with the number of consultations per provider per day — each 1 % increase in the number of consultations results in a 0.50% decrease in consultation time per under-five child ($p<0.0001$), see Table 8. 3 and Figure 8. 1a.

Figure 8. 1 Effect of (a) number of consultations per provider per day or (b) hours worked per day (X axes) on consultation time with under-fives (Y axis), estimated separately for each state and for IMCI-trained and non-trained providers.



Per=Pernambuco, Ce=Ceará and Pa=Paraíba

It is worth noting that the number of consultations per provider per day may be mediated by IMCI, i.e., people may choose to use IMCI providers due to perceived higher quality associated with IMCI training. This did not seem to be the case in our sample, as the mean consultations per provider per day was similar in facilities with and without an IMCI-trained provider (31 and 36 per day respectively, p=0.09). Accordingly, we treat this variable as a potential exogenous confounder rather than endogenous to the model.

Table 8. 3 Results of the multivariate regression model (dependent variable: natural log of time (minutes) per consultation with under-five child).

Adjusted $R^2 = 0.32$ F-statistic = 38.92 p <0.0001 N = 730

Variable	$\overline{\beta}$ coef.	S E	t	P	95%confidence-Interval	
					Low	high
IMCI	0.36	0.04	9.11	<0.0001	0.28	0.43
Interruption	0.14	0.08	1.74	0.081	-0.02	0.30
Single visit	-0.40	0.05	-8.65	<0.0001	-0.49	-0.31
Ln visits per provider per day	-0.50	0.05	-10.34	<0.0001	-0.60	-0.41
Ceará	-2.96	0.63	-4.69	<0.0001	-4.20	-1.72
Paraíba	-13.04	4.81	-2.71	0.007	-22.48	-3.59
Ceará_lnhours	1.32	0.28	4.69	<0.0001	0.77	1.88
Paraíba_lnhours	5.84	2.18	2.68	0.008	1.56	10.13
sex	-0.09	0.04	-2.27	0.024	-0.17	-0.01
Constant	4.12	0.19	22.08	<0.0001	3.75	4.48

Consultation time was lower in Ceará and Paraíba than in Pernambuco, for both IMCI-trained and comparison providers. It is also worth noting that controlling for confounding factors showed that IMCI-trained providers spent longer per consultation with under-fives than the comparison group in Ceará as well as in the other states. This means that the findings of the univariate analysis presented in Table 8. 2, that IMCI providers seemed to spend less time than the comparison group, were due to confounding factors and not IMCI.

The effect of the number of consultations per provider per day did not vary by state, but the number of working hours per day did. This is illustrated in Figure 8. 1b. The lines show the estimated consultation time based on the reported regression, controlling for all variables except hours worked per day, shown on the X axis in natural units. Consultation time (Y axis) was converted into natural units (minutes) using the antilog of the dependent variable, and was estimated separately for IMCI-trained and comparison groups and for each of the three states. It was estimated for a single visit, uninterrupted, and performed at the health facility (e.g. not a home or outreach visit). The number of visits per provider per day was set at the mean value from the sample (34 per day).

The Figure shows that working more hours per day did not have any impact on consultation time in Pernambuco but did in Ceará and Paraíba (confirmed by the fact that hours per day was not significant in the final model while the joint effect of hours per day and the state variables was significant for Ceará and Paraíba). In those states, each 1% increase in hours worked per day per provider led to 1.32% and 5.84% increase in consultation time per under-five child respectively (Table 8. 3).

Finally, the results show that male providers spent 55 seconds less per consultation with under-fives than female providers. No statistically significant differences were found in consultation time between vertical versus integrated-service facilities, for "under-five" days versus "other" days within vertical-service facilities, for day one versus subsequent days of observation, or for home and outreach visits compared with consultations which took place at health facilities.

8. 4 Discussion

The main purposes of this study were to explore the effect of IMCI training on the length of consultation time with under-five children, and to examine how providers cope with any capacity constraints on their time. The multivariate analysis confirms that IMCI-trained providers spent more time on the average consultation than non-IMCI-trained providers. After controlling for other determinants, the difference was 1 minute and 26 seconds per consultation ($p<0.0001$), approximately 20% higher than the average consultation time of non-IMCI trained providers. This is substantially less than the difference of 4 minutes suggested by the univariate analysis and emphasizes the need to control for confounders in studies such as this.

The size of the difference is, however, crucially determined by constraints on providers' time. The higher the workload, the lower the difference, to the extent that the difference is relatively small where the workload exceeds 50 consultations per provider per day. This has key policy and methodological implications. In terms of policy, the average quality of care of IMCI providers has been shown to be higher than that of non-IMCI providers.[70;73;91;205] It is important to determine if quality is also a function of the time spent per consultation, and if the quality of IMCI providers still exceeds that of non-IMCI providers where workload is high.

There is already some evidence that quality might well be a function of the time spent per consultation. The average time per consultation in Pernambuco was higher than that in the two other states. At the same time, the component of the MCE study exploring quality of care found that IMCI providers in Pernambuco consistently made a higher percentage of correct disease classifications than IMCI providers in the other states. Significant differences between health workers who were trained or not trained in IMCI were found in the assessment of the sick child, classification of illness, treatment, as well as in communication with the caretaker.[205] Whether this is due to more frequent supervision and more regular availability of drugs and vaccines in Pernambuco,[205] to the fact that Pernambuco implemented IMCI earlier than the other states, to the strong support for IMCI from the State Health Secretariat in Pernambuco, or to the higher time per consultation is yet to be determined. It is important that the answer is found rapidly so as to design strategies to support providers facing capacity constraints on their time, if necessary, in all parts of Brazil where IMCI is being introduced.

Methodological implications relate to the way costs are estimated for the purposes of cost-effectiveness analysis, for estimating the costs of benefits packages for health insurance, or for budgeting for the scale up of interventions.[180;188] As illustrated in this study, the cost per visit depends on the throughput of the health facility. Where patient load is high, unit costs are relatively low because providers must spend less time per patient, and vice versa. Moreover, the incremental cost of expanding is always less than the average cost of providing existing services, as expected from theory, and falls more rapidly where patient load is low. General policy implications should not be based on the results of costing studies that do not report capacity utilization and studies of scale-up costs will not be useful to policy-makers if they are based on the current costs of providing care.

This result is also important for policy relating to child health. Where current case loads are relatively low, providers spend additional time to provide child

health services based on IMCI as part of their current activities. In our sample, the mean number of consultations per provider per day was 34 and 95% of the providers had case loads lower than 54 patients per day. If this is representative of the rest of Brazil, it would be possible to introduce IMCI relatively easily throughout the country without encountering capacity constraints in terms of provider time. Interestingly, an unthinking cost analysis would suggest the opposite. Because the observed additional time is greater in areas of low workload, it would suggest that the incremental financial costs of IMCI are higher where workload is low than where workload is high. We believe the results can best be interpreted as showing that IMCI can be introduced without significant financial implications when capacity constraints on time are relatively low.

It is interesting to find no evidence that the time spent per child in vertical-service facilities differed from that in integrated facilities, or that time spent per child on vertical days differed to the time spent on other days in vertical-facilities. Indeed, we could find no evidence that the mix of patients varied by day in vertical-service facilities which is a relief from the medical perspective - it implies that patients in need of urgent care could obtain it on any day of the working week. This is useful information for future research in this region of Brazil, where it might not be necessary to account for the two different types of facilities in the design of future research.

Another interesting finding is that there are sex differences in average consultation time by providers. It would be useful to explore the possible reasons for this in future studies. Of particular interest are questions such as the effect and suitability of IMCI-training materials to both sexes, and if the difference in time is also associated with difference between the sexes in quality of care.

Finally, the fact that the times reported here are consistent with those observed in the MCE study in Tanzania, the only available comparative study, gives us some confidence in the results.[212] The average consultation time with under-fives in Paraíba and Ceará were very close to the 8.2 minutes observed in Tanzania, and only the consultation time in Pernambuco was somewhat longer. In addition, Tanzanian workers spent longer on average with children under-five than with older people, something also observed in general in Brazil. The times are lower than those reported in an earlier study from Bangladesh,[75] where IMCI providers spent on average 16 minutes per consultation with under-fives. That study did not try to determine how long they would have spent had they not been trained in IMCI. It is not strictly comparable to the Brazilian study as it was undertaken immediately after training, while the Brazilian study was undertaken in settings some years after

the introduction of IMCI when the immediate effects of training are likely to have waned.

In conclusion, the critical implication of this study is that at relatively low patient load, adoption of new interventions that improve quality can promote efficiency at primary care facilities - they can be delivered by the already available human resource capacity of the health system without the need to hire additional staff or to reduce other activities. At high patient load, IMCI-trained providers do not spend as much additional time with children as providers with low patient loads. It is important to determine if the quality of IMCI-trained providers is maintained in those settings. This is an urgent need as countries move to scale with their delivery of child health services based on IMCI.

9

Achieving the WHO/UNAIDS Antiretroviral Treatment "3 by 5" Goal:

What Will it Cost?

Reprinted from the Lancet: Gutierrez JP, Johns B, Adam T, Bertozzi S , Tan-Torres Edejer T, Greener R, Hankins C and Evans DB. Achieving the WHO/UNAIDS antiretroviral treatment "3 by 5" goal: what will it cost? 2004;(364):63-4 ©, with permission from the Lancet.

CHAPTER 9

Chapter 9. Achieving the WHO/UNAIDS antiretroviral treatment "3 by 5" goal: What will it cost?

Summary

The "3 by 5" goal to have 3 million people in low and middle income countries on antiretroviral therapy (ART) by the end of 2005 is ambitious. Estimates of the necessary resources are needed to facilitate resource mobilisation and rapid channelling of funds to where they are required. We estimated the financial costs needed to implement treatment protocols, by use of country-specific estimates for 34 countries that account for 90% of the need for ART in resource-poor settings. We first estimated the number of people needing ART and supporting programmes for each country. We then estimated the cost per patient for each programme by country to derive total costs. We estimate that between US$5.1 billion and US$5.9 billion will be needed by the end of 2005 to provide ART, support programmes, and cover country-level administrative and logistic costs for 3 by 5.

In September 2003, at the second UN General Assembly Special Session on HIV/AIDS, the World Health Organization (WHO) and the Joint United Nations Programme on HIV/AIDS (UNAIDS) declared the lack of treatment in low- and middle-income countries a global public health emergency and launched the '3 by 5' Initiative aimed at enrolling 3 million people on antiretroviral therapy (ART) by the end of 2005. For this ambitious goal to be reached, starting from a base of fewer than 200,000 patients on treatment, countries, donors, and multilateral agencies must know what resources need to be rapidly mobilised.

Since previous estimates of the cost of scaling up interventions against HIV and AIDS were made[202] new WHO recommended treatment protocols for resource limited settings have been published.[213] We estimate the financial

costs to implement these protocols using country-specific estimates for 34 countries that account for 90% of the need for ART in resource poor settings. The '3 by 5' strategy includes standardised treatment protocols, simplified clinical monitoring and record keeping, optimal use of existing health system resources, active involvement of communities and people living with HIV, plus efforts to minimise the cost of drugs and diagnostics.[213;214] No major changes to the health system infrastructure, the quantity of available personnel, or in transmission as a result of ART are deemed possible given the short time frame.

The number of people needing treatment is defined as those expected to die within two years in the absence of ART. People needing treatment in 2004, for example, are those expected to die before the end of 2006 without treatment. People who stop treatment or die while on treatment are replaced so that 3 million people are alive and receiving ART at the end of 2005 *(see web document at lancet.com for more details).*

The three main entry points for recruitment of eligible patients are: tuberculosis clinics, health facilities (in-patient and out-patient) and mother to child transmission prevention programmes in antenatal care clinics. Drug regimens and testing procedures vary by entry point *(see web document at Lamvet.com).* Two assumptions of the growth rates in coverage to reach the target are employed. Slower scale-up reaches 10% of the target in 2004 and 90% in 2005. More rapid scale-up assumes 20% of the target met in 2004.

Patient enrolment requires confirmation of positive HIV status by rapid testing with counselling. A doctor or nurse confirms clinical eligibility. To stabilise patients starting ART and ensure continued well-being while on therapy, patients are diagnosed and treated for opportunistic infections (OIs). Medicines for prevention of OIs and laboratory tests for suspected toxicity help ensure successful ART, while those failing treatment require palliative care.

Support costs at country level include training of existing health personnel, supervision of ART delivery, and volunteers (remunerated) providing adherence support to patients. Universal precautions and post-exposure prophylaxis are included, as are costs for limited upgrades to laboratories and drug storage and distribution systems.

The number of facility-based visits for ART initiation and monitoring, frequency of adherence counselling and monitoring, laboratory tests for toxicity and estimated times associated with each of these activities are based on reports from a consensus meeting.[213] Quantities of inputs used in other

interventions are based on recommendations in published guidelines supplemented by expert opinion.

Unit cost data were obtained from studies conducted in sub-Saharan Africa or Asia or, where no data existed, extrapolated using methods described elsewhere.[188;215] Details of assumptions and the unit costs by intervention and activity for Botswana are in the *web document.*

Two assumptions for the cost of drugs are used. The high cost option assumes:[216;217] $304 for first line therapy (fixed-dose combination); $1108 for second line therapy; $706 for a switch due to toxicity; $505 for patients with TB ($353 for 6 months then first line for the remaining 6 months); $831 for pregnant patients with TB ($679 for the 6 month switch). The low cost option uses $140 per patient per year, the price negotiated by the Clinton Foundation for standard first-line treatment in selected countries and which may become more generally available. Drug costs for other categories are assumed to undergo a similar cost reduction.

The estimated costs for 2004-2005 are summarised in the Table 9. 1 for four combinations of assumptions of the rate of scale-up and pharmaceutical costs. The total cost is between US$ 5·1 and 5·9 billion for the two years. The mid-point of US$ 5·5 billion dollars is used for advocacy purposes.

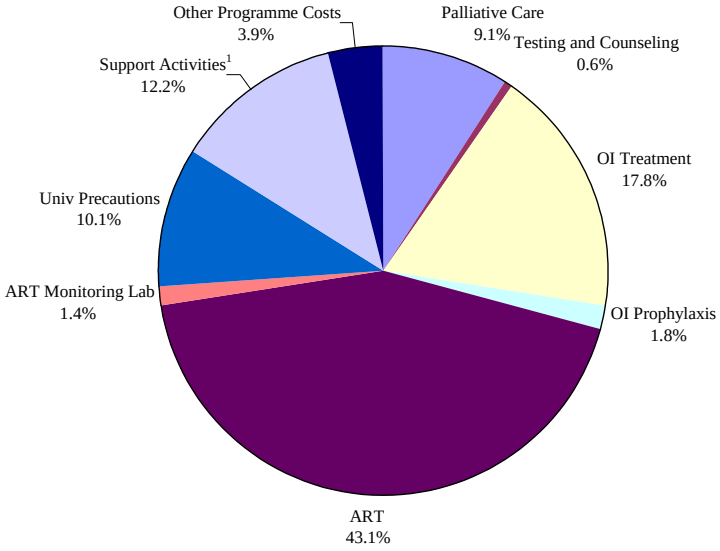
Figure 9. 1 shows the breakdown of total two-year costs by category for scenario 2A of Table 9. 1. Patient costs account for over 77% of the total costs, with purchase and provision of ART responsible for over 43%. Universal precautions, post-exposure prophylaxis, and other programme costs account for less than 23%.

Table 9. 1 Estimated cost of "3 by 5", 2004-2005 (US$ billions)

Scenario:	2004	2005	TOTAL
1A: 10%/90% with higher drug costs*	$2·0	$3·8	$5·7
1B: 10%/90% with lower drug costs	$1·9	$3·2	$5·1
2A: 20%/80% with higher drug costs	$2·2	$3·7	$5·9
2B: 20%/80% with lower drug costs	$2·0	$3·1	$5·1

10%/90% refers to 10% coverage in 2004 and 90% in 2005

Figure 9. 1 Distribution of total costs of "3 by 5" for 2004-2005 (Scenario 2A)



[1]Support Activities include post-exposure prophylaxis (PEP), and scale-up of preventing mother to child transmission of HIV (PMTCT)

The primary purpose of these estimates is to provide a global aggregate target to guide short run resource mobilization. Precise country estimates are being developed to help local planning. The figures presented here show how critical it is to intensify resource mobilization activities. This is even more important for the longer run because the need for additional resources will continue to grow after 2005 as more patients are recruited, adding to surviving cohorts of patients.

10

Achieving the Millennium Development Goals for health:

Methods to Assess the Costs and Effects of Interventions for Improving Health in Developing Countries

Based on: Evans DB, Tan-Torres Edejer T, Adam T, Lim SS, and the WHO-CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Methods to Assess the Costs and Effects of Interventions for Improving Health in Developing Countries. BMJ 2005 (In press) ©. Published with permission from BMJ.

CHAPTER 10

Chapter 10. Achieving the Millennium Development Goals for health: Methods to assess the costs and effects of interventions for improving health in developing countries

Summary

Five years after signing of the Millennium Declaration, the United Nations is reviewing progress towards achievement of the Millennium Development Goals (MDGs). Three are exclusive to health, focusing on maternal and child health, HIV/AIDS, tuberculosis and malaria. This is the first in a series of papers considering whether existing strategies targeting these diseases and conditions are the most efficient way of spending resources, and asking what should be done if additional resources become available.

Sandardized analytical methods developed for cost-effectiveness analysis in the WHO-CHOICE project are used. They allow the efficiency of the current use of resources to be assessed at the same time as indicating which interventions should be given priority when new resources become available. The papers also account for interactions between interventions undertaken simultaneously, in terms of costs and effects. This is rarely done in cost-effectiveness analysis despite the fact that interventions are seldom done in isolation.

This paper provides a detailed description of the methods to assess the costs and effects of heath interventions in this series and how the results could lead to significant improvements in population health.

10. 1 Introduction

The methods chosen to assess the cost-effectiveness of interventions designed to achieve the health MDGs must be capable of revealing the efficiency of current and possible new resource use, incorporating interactions between interventions undertaken at the same time, and incorporating the impact of expanding coverage on unit costs.[218] They should also allow valid comparisons across a wide range of interventions. Here we describe how the standardized cost-effectiveness methods used in this series have addressed these issues.[219-223]

10. 2 Methods

10.2. 1 Level of the analysis

The project divided countries into 14 regions based on geographical proximity and rates of child and adult mortality (see table A on bmj.com). Results for all regions are available at www.who.int/choice , but the papers in this series give details for just two regions: Afr-E, which includes countries in sub-Saharan Africa with high child mortality and very high adult mortality, and Sear-D, which comprises countries in South East Asia with high child and adult mortality.

10.2. 2 Definition and selection of interventions

The term intervention is defined to include any preventive, promotive, curative or rehabilitative action that improves health. They are analysed individually and then in combinations or packages that could be undertaken together (see Box 10. 1), taking into account interactions in costs and/or effectiveness.

Interventions were chosen for analysis either because they are commonly used or because disease control experts have advocated their introduction. In each case, some evidence was needed that the intervention could be effective. The list is not exhaustive, and excluding an intervention does not imply it is cost ineffective.

All interventions and combinations are assessed assuming they are implemented for 10 years starting in 2000, the year the Millennium Declaration was signed. Good policy making would then require a reassessment of

strategies. Costs, therefore, are only incurred over 10 years, but all benefits accruing because of actions taken between 2000 and 2010 are included.

Box 10. 1 Analysis of individual and combined interventions

Interventions are analysed individually and then in combinations that could be undertaken together. This requires a decision about whether interventions are independent (they can be done at the same time in a population, with or without interactions) or mutually exclusive (if one is chosen, another cannot be). Independent interventions can be added to existing interventions, while mutually exclusive interventions must replace an existing intervention.

Consider three individual antenatal interventions: tetanus immunisation, screening for pre-eclampsia, and screening and treatment for asymptomatic bacteruria. They are independent, so all possible combinations of the interventions would be:
Tetanus immunisation + screening for pre-eclampsia
Tetanus immunisation + screening and treatment for asymptomatic bacteruria
Screening for pre-eclampsia+ screening and treatment for asymptomatic bacteruria
Tetanus immunisation + screening for pre-eclampsia and asymptomatic bacteruria

All the scenarios are assessed for different coverage levels (50%, 80%, 95%), introducing the idea of mutual exclusivity. Coverage at 95% must replace coverage at 80%. Each of the combinations of interventions are analysed taking into account all possible interactions in costs or effectiveness. Combinations of interventions must be plausible. For example, emergency obstetric care interventions are not included unless skilled birth attendance is implemented at the same time. The number of possible combinations increases rapidly with the addition of each new intervention

10.2. 3 Intervention Costs

Costs are measured from the perspective of society as a whole, to understand how best to use resources regardless of who pays for them, or indeed, whether they are paid for at all. For example, village volunteers working for maternal health must be included because they could be working for some other health

programme if they did not work for maternal health. All resources used for each intervention or combination are included and valued.

Data on the costs incurred by people to access services (e.g. travel costs) are rarely available, and we, like most other studies, have excluded them.[224;225] Domestic taxes were also excluded. From the perspective of the society, they simply transfer financial resources from one person to another and do not use up a physical resource such as capital or labour. The impact of interventions on the time and potential earnings of patients and unpaid carers - i.e. work time lost - is a vexing question in cost-effectiveness analysis but, as here, they are often excluded on ethical grounds. Inclusion would give priority to extending the life of people who earn more (seeBox 10. 2).[226;227]

Box 10. 2 Cost-effectiveness analysis and economic production

Cost-effectiveness analysis (CEA) developed partly in reaction to ethical concerns about the implications of cost-benefit analysis (CBA). CBA values health improvements in money terms. If it is applied properly, it gives higher weight to health gains in people who earn more. CEA values each unit of health improvement equally regardless of the income of the recipient. Accordingly, production gains resulting from health interventions have not generally been included in CEA, although this is not by any means a universal practice.[2] We exclude them in this series for ethical reasons, arguing that health professionals and planners should seek to improve population health to the greatest extent possible for the available resources.[228] Questions of the impact of health actions on national income require interaction with other sectors of the economy. However, it is important to acknowledge that CEA provides only one, though important, part of the information set required to decide how best to allocate resources. In this series of papers, economic production concerns are most relevant to HIV/AIDS, where it has been argued that lack of action could even lead to the complete breakdown of societies.[229] This is discussed further in that paper.[222]

Costs are evaluated assuming capacity utilization (i.e. technical efficiency) is relatively high, though achievable - 80% for all interventions. This is to ensure that differences in cost-effectiveness ratios are due to fundamental differences in the technologies or strategies adopted, and not simply because an intervention has been implemented poorly in a dysfunctional health system.

Accordingly, the results provide practical information about the appropriate mix of interventions for various levels of resource availability, assuming interventions are undertaken relatively efficiently.

10.2.3. 1 Classification of costs

Costs are divided into those incurred at the patient or programme level. Patient-level costs involve face to face delivery by a health provider (broadly defined) to a recipient - e.g. medicines, outpatient visits, in-patient stays, individual health education messages. Programme-level costs include all resources required to establish and maintain an intervention - administration, publicity, training, delivery of supplies. Interventions like radio delivery of health education messages largely involve the former, while treatment at health centres largely involves the latter.

10.2.3. 2 Cost measurement

A standardized ingredients approach was used, requiring information on the quantities of physical inputs needed and their unit cost. Total costs are quantities of inputs multiplied by their unit costs.
For programme-level costs, the physical inputs - human resources, office space, vehicles, electricity, other services, and a variety of consumables - required to introduce and run a programme were collected in 17 countries by costing experts commissioned for this purpose, using a standard template.[188;215;230] This was supplemented by information from programme managers in other countries known by WHO staff.

For patient-level costs, quantities were taken from a variety of sources. Where effectiveness estimates were available from published studies, the resources necessary to ensure the observed level of effectiveness were identified. In other cases, the resources implied by the activities outlined in WHO treatment practice guidelines[231-234] were estimated. Programme experts checked the face validity of all estimated quantities. It was not always possible to identify the quantities of primary inputs (human resources, consumables) necessary for patient-level costs, so quantities and prices were estimated at an intermediate level for several inputs - inpatient days at different hospital levels, outpatient visits and health center visits.

Unit costs for each input were derived from an extensive search of published and unpublished literature and databases along with consultation with costing experts. The full costing data base is available at www.who.int/choice.

For goods that are traded internationally, the most competitive price available internationally was used. For example, the prices of medicines were taken

from the latest WHO negotiated prices. A mark up was included for transportation costs.
For goods available only locally (e.g. human resources, inpatient bed days) unit costs have been shown to vary substantially within countries. As a result, cross country regressions were run using the collected data to estimate the average cost (with adjustments for capacity utilization) for each setting.[188;215;230;235]

10.2.3. 3 Variations in scope and scale

Some interventions can share resources like vehicles, buildings, and administrative staff, so the costs of doing both together is less than the sum of the costs of the two individual programmes. Careful attention was paid to identifying possible shared costs when combinations of interventions were evaluated (see Box B on bmj.com).

Costs were estimated at three standard levels of coverage - 50%, 80% and 95%. We assumed that interventions are first provided to, or obtained by, easy-to-reach populations. Coverage then expands to more outlying, sparsely populated areas. Facilities still need to be built despite the lower population density, so capacity utilization will be lower, and the costs per patient treated correspondingly higher at higher levels of coverage. On the other hand, some interventions require a fixed number of staff, or office space, or equipment, to begin work. The same people and resources cope with increases in coverage so the programme-level costs can be spread over a larger population, reducing the costs per person covered. Both types of effects are included in the costing exercise, and the net effect varies by intervention.

Costs are reported in international (I$) rather than US dollars (US$) with 2000 the base year (see Box 10. 3). Future costs are discounted to 2000 values using a 3% discount rate.[2] These costs do not translate directly into the financial or cash requirements to run or expand interventions. Estimates of the cash requirements are already available.[4;187;236-238] Here we estimate the opportunity cost of all resources required to provide interventions, regardless of who pays for them, to explore the combination of interventions that makes the best use of these resources.

10.2. 4 Assessing the health impact of interventions

Interventions improve health through their effect on incidence, remission, severity and case-fatality. Efficacy data were obtained, in order of preference, from systematic reviews with meta-analysis; randomized studies; and before-and-after evaluations of country programs.

Effectiveness was obtained by adjusting efficacy by a factor between zero and one to allow for less than perfect adherence to recommended practices. The adjustment factor was taken from the literature wherever possible, with expert opinion the last resort. Finally, the eligible population was multiplied by coverage to determine the total numbers benefiting from estimated effectiveness. Although some information on the effect of increasing coverage on unit costs was available, this was not the case for health impacts. We assumed, therefore, that the health benefit of each child treated for pneumonia (i.e. unit effectiveness), for example, did not vary with coverage.

Box 10. 3 International dollars

International dollars are used to account for differences in price levels across countries. The exchange rate for domestic currency into international dollars is the amount of domestic currency required to purchase the same quantity of goods and services as $1 could purchase in the US. For low income countries, national income measured in terms of international dollars is generally higher than it is in US dollars because domestic costs for many goods are lower than in the US. Table B on bmj.com gives conversions from international to US dollars.

10.2.4. 1 The impact of interventions undertaken simultaneously

It was rarely possible to obtain information on the joint effectiveness of multiple interventions undertaken simultaneously. In its absence, the health impact was assumed to be additive for interventions that impact on different health outcomes. For interventions acting on the same outcome the joint effect was estimated by multiplying the individual relative risks (see Box C on bmj.com)

10.2.4. 2 Population health effects of interventions

A population model, PopMod, was used to project the impact of interventions on the aggregate healthy years of life lived by a population, combining incidence, prevalence and mortality rates, and estimates of disease severity, with information on intervention coverage and effectiveness.[239] The exception was for HIV/AIDS, where a model that allows death rates to increase with time spent ill was used. Regional epidemiology was taken from the latest internal WHO update of the Global Burden of Disease study.[240]

PopMod projects regional population over time, allowing people to move in and out of disease states, or die, in accordance with incidence and remission

rates. To derive a single indicator of population health, time spent in each state is given a weight (i.e. health-state valuation or disability weight) reflecting disease severity. Weights were taken from Murray and Lopez.[241] The population is projected for the length of time necessary for people affected by the intervention to live their full life course. The difference between the aggregate number of healthy years lived by the projected population in the intervention and in the do-nothing scenario is the population health gain attributable to the intervention.

The outcome indicator is essentially the number of healthy years of life gained by an intervention, also called the number of disability adjusted life years (DALYs) averted. The mechanics of estimating DALYs are virtually identical to those required to estimate quality adjusted life years (QALYs), another common outcome indicator, although there are some differences in the interpretation (see box D on bmj.com). In the base case analysis discounts DALYs averted in the future at 3% and gives greater weight to DALYs averted during adulthood. This is relaxed in the sensitivity analysis.

10.2. 5 Calculating cost-effectiveness

Traditional CEA generally considers the costs and effects of adding new interventions to current practice or of replacing one existing intervention with another targeting the same condition. Here we evaluate the full set of existing interventions by first considering what would happen to population health if they all ceased to be implemented today. This is the "null" or do-nothing scenario (see Box E on bmj.com). The current population is followed over time in PopMod assuming that all interventions cease, using the information on epidemiology, effectiveness and the current coverage of interventions described earlier.

Next, we trace the implications for population health of adding all possible interventions singly and in various combinations, against the baseline of doing nothing. The difference is the gain in health due to the reduction in disease burden from the intervention(s) (DALYs averted). The costs of each scenario are then compared with the gain to identify the most cost effective set of interventions at different levels of resource availability. The comparison of the current mix against the optimal set for the resources currently available reveals areas of inefficiency. The optimal set for higher resource levels shows what should be done if existing efforts to raise more resources are successful.

10.2. 6 Interpreting results

For each paper in this series, the cost of every intervention and package is plotted against the corresponding population health impact relative to the do-nothing scenario for that condition. We then determine the order in which interventions would be chosen and combined at given levels of resource availability if cost-effectiveness were the only consideration. The graphical depiction we call the *Expansion Path*. To understand this process, two cost-effectiveness ratios are defined. The *incremental cost-effectiveness ratio* (ICER) denotes the additional costs and benefits of a new intervention or package relative to what existed before - the previous point on the expansion path. The ICER compared to doing nothing is sometimes called the *average cost-effectiveness ratio* (ACER). Box F on bmj.com provides an example of how the expansion path is determined and interpreted.

10.2. 7 Uncertainty

Cost-effectiveness is an inexact science, particularly where data are limited and quality poor. Many sources of uncertainty cannot be captured by statistical confidence intervals because no sampling is involved. Probabilistic uncertainty analysis has been developed for these circumstances, but technical limitations do not yet allow it to include the complex correlations inherent in combining the very large set of interventions in this series (see box G on bmj.com).[242;243]

Accordingly, the individual papers incorporate relatively simple forms of sensitivity analysis that are easier to understand. In the final paper, the results of all the disease-specific analysis are integrated to develop priorities from the perspective of a policy-maker responsible for all health MDGs, not just one. For this cross-disease analysis, we believe it is not possible to recommend that an intervention shown to cost $45 per DALY averted is more efficient than one costing $60 given the nature of the uncertainties involved. However, we are much more confident that $45 per DALY is better than $450 per DALY. For the broad comparison across MDGs we, therefore, interpret cost-effectiveness figures in broad order of magnitude ranges (see box H on bmj.com).

10.2. 8 Documentation and transparency

Details of the specific data sources and assumptions made to obtain costs and effects are presented in the individual papers and are available from the WHO-CHOICE website (www.who.int/choice). Data inputs are also available in the form of a contextualization tool, currently being tested in a number of

countries, that allows analysts to adapt and modify the regional parameters to their own settings. They can modify prices, capacity utilization, epidemiology and effectiveness assumptions. This tool is available from the authors on request.

10. 3 Conclusions

Progress in the five years since the Millennium Declaration was signed has been disappointing.[218] More funds to improve health in poor countries is urgently required, but this alone will not be sufficient to achieve the MDGs. Open discussion of ways of increasing their impact is required, both to allow countries to improve the health of their populations more quickly, and to give potential donors confidence that new funds would be used effectively. The methods outlined in this paper are applied in the next six papers to facilitate open debate and encourage changes in strategies where necessary.

11

Achieving the Millennium Development Goals for Health:

Cost-effectiveness of Strategies for Maternal and Neonatal Health in Developing Countries

Based on: Adam T, Lim SS, Mehta SS, Bhutta ZA, Fogstad H, Mathai M, Zupan J and Darmstadt GL. Achieving the Mellinnium Development Goals for Health: Cost effectiveness of Strategies for Maternal and Neonatal Health in Developing Countries. BMJ 2005 (In press) ©. Published with permission from BMJ.

CHAPTER 11

Chapter 11. Achieving the Millennium Development Goals for health: Cost-effectiveness of strategies for maternal and neonatal health in developing countries

Summary

To determine the costs and benefits of interventions for maternal and newborn health in order to assess the appropriateness of current strategies and guide future plans to attain the Millennium Development Goals.

The standard methods of the WHO-CHOICE project were used. The analyses included 21 maternal and newborn health interventions and almost 300 combinations, taking into account joint costs and effects when they are implemented together. Interventions included in this analysis are those delivered during pregnancy, childbirth and the newborn period for which evidence of effectiveness exists. Results are presented as cost per disability adjusted life year (DALY) averted in 2000 international dollars.

Overall, the most cost-effective mix of interventions was similar in the African (Afr-E) and Southeast Asian (Sear-D) regions. These were the community-based newborn care package, followed by antenatal care, skilled attendance at birth offering basic maternal and neonatal care around childbirth and finally emergency neonatal and obstetric care around and after birth. There were some potentially important differences for particular interventions. Screening and treatment of maternal syphilis was relatively less cost-effective in Sear-D while community-based management of neonatal pneumonia was relatively more cost-effective in Afr-E. Scaling up the above set of interventions to 95% coverage would halve neonatal and maternal deaths in these two regions.

Preventive interventions at the community level for newborns and the primary care level for mothers and newborns are extremely cost-effective. However, universal access to clinical services is also essential if the Millennium Development Goals for maternal and child health are to be met.

11. 1 Introduction

Each year over 500,000 mothers die during pregnancy or childbirth[244] and more than 4 million babies die in the first four weeks of life, accounting for 38% of under-5 child mortality worldwide.[245] The contrast between countries is stark. Of every 1000 children born in Africat and South East Asia, 44 and 38 die in the neonatal period respectively, compared with four deaths in high income countries. A similar gulf exists for maternal mortality with rates in Sub-Saharan Africa more than 2.5 times those in Asia, which are in turn more than 20 times those in developed countries.[244] Effective interventions to reduce maternal and neonatal deaths exist,[246] but the unacceptable fact is that they are not available to people living in the poorest parts of the world.[187]
Recognizing this, the member countries of the United Nations (UN) agreed to reduce child mortality by two thirds, and maternal mortality by three quarters by the year 2015 as part of the Millennium Development Goals (MDG goals four and five respectively). Progress has been slow and few of the poorest countries in the world will achieve this at current rates. Although many bilateral and multilateral donors have committed funds, they have not been sufficient. Policy makers must decide every day how best to allocate the limited resources they have and, hopefully, what they should do when additional resources become available.

To aid these decisions, information on the costs and effectiveness of current and possible new interventions is critical to show what would be the improvement in health associated with different expenditure options. This information should, however, be realistic and include interventions in the type of combinations that would be undertaken in practice, rather than assuming that each intervention can be evaluated in isolation from others - the usual practice in cost-effectiveness analysis.

Indeed, cost-effectiveness analyses (CEA) of maternal and newborn interventions have usually been restricted to the analysis of individual interventions[247-249] with considerable variation in the analytical methods used. This, combined with variability in the settings in which the analyses have been undertaken, limits the value of the existing literature.

Recently, as part of a series on neonatal survival, the cost-effectiveness of antenatal, intrapartum and postnatal interventions of proven benefit for

reducing neonatal mortality was estimated.[246] It used a standardized methodology and a form of analysis that allows existing interventions to be evaluated at the same time as possible new interventions.[246] In this paper and the current series on the health MDGs, we develop this work further by including a more comprehensive list of maternal interventions provided during pregnancy, childbirth and the neonatal period. As sufficient information on effectiveness is not available for all maternal and newborn interventions, this analysis focuses on those for which evidence is available.

By using the standardized methods of the WHO-CHOICE (**CHO**osing **I**nterventions that are **C**ost-**E**ffective) project,[226] our work allows valid comparisons not only among the interventions presented in this paper but also the published neonatal series,[246] previous WHO-CHOICE analyses for other diseases and risk factors,[250-252] and papers in the present series.[219;221-223] A summary paper to this series will consider the overall implications of the results presented in this MDG series for priority setting in the health sector as a whole.[253] This paper, however, provides policy-makers with information necessary to evaluate if they are using the resources currently available for maternal and neonatal conditions effectively and efficiently, and how they can best achieve MDGs four and five as new resources become available.

11. 2 Methods

11.2. 1 Interventions

The analysis included 21 interventions and all possible combinations that could be undertaken programmatically, taking into account interactions in costs or effectiveness when interventions are implemented together. Table 11. 1 lists the interventions, categorized according to the level of care required to deliver them (first level maternal and newborn care vs referral level maternal and newborn care vs community-based newborn care), and the time period of implementation (antenatal vs intrapartum vs post-partum care vs newborn care).

All interventions are assumed to run for a period of 10 years (2000-2010), after which policy-makers would typically re-evaluate their strategies. Interventions are compared against a scenario of doing none of the interventions in Table 11. 1.[254] This shows the most cost-effective mix for any level of resource use. Comparison of the current mix of interventions with the ideal mix reveals inefficiencies in current resource use to be identified.[226;254] The uses and limitations of this form of analysis are further discussed below and in the first paper of this series.[254]

Table 11. 1 Description of maternal and neonatal intervention packages and components

Intervention	Description
Primary level care, including outreach	
Antenatal care:	
Tetanus toxoid	Two tetanus toxoid immunisations
Screening for pre-eclampsia	Blood pressure measurements for all pregnant women, urine test for proteinuria, and pre-referral care of women with pre-eclampsia or eclampsia
Screening and treatment	
Asymptomatic bacteriuria	Screening urine of all pregnant women at antenatal visits and treatment of bacteriuria with amoxicillin
Syphilis	Screening all pregnant women by rapid plasma reagin test and treatment of syphilis with benzathine penicillin
Skilled maternal care and immediate care of newborn (intrapartum):	
Normal delivery by skilled attendant	Includes safe delivery, cord care, identification of complications, first aid, and referral of complicated cases
Active management of third stage of labour	Administration of prophylactic oxytocin, cord clamping, and delivery of placenta by controlled cord traction
Initial management of post-partum haemorrhage	Management of post-partum haemorrhage with additional oxytocin, uterine massage, manual removal of placenta, repair of lacerations, and management of shock
Neonatal resuscitation	Detection of breathing problems and resuscitation of newborn when required
Referral level care	
Treatment of severe pre-eclampsia or eclampsia (antenatal and intrapartum)*	Inpatient care, including airway management, treatment with magnesium sulphate, treatment with antihypertensives, and birth care when undelivered
Antibiotics for preterm premature rupture of membranes (antenatal and intrapartum)*	Administration of oral antibiotics to women with preterm premature rupture of membranes, and care during labour

Steroids for preterm births (antenatal and intrapartum)*	Administration of steroids and inpatient care of women with suspected preterm labour
Management of obstructed labour, breech presentation, and fetal distress*	External cephalic version for breech presentation; management of obstructed labour, persistent breech presentation, and fetal distress by operative delivery (vacuum extraction, forceps and vaginal breech delivery, and caesarean section)
Management of severe post-partum haemorrhage (intrapartum and post partum)*	Inpatient care of post-partum haemorrhage, including blood transfusion, treatment for shock, and hysterectomy
Management of maternal sepsis (intrapartum and post partum)*	Inpatient care of maternal sepsis, including treatment with intravenous or intramuscular antibiotics
Emergency neonatal care:	
Management of very low birthweight babies*	Inpatient care for very low birthweight babies, including special feeding support, additional warmth, close monitoring, and treatment with oxygen if necessary
Management of severe neonatal infections*	Inpatient care for severe neonatal infections, including treatment with intravenous or intramuscular antibiotics
Management of severe neonatal asphyxia*	Inpatient care for neonatal encephalopathy including treatment with oxygen
Management of neonatal jaundice*	Inpatient care for severe neonatal jaundice, including phototherapy
Community care of newborn	
Community newborn care package (first two components):	
Support for breastfeeding mothers (antenatal and neonatal)	Home visits to promote early and exclusive breast feeding provided by skilled care providers and community health workers
Support for low birth weight babies	Home visits to promote extra warmth for low birth weight babies and to support breastfeeding mothers provided by skilled care providers and community health workers
Community based management of neonatal pneumonia	Home visits for diagnosis and management of pneumonia in neonates and treatment with oral antibiotic therapy provided by community health workers

*Includes costs for transportation.

11.2. 2 Regions

The costs, population health effects and cost-effectiveness of each of the interventions are evaluated for 14 epidemiological regions of the world (see www.who.int/choice for the list). Two are discussed in detail here – Afr-E, comprising countries in sub-Saharan Africa with very high adult and child mortality rates, and Sear-D, countries in South-East Asia with high adult, and very high child mortality.

11.2. 3 Intervention effects

The effectiveness of all interventions and the estimated coverage levels in the year 2000, the starting point for the analysis, are provided in Annex 11 A and B. Interventions are evaluated at 50%, 80%, and 95% population coverage levels. Effects are estimated through their impact on incidence, remission and case-fatality (see table D on bmj.com). In the absence of clinical trials showing the effectiveness of interventions implemented in combination, joint effects of interventions addressing the same cause of death are assumed to be a multiplicative function of the individual effects.[254] The impact of interventions on maternal mortality and morbidity, and on neonatal mortality, where available, are included. A lack of reliable data prevents inclusion of the impact on neonatal morbidity or stillbirths, so the benefits of some interventions isunder-estimated.

The population health effects of the interventions compared to the no intervention scenario are assessed using the population model PopMod, [239;254] with effects measured as the number of disability adjusted life years (DALYs) averted. This can also be interpreted as the number of healthy life years (HLYs) gained by the intervention. Age-weights and a 3% discount rate are applied to the calculation of DALYs averted.

11.2. 4 Intervention costs

Costs are divided into programme and patient costs. Programme costs are those involved in establishing and managing an intervention or package, and includes planning, administration, supervision, training, monitoring and evaluation. Patient costs are those involved with delivering a direct service to patients or beneficiaries and include items such as remuneration, equipment, supplies, medicines and facility. Costing methods are based on a previously published standard ingredients approach to ensure comparability of

results.[188;215] More details are in the first paper of this series (Chapter 10) and at www.who.int/choice .

The quantities of resources used are based on WHO evidence-based guidelines[231-234] as well as information obtained from the studies used for the estimates of effectiveness, to ensure consistency between costs and effectiveness. When these could not be used, expert opinion on the resource needs to introduce and run a programme was sought. More detail on costing methods and the main assumptions on resource use are available in Table E on bmj.com.

Costs are calculated for the 10 year implementation period of the intervention and are presented in international dollars (I$) referenced to the year 2000, with future costs discounted at 3% per annum.

11. 3 Results

The costs and benefits of almost 300 combinations of the individual interventions were examined. Selected results for the most cost-effective set of interventions are presented here (Tables 11. 2 and 11. 3 and Figures 11. 1 and 11. 2) with full results on www.who.int/choice or www.bmj.com. The order in which interventions would be purchased at given levels of resource availability, if cost-effectiveness is the only consideration, is called the *Expansion Path*. Tables 11. 2 and 11. 3 show the intervention packages and the additional interventions chosen at each step of the expansion path. The *average cost-effectiveness ratio* (ACER) denotes the cost-effectiveness of each package relative to the no intervention scenario - the incremental costs and benefits compared to doing nothing. The *incremental cost-effectiveness ratio* (ICER) denotes the additional costs and benefits of the new intervention or package relative to what has already been purchased - the previous point on the expansion path.

Table 11. 2 Annual costs, effects and cost-effectiveness of interventions for Afr-E in 2000

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
Intervention packages on the optimal expansion path							
A1	Community-based case management for neonatal pneumonia (95%)		1	1	1	1	1
A2	Community newborn package + Community-based pneumonia (95%)	Community newborn package: Support for breastfeeding mothers Support for low birth weight babies	9	58	7	8	57
A3	Community newborn package + Tetanus Toxoid + Community-based pneumonia (95%)	Tetanus toxoid	12	125	11	22	67
A4	Community newborn package + ANC + Community-based pneumonia (95%)	Screening for pre-eclampsia Screening & treatment of	13	160	12	27	35

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
		asymptomatic bacteruria Screening & treatment of syphillis					
A5	Community newborn package + ANC + Community-based pneumonia + SMNC (95%)	Normal delivery by skilled attendant Active management of the third stage of labour Initial management of post-partum haemorrhage Neonatal resuscitation	16	284	18	40	124
A6	Community newborn package + ANC + Community-based pneumonia + SMNC + Treatment of severe PEE (95%)	Treatment of severe pre-eclampsia/eclampsia	16	306	19	42	22
A7	Community newborn package + ANC + Community-based pneumonia + SMNC + Treatment	Facility-based care of very low birth weight babies, severe	20	498	25	61	192

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	of severe PEE + ENC (95%)	neonatal infections, severe neonatal asphyxia and neonatal jaundice					
A8	Community newborn package + ANC + Community-based pneumonia + SMNC + Treatment of severe PEE + Management of OL + ENC (95%)	Management of obstructed labour, breech and fetal distress	21	589	28	73	91
A9	Community newborn package + ANC + Community-based pneumonia + SMNC + Treatment of severe PEE + Management of OL + Steroids pre-term + ENC (95%)	Antenatal steroids for pre-term births	22	706	32	117	117
A10	Community newborn package + ANC + ENC care + SMNC + Treatment of severe PEE + Management of OL + Steroids pre-term + Maternal sepsis +	Management of maternal sepsis	22	748	34	125	42

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	Community-based pneumonia care (95%)						
A11	Community newborn package + ANC + ENC care + SMNC + Treatment of severe PEE + Management of OL + Steroids pre-term + Maternal sepsis + Antibiotics pPROM + Community-based pneumonia care (95%)	Antibiotics for pre-term premature rupture of membranes (pPROM)	22	781	35	178	34
A12	Community newborn package + ANC + ENC care + SMNC + Treatment of severe PEE + Management of OL + Steroids pre-term + Maternal sepsis + Antibiotics pPROM + PPH referral + Community-based pneumonia care (95%)	Referral care for severe post-partum haemorrhage	22	801	36	223	19

Average GDP per capita in AfrE= Int $ 1576

Table 11. 3 Annual costs, effects and cost-effectiveness of interventions for Sear-D in 2000

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
Intervention packages on the optimal expansion path							
D1	Support for breastfeeding mothers (50%)	Support for breastfeeding mothers (50% coverage)	8	49	6	6	49
D2	Support for breastfeeding mothers (80%)	Support for breastfeeding mothers (expanded to 80% coverage)	14	80	6	6	31
D3	Support for breastfeeding mothers (95%)	Support for breastfeeding mothers (exapnded to 95% coverage)	16	98	6	7	18
D4	Breast feeding + Tetanus toxoid (80%)	Tetanus toxoid (80% coverage)	24	155	7	8	57
D5	Breast feeding + Tetanus toxoid (95%)	Tetanus toxoid (expanded to 95% coverage)	28	194	7	9	39
D6	Community newborn package + Tetanus Toxoid (95%)	Support for low birth weight babies	28	195	7	20	1
D7	Community newborn	Normal delivery by skilled	31	426	14	88	231

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	package + Tetanus Toxoid + Normal delivery by skilled attendant + Active management and initial treatment of PPH (95%)	attendant Active management of third stage & initial treatment of post-partum haemorrhage					
D8	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + Normal delivery by skilled attendant + Active management and initial treatment of PPH (95%)	Screening for pre-eclampsia Screening & treatment of asymptomatic bacteruria	31	476	15	123	49
D9	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + Normal delivery by skilled	Community-based management of pneumonia	31	485	16	144	9

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	attendant + Active management and initial treatment of PPH + Community-based pneumonia (95%)						
D10	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + SMNC + Treatment of severe PEE + Community-based pneumonia (95%)	Neonatal resucitation Treatment of severe pre-eclampsia/eclampsia	31	537	17	218	52
D11	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + SMNC + Treatment of severe PEE + PPH referral + Community-	Referral care for severe post-partum haemorrhage	32	571	18	261	34

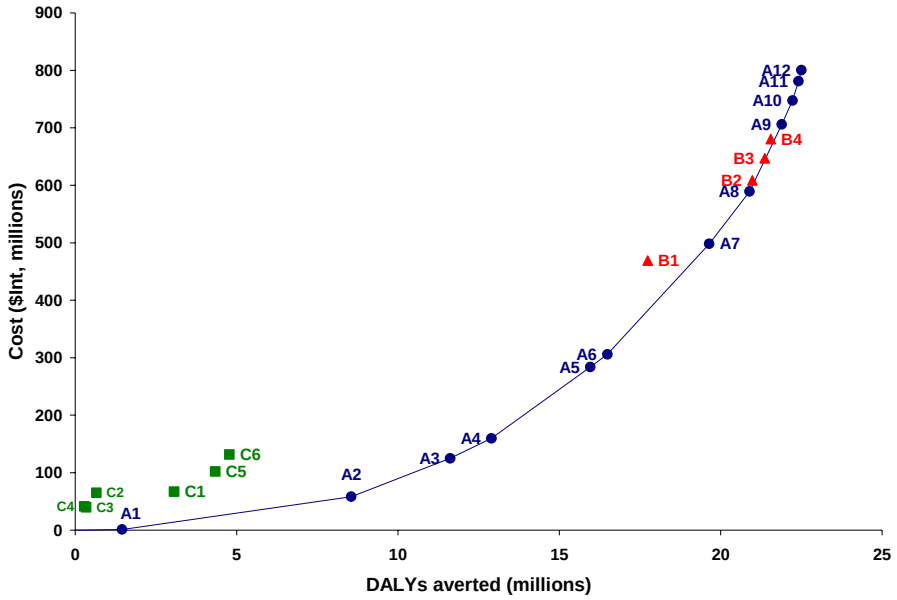
	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	based pneumonia (95%)						
D12	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + SMNC + Treatment of severe PEE + PPH referral + Community-based pneumonia + Maternal sepsis (95%)	Management of maternal sepsis	32	654	21	290	83
D13	Community newborn package + Tetanus Toxoid + Pre-eclampsia screening + Asymptomatic bacteruria screening + SMNC + Treatment of severe PEE + PPH referral + Community-based pneumonia + Maternal sepsis + ENC (95%)	Facility-based care of very low birth weight babies, severe neonatal infections, severe neonatal asphyxia and neonatal jaundice	32	1,039	32	614	385

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
D14	Community newborn package + ANC + SMNC + Treatment of severe PEE + PPH referral + Community-based pneumonia + Maternal sepsis + ENC (95%)	Screening & treatment of syphillis	33	1,049	32	699	9
D15	Community newborn package + ANC + SMNC + Treatment of severe PEE + Management of OL + PPH referral + Community-based pneumonia + Maternal sepsis + ENC (95%)	Management of obstructed labour, breech and fetal distress	33	1,234	38	2,638	186
D16	Community newborn package + ANC + SMNC + Treatment of severe PEE + Management of OL + PPH referral + Community-	Antibiotics for pre-term premature rupture of membranes (pPROM)	33	1,299	40	2,808	65

	Intervention package	Additional intervention components	Yearly DALYs averted (millions)	Yearly Cost (millions $Int)	ACER	ICER	Additional yearly resources required (millions $Int)
	based pneumonia + Maternal sepsis + ENC + Antibiotics pPROM (95%)						
D17	Community newborn package + ANC + ENC care + SMNC + Treatment of severe PEE + Management of OL + Steroids pre-term + Maternal sepsis + Antibiotics pPROM + PPH referral + Community-based pneumonia care (95%)	Antenatal steroids for pre-term births	33	1,619	50	16,93	319
Average GDP per capita in SearD= Int $ 1449							

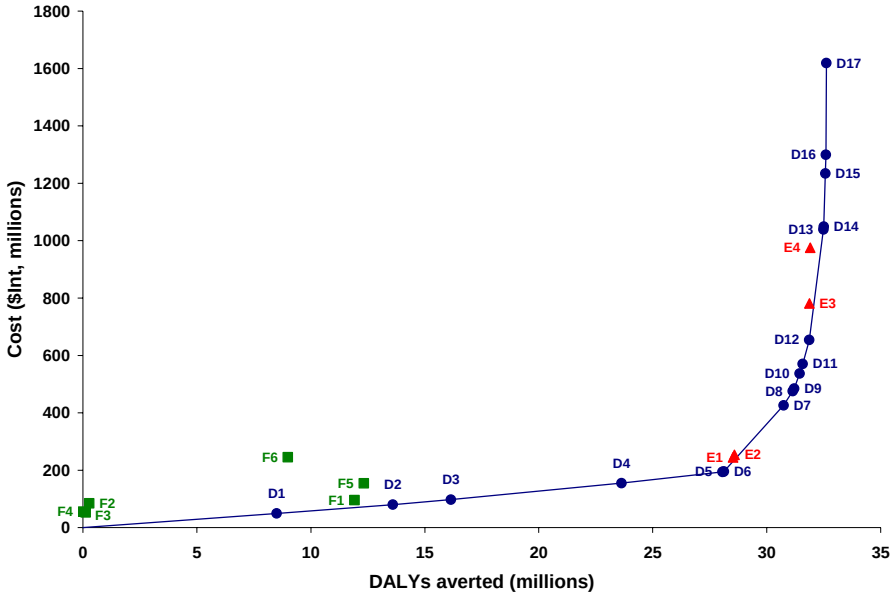
The expansion path is illustrated by the solid diagonal line in Figures 11. 1 and 11. 2. It is determined by first selecting the intervention or package with the lowest ACER. Subsequent interventions on the expansion path are chosen by selecting from the remaining options the one with the lowest ICER compared to those already purchased.

Figure 11. 1 Expansion path of the most cost-effective mix of interventions in Afr-E region in 2000



The expansion paths for both regions suggest that interventions for newborn care at the community level are highly cost-effective (e.g. promotion of breastfeeding), followed by selected antenatal care interventions (e.g. tetanus toxoid), interventions deliverable by a skilled attendant at birth in a health facility (e.g. normal delivery care by a skilled attendant), then by more complex interventions that require referral to a higher level health facility. There are, however, important differences between regions. Given the lower prevalence of syphilis in Sear-D, screening and treatment of syphilis is relatively less cost-effective in this region. Community-based management of neonatal pneumonia is relatively more cost-effective in Afr-E because of the greater contribution of pneumonia deaths there compared with Sear-D.

Figure 11. 2 Expansion path of the most cost-effective mix of interventions in Sear-D region in 2000



The expansion path shows not only the relative cost-effectiveness of interventions but also allows decision makers to see the absolute value of resources necessary to move to the next point on the expansion path. Sometimes, however, there may be insufficient resources to move to the next point. For example, if **A9** is not affordable, a decision maker may choose to implement a lower cost, but less cost-effective package such as **B2** (purchasing referral care of post-partum haemorrhage (PPH)), **B3** (purchasing referral care of maternal sepsis and PPH) or **B4** (purchasing referral care of maternal sepsis and PPH and antibiotics for pre-term premature rupture of membranes (pPROM)). Further examples are provided of alternative interventions for scaling up maternal and newborn health services in the event that the preferred intervention is unaffordable (see Tables H and I on bmj.com).

There is considerable uncertainty around the inputs used in this analysis but rather than undertake a complex multivariate uncertainty analysis, for practical policy purposes we prefer to interpret the results by ICER bands. For example, it is difficult to say with certainty that tetanus toxoid (ICER Int$22 per DALY averted in Afr-E) is more cost-effective than other ANC interventions (ICER Int$27 per DALY averted in Afr-E). We can be more certain, however, that interventions under Int$50 per DALY, such as community newborn care, ANC and SMNC, are more cost-effective than those

greater than Int$100 per DALY averted such as antibiotics for pPROM. This means that the order in which the most cost-effective interventions are introduced is up to the specific circumstances of a country. The important thing is to obtain high coverage with the group of very cost-effective interventions before those of high cost and low effectiveness are implemented.

Eliminating discounting for health benefits resulted in approximately 2.5 times more DALYs averted (data not shown). No age-weighting decreased the DALYs averted by 10-15%. Costs were 5-20% higher without discounting. While removing age weighting and discounting of DALYs favours interventions for newborns over those for mothers, the ranking of interventions and the expansion path, on the whole, remains the same.

Implementation of all the interventions covered in this analysis at 95% coverage would avert 52% of the year 2000 neonatal deaths and 51% of maternal deaths in Afr-E, and 56% of neonatal deaths and 51% of maternal deaths in Sear-D.

11. 4 Discussion

There are several strengths of the analysis undertaken for this paper and the series that set it apart from traditional cost-effectiveness analysis, and increase the value of this information for decision making. Firstly, the methods explicitly account for the fact that the total cost of a package of interventions is often less than the sum of the costs of each component evaluated separately. This is mainly due to economies of scale as many costs can be shared between activities. Total benefits are also not additive across interventions in some cases. Largely due to synergies on costs, packages of maternal and newborn interventions proved, on the whole, more cost-effective than individual interventions. This highlights the importance of considering effective integration of services and implementation of maternal and newborn interventions in parallel, particularly those with common delivery modes.

Secondly, the method allows the cost-effectiveness of interventions that are currently implemented to be evaluated as well as those that could be introduced if new resources become available. It identifies whether maternal and newborn health could be improved by scaling down some interventions, while scaling up others, even without additional resources.

Overall, community-based and antenatal care packages were found to be highly cost-effective. However, it is only with accessible and good quality clinical services - skilled attendance to allow appropriate early diagnosis and

treatment of complications, along with appropriate, timely referral to hospitals for more complex care - that the large numbers of maternal and neonatal deaths will be reduced. Although these services require much more resources, they are also very effective in reducing maternal and neonatal morbidity and mortality, and as such, are also highly cost-effective. Sustained efforts to scale up coverage of skilled attendance at birth from the 44% in Afr-E and 28% in Sear-D (in 2000) are therefore crucial to meet the MDG goals for maternal and child health. In addition, while increasing efforts have been directed towards improving antenatal coverage,[187] implementation of key components such as tetanus toxoid remains sub-optimal (51% in Afr-E and 77% in Sear-D). Furthermore, in regions where pneumonia is a leading cause of neonatal mortality, coverage of pneumonia case management is also very low (13% in Afr-E).[246] Where resource availability is unlikely to increase in the near future, it may be worthwhile to scale down implementation of less cost-effective interventions such as antibiotics for pPROM and antenatal steroids for pre-term births, and to reallocate these resources to more cost-effective options such as community-based newborn packages, ANC and skilled attendance.

As resource allocation decisions are not based solely on considerations of cost-effectiveness, the results presented here should not be used in a formulaic way, but rather cost-effectiveness information enters policy debates to be considered alongside other health system goals such as equity and acceptability to stakeholders. Of particular importance is the need to consider the feasibility of implementing these interventions. Clinical services for maternal and newborn health, in particular, require well functioning health systems, appropriate human resources, timely referral systems and institutional infrastructure.

In line with accepted practice for cost-effectiveness studies, the resources used to provide the interventions were valued from an economic perspective. They should not, therefore, be interpreted as the extra cash that countries will have to pay to provide the intervention. But taking a financial perspective, the incremental costs of scaling up maternal and newborn health services, from the current coverage level of 43% with a limited package of care, to 73% in 2015 with a full package of care, was estimated recently for 75 countries to require an initial investment of US$0.22 per capita in 2006 rising to US$1.18 per capita by 2015.[4]

As with any cost-effectiveness analysis, possible limitations of the analysis need to be carefully considered. Due to the paucity of large-scale effectiveness trials as well as the difficulty of measuring efficacy of some of the key interventions, particularly those done in combination, many of the interventions analysed are based on limited efficacy trials and/or expert opinion. These sources of treatment efficacy are often derived from studies of

good quality services provided by highly skilled professionals in developed settings and care needs to be taken when extrapolating to less developed countries. In these circumstances, feasibility studies are recommended before wider implementation is undertaken. The lack of information on neonatal morbidity and stillbirths meant that only the impact of interventions on neonatal mortality could be included (although we were able to include the impact on maternal morbidity). As a result, our analysis underestimates the total benefits of some of the interventions in this paper.

Moreover, some interventions which are beneficial or are potentially beneficial were not included in this analysis. Exclusion of these interventions was largely due to a lack of information on either intervention effectiveness or disease burden necessary for cost-effectiveness analysis, but this does not imply that they are not necessarily cost-effective. They include, among others, safe abortion, family planning, and surfactant therapy for respiratory distress syndrome.[255-257] Several of these interventions, such as the prevention of mother-to-child transmission of human immunodeficiency virus (HIV), malaria prevention and treatment, maternal and infant micro- and macro-nutrient supplementation, and vaccine preventable diseases are covered by other papers that will appear as part of this series, as well as previously published WHO-CHOICE analyses.[219;221;223;258] It should be noted that provision of these other interventions should be delivered in an integrated manner along with the ones covered in this paper to ensure a better coverage of care, as well as maximize on economies of scale and cost-effectiveness.

This paper demonstrates that while effective and efficient maternal and newborn health services are available at different resource levels, universal access to the clinical facility-based health services covered here are required to halve current levels of maternal and newborn mortality. While this will go some way to achieving MDGs 4 and 5, they are by themselves not enough to reach the targets. A coordinated and inter-sectoral response with other child and reproductive health services as well non-health sectors to reduce poverty and improve education is needed. Although not insubstantial resources are required, the interventions outlined here are highly cost-effective, which highlights the importance of overcoming other barriers to attaining the Millennium Development Goals Four and Five such as health system strengthening, financial commitment as well as the moral and political will to do so.

Annex 11. A 1 Summary of intervention effectiveness on neonatal and maternal outcomes.

No.	Intervention	Neonatal outcome(s)	Risk reduction on neonatal outcome(s)	Maternal outcome(s)	Risk reduction on maternal outcome(s)**	Reference(s)
Primary-level health facility or Outreach care						
1	Tetanus toxoid	Deaths from tetanus	90%			1-4
2	Screening for pre-eclampsia	Pre-term deaths	15%*	Deaths from hypertensive disorders during pregnancy	48%*	5-6,30
3	Screening & treatment of asymptomatic bacteruria	Pre-term deaths	10%	Deaths from sepsis and cases of infertility	10%	6-8,30
4	Screening & treatment of syphilis	Deaths from severe infection & congenital abnormality	Depends on prevalence level. Low: <1%, moderate: 1-2%, high: 4-5%.			6, 9-11,30
5	Normal delivery by skilled attendant	Deaths from severe infection & tetanus	15% (severe infection) 60% (tetanus)	Deaths from sepsis	40%*	4,6, 12-14,30
6	Active management of the third stage of labour			Deaths from PPH & cases of anaemia	62%*	15
7	Initial management of post-partum haemorrhage			Deaths from PPH & cases of anaemia	75%*	15
8	Neonatal resuscitation	Deaths from asphyxia	38%			6, 16-18,30
Referral care interventions at secondary or tertiary health care level						
9	Treatment of severe pre-eclampsia/eclampsia	Deaths from asphyxia		Deaths from hypertensive disorders during pregnancy	59%	19-20
10	Antibiotics for pre-term premature rupture of membranes (pPROM)	Deaths from severe infection	6%			6, 21-23,30

No.	Intervention	Neonatal outcome(s)	Risk reduction on neonatal outcome(s)	Maternal outcome(s)	Risk reduction on maternal outcome(s)**	Reference(s)
11	Antenatal steroids for pre-term births	Pre-term deaths	38%			6, 24,30
12	Management of obstructed labour, breech & fetal distress (OL)	Deaths from asphyxia	40%	Deaths from obstructed labour & Cases of urinary incontinence and obstetric fistula	95%*	6, 25-28,30
13	Referral care for severe post-partum haemorrhage (PPH)			Deaths from PPH & cases of anaemia	75%*	15
14	Management of maternal sepsis			Deaths from sepsis & cases of infertility	90%*	29
15	Management of very low birth weight babies (vLBW)	Pre-term deaths	25%			6, 30
16	Management of severe neonatal infections	Deaths from severe infection	50%			6, 30
17	Management of severe neonatal asphyxia	Deaths from asphyxia	3%			6, 30
18	Management of neonatal jaundice	Deaths from jaundice as part of other causes of death	4%			6, 30
Community newborn care						
19	Support for breastfeeding mothers	Deaths from severe infection & diarrhoea	Region-specific see Table 11.A.2 below			6, 31-37
20	Support for low birth weight babies	Pre-term deaths	40%			6, 30,38-43
21	Community-based case management for neonatal pneumonia	Deaths from severe infection	40%			6, 30,44-48

*Based on expert panel assessment of available evidence

**Impact is assumed to be the same for reduction in mortality and morbidity outcomes unless otherwise specified

Annex 11. A 2 Region specific effectiveness of breastfeeding promotion[1]

Region	% mortality reduction attributable the intervention (through reduction of diarrhea and ARI mortality) Year 1			% reduction attributable the intervention (through reduction of all cause mortality) Year 2		
			50 % coverage			
	% Total deaths	% female deaths	% male deaths	% Total deaths	% female deaths	% male deaths
AfrD	12%	13%	10%	19%	19%	19%
AfrE	13%	14%	11%	19%	19%	19%
AmrB	10%	10%	11%	20%	20%	20%
AmrD	10%	12%	8%	17%	17%	17%
EmrB	15%	14%	16%	37%	37%	37%
EmrD	14%	11%	16%	22%	22%	22%
SearB	6%	7%	6%	14%	14%	14%
SearD	9%	9%	9%	9%	9%	9%
WprB	11%	15%	7%	18%	18%	18%
			80% coverage			
	% Total deaths	% female deaths	% male deaths	% Total deaths	% female deaths	% male deaths
AfrD	12%	13%	11%	21%	21%	21%
AfrE	13%	15%	12%	21%	21%	21%
AmrB	11%	10%	11%	11%	11%	11%
AmrD	10%	12%	8%	18%	18%	18%
EmrB	17%	16%	18%	39%	39%	39%
EmrD	15%	13%	17%	18%	18%	18%
SearB	7%	8%	7%	19%	19%	19%
SearD	10%	10%	11%	9%	9%	9%
WprB	12%	18%	6%	8%	8%	8%
			95% coverage			
	% Total deaths	% female deaths	% male deaths	% Total deaths	% female deaths	% male deaths
AfrD	12%	13%	11%	21%	21%	21%
AfrE	14%	15%	12%	21%	21%	21%
AmrB	11%	10%	12%	7%	7%	7%
AmrD	11%	12%	9%	17%	17%	17%
EmrB	17%	16%	18%	39%	39%	39%
EmrD	15%	13%	18%	18%	18%	18%
SearB	7%	7%	6%	11%	11%	11%

[1]Adam T and Lauer JA. Modeling of breastfeeding-attributable reductions in neonatal mortality, diarrhea, and pneumonia by region (unpublished). World Health Organization, 2004.

References for Table 11. A

1 Black R, Huber D, Curlin G. Reduction of neonatal tetanus by mass immunization of non-pregnant women: duration of protection from one or two doses of aluminum-adsorbed tetanus toxoid. Bull World Health Organ 1980;58:927-930.

2 Kapoor S, Reddaiah V, Lobo J. Control of tetanus neonatorum in a rural area. Indian J Pediatr 1991;58:341-344.

3 Rahman M, Chen L, Chakraborty J, Yunus M, Faruque A, Chowdhury A. Use of tetanus toxoid for the prevention of neonatal tetanus. 1. Reduction of neonatal mortality by immunization of non-pregnant and pregnant women in rural Bangladesh. Bull World Health Organ 1982;60(1):261-267.

4 Rahman S. The effect of traditional birth attendants and tetanus toxoid in reduction of neonatal mortality. J Trop Pediatr 1982;28:163-165.

5 Abalos E, Duley L, Steyn D, Henderson-Smart D. Antihypertensive drug therapy for mild to moderate hypertension during pregnancy. Cochrane Database of Systematic Reviews 2001;Issue 1.

6 Bhutta ZA, Darmstadt GL, Hasan B, Haws R. Community-based interventions for improving perinatal and neonatal health outcomes in developing countries: review of the evidence. Pediatrics 2005;115:519-617.

7 Gichangi P, Ndinya-Achola J, Ombete J, Nagelkerke N, Temmerman M. Antimicrobial prophylaxis in pregnancy: a randomized placebo-controlled trial with cefetamet-pivoxil in pregnancy women with a poor obstetric history. Am J Obstet Gynecol 1997;177:680-684.

8 Smaill F. Antibiotics for asymptomatic bacteriuria in pregnancy (Cochrane Review). The Cochrane Library. Oxford: Update Software, 2002.

9 Wawer M, Sewankambo N, Serwadda D, et al. Control of sexually transmitted diseases for AIDS prevention in Uganda: a randomised community trial. Lancet 1999;353:525-535.

10 Osman N, Challis K, Cotiro M, Nordahl G, Bergstrom S. Maternal and fetal characteristics in an obstetric cohort in Mozambique. Afr J Reprod Health 2000;4:110-119.

11 Temmerman M, Gichangi P, Fonck K, et al. Effect of a syphilis control programme on pregnancy outcome in Nairobi, Kenya. Sex Transm Infect 2000;76:117-121.

12 Kapoor S, Reddaiah V, Lobo J. Control of tetanus neonatorum in a rural area. Indian J Pediatr 1991;58:341-344.

13 Meegan M, Conroy R, Lengeny S, Renhault K, Nyangole J. Effect on

References for Table 11. A

neonatal tetanus mortality after a culturally-based health promotion programme. Lancet 2001;358:640-641.

14 Chongsuvivatwong V, Bujakorn L, Kanpoy V, Treetrong R. Control of neonatal tetanus in southern Thailand. Int J Epidemiol 1993;22:931-935.

15 Gulmezoglu, Villar et al. WHO multicentre randomised trial of misoprostol in the management of the third stage of labour. Lancet 2001; 358:689-95

16 Bang A, Bang R, Baitule S, Reddy M, Deshmukh M. Effect of home-based neonatal care and management of sepsis on neonatal mortality: field trial in rural India. Lancet 1999;354:1955-1961.

17 Zhu X, Fang H, Zeng S, Li Y, Lin H, Shi S. The impact of the neonatal resuscitation program guidelines (NRPG) on the neonatal mortality in a hospital in Zhuhai, China. Singapore Med J 1997;38:485-487.

18 Sibley L, Sipe T. Traditional birth attendant training effectiveness: A meta-analysis (Final Technical Report). Washington, DC: Academy for Educational Development/Sara Project, 2002.

19 Magpie Trial Collaborative Group. Do women with pre-clampsia and their babies benefit from magnesium sulphate? The Magpie Trial: a randomised placebo controlled trial. Lancet 2002; 359: 1877-90

20 Duley L, Gülmezoglu A, Henderson-Smart D. Magnesium sulphate and other anticonvulsants for women with pre-eclampsia. Cochrane Database Systematic Review 2003;Issue 2.

21 Magwali T, Chipato T, Majoko F, Rusakaniko S, Mujaji C. Prophylactic augmentin in prelabor preterm rupture of the membranes. Int J Gynaecol Obstet 1999;65:261-265.

22 Kenyon S, Boulvain M, Neilson J. Antibiotics for preterm premature rupture of membranes (Cochrane Review). Oxford: Update Software, 2004.

23 Almeida L, Schmauch A, Bergstrom S. A randomised study on the impact of peroral amoxicillin in women with prelabour rupture of membranes preterm. Gynecol Obstet Invest 1996:82-84.

24 Crowley P. Prophylactic corticosteroids for preterm birth (Cochrane Review). Cochrane Database Syst Rev. Chichester, UK: John Wiley & Sons, Ltd., 2004.

25 Johanson, R.B. & Menon, B.K. Vacuum extraction versus forceps for assisted vaginal delivery. Cochrane Database of Systematic Reviews. (2): CD000224 (2000).

26 Hofmeyr, G.J. and R. Kulier, External cephalic version for breech

References for Table 11. A

presentation at term. Cochrane Database of Systematic Reviews., 2000(2):

27 Schuitemaker, N. et al. Maternal mortality after cesarean section in The Netherlands. Acta Obstetricia et Gynecologica Scandinavica. 76 (4): 332-4 (1997).

28 Hofmeyr, G.J. and M.E. Hannah, Planned Caesarean section for term breech delivery. Cochrane Database of Systematic Reviews., 2001(1)

29 French, L.M., Smaill, F.M. Antibiotic regimens for endometritis after delivery.Cochrane Database of Systematic Reviews., 2001(1)

30 Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L. Evidence-based, cost-effective interventions: how many newborn babies can we save? Lancet 2005;365:977-88.

31 Augustine T, Bhatia B. Early neonatal morbidity and mortality pattern in hospitalized children. Indian J Matern Child Health 1994;5:17-19.

32 Habicht J, DaVanzo J, Butz W. Does breastfeeding really save lives, or are apparent benefits due to biases? Am J Epidemiol 1986;123:279-290.

33 WHO. Collaborative Study Team on the Role of Breastfeeding in the Prevention of Infant Mortality. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. Lancet 2000;355:451-455.

34 Singh K, Srivastava P. The effect of colostrum on infant mortality: Urban rural differentials. Health Pop 1992;15:94-100.

35 Sikorski J, Renfrew MJ, Pindoria S, Wade A. Support for breastfeeding mothers: a systematic review. Paediatr Perinat Epidemiol 2003;17(4):407-17.

36 Lauer J, Betran A, Victora C, de Onis M, Barros A. Breastfeeding patterns and exposure to suboptimal breastfeeding among children in developing countries: review and analysis of nationally representative surveys. BMC Medicine 2004;2(1):26.

37 Lauer JA, Betrán AP, Barros AJD, de Onís M. Deaths and years of life lost due to suboptimal breastfeeding among children in the developing world: a global ecological risk assessment (submitted).

38 Charpak N, Ruiz-Pelaez J, Figueroa Cd, Charpak Y. Kangaroo mother versus traditional care for newborn infants <2000 grams: a randomized controlled trial. Pediatrics 1997;100:682-688.

39 Charpak N, Ruiz-Pelaez J, Figueroa Cd, Charpak Y. A randomized, controlled trial of kangaroo mother care: results of follow-up at 1 year of corrected age. Pediatrics 2001;108(5):1072-1079.

References for Table 11. A

40 Charpak N, Ruiz-Pelaez J, Charpak Y. Rey-Martinez Kangaroo-mother programme: an alternative way of caring for low birth weight infants? One year mortality in a two cohort study. Pediatrics 1994;94(804-810).

41 Bergman N, Jurisoo L. The "kangaroo-method" for treating low birth weight babies in a developing country. Trop Doct 1994;24:57-60.

42 Kambarami R, Chidede O, Kowo D. Kangaroo care versus incubator care in the management of well preterm infants--A Pilot Study. Ann Trop Paediatr 1998;18(2):81-86.

43 Lincetto O, Nazir A, Cattaneo A. Kangaroo Mother Care with limited resources. J Trop Pediatr 2000;46:293-295.

44 Sazawal S, Black R. Meta-analysis of intervention trials on case management of pneumonia in community settings. Lancet 1992;340(8818):528-533.

45 Bang A, Bang R, Tale O, et al. Reduction in pneumonia mortality and total childhood mortality by means of community based intervention trial in Gadchiroli India. Lancet 1990;336:201-206.

46 Bang A, Bang R, Morankar V, Sontakke P, Solanki J. Pneumonia in neonates: can it be managed in the community? Arch Dis Child 1993;68:550-556.

47 Bang A, Bang R, Sontakke P. Management of childhood pneumonia by traditional birth attendants. The SEARCH Team. Bull World Health Organ 1994;72(6):897-905.

48 Sazawal S, Black R, Group PCMT. Effect of pneumonia case management on mortality in neonates, infants, and preschool children: a meta-analysis of community-based trials. Lancet Infect Dis 2003;3(9):547-556.

Annex 11.B 1: Current coverage of antenatal care, tetanus toxoid and institutional delivery for AfrE and SearD

Intervention (number)	AfrE	SearD	Source
Antenatal care (ANC)	76%	50%	1-2
Tetanus toxoid	51%	77%	3
Institutional delivery (ID)	44%	28%	3

Annex 11.B 2: Estimated coverage for remaining interventions in 2000

Intervention	Estimated coverage*
Antenatal care (ANC)	
Screening & treatment of asymptomatic bacteriuria	ANC coverage x 0.2
Other ANC interventions	ANC coverage x 0.5
Institutional delivery (ID)	
Neonatal resuscitation	ID coverage x 0.2
Antibiotics for pPROM	ID coverage x 0.5
Referral care of neonatal sepsis	ID coverage x 0.7
Other skilled attendant & referral care interventions	ID coverage x 0.5

* Based on expert panel estimates for effective coverage of these interventions in the base year (2000).

References for Annex 11. B

1 UNICEF Global Database - Antenatal care: http://www.childinfo.org/eddb/antenatal/database1.htm

2 The State of the World's Children 2004: http://www.unicef.org/files/Table8.pdf

3 Weighted average for the WHO sub-region from the Demographic and Health Surveys (DHS): http://www.measuredhs.com

12

Achieving the Millennium Development Goals for Health:

Evaluation of Current Strategies and Future Priorities for Improving Health in Developing Countries

Evans DB, Lim SS, Adam T, Tan-Torres Edejer T and the WHO-CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Evaluation of current strategies and future priorities for improving health in developing countries. BMJ 2005 (In press) ©. Published with permission from BMJ.

CHAPTER 12

Chapter 12. Achieving the Millennium Development Goals for health: Evaluation of current strategies and future priorities for improving health in developing countries

---

Summary

This September, the United Nations is reviewing progress towards the Millennium Development Goals (MDGs), five years after signing of the Millennium Declaration. A series of papers has already been published in this Journal highlighting the areas in which current strategies aimed at an individual condition or disease could be modified to move more rapidly towards the targets. This paper, the last of the series, summarized the implications of all papers taken together for decision makers wishing to make the best use of the resources available across all diseases and conditions.

The standardized methods developed for the WHO-CHOICE project are used. They allow the efficiency of the current use of resources to be assessed at the same time as indicating which interventions should be given priority should new resources become available. The analysis also takes interactions between interventions undertaken at the same time into account, something rarely found in cost-effectiveness studies but important for understanding what actually happens in countries.

At present, a set of very cost-effective interventions are not fully implemented while money is spent in the health sector on higher cost, less effective activities. The paper identifies which interventions should be given priority now, and in the future if new resources become available.

12. 1 Introduction

Five years after the Millennium Declaration was signed, few of the poorest countries in the world are on track to achieve the Millennium Development Goals (MDGs) for health.[218;257] In September 2005, Heads of State renewed, in the UN, their commitment to these goals and to finding the resources to achieve them. The needs are substantial. An additional US$73 billion in external aid will be needed in 2006 alone for all the MDGs, with approximately $18.5 billion for health.[254] In this series of papers we have considered a different, but related question. We have examined whether the strategies adopted to date for using the available resources, and those planned for the future as more resources become available, are appropriate in view of the disappointing progress, changing circumstances and new evidence.

Five papers each focused on one of the conditions mentioned in the Millennium Development Goals or their targets - child health, maternal and neonatal health, HIV/AIDS, malaria and tuberculosis (TB).[219-223] Each analysed the costs and effects of key interventions.

This paper summarizes their key findings, providing information for programme managers and donors focusing on a particular area. We then take the perspective of a policy-maker concerned about how best to achieve the health MDGs as a group. Resources must be allocated across the three health goals and five conditions, balancing the costs and effects of a large number of technically feasible interventions addressing quite different health problems. This requires explicit comparison of all the interventions considered in the earlier papers.

12. 2 Methods

The analysis uses the results of the preceding papers.[219-223] Methods were described in Evans et al.[254] and in earlier publications.[115;188;215;239] Box 12.1 provides only essential details. Two innovations were involved to ensure the results had more relevance to practical policy decisions than traditional cost-effectiveness analysis. First, the cost-effectiveness of the existing resource use could be evaluated at the same time as the cost-effectiveness of possible future courses of action should new resources become available. Traditional cost-effectiveness analysis has usually considered only the second. Second, interactions in costs and effects between interventions that are undertaken simultaneously, as they would be in practice, were incorporated. Previous studies have generally assumed, mostly implicitly, that every intervention is undertaken in isolation from related activities.

Box 12.1. Summary Methods*

- Health gains are measured as the number of disability adjusted life years (DALYs) averted by an intervention or combination.
- Costs are reported in international dollars (I$). Web Table 1 allows readers to translate the results to US dollars if desired.[254]
- Afr-E includes countries from sub-Saharan Africa with high child and very high adult mortality rates. Sear-D consists of countries in South-East Asia with high child and adult mortality. Malaria is not a major cause of mortality in Sear-D so that set of interventions was not evaluated.
- Costs and DALYs averted depicted in Figures 1 and 2 are the additional costs and additional effects of adding the intervention to those already existing. The exceptions are A1 and B1, where costs and effects are incremental on doing nothing.

* Full details were provided in Evans et al.[254]

The major addition for the analysis in this paper is that more synergies between interventions undertaken concurrently were included. For example, different interventions that would be delivered as part of a basic obstetric package, often by the same person during the same visit, had been analysed separately in the maternal and neonatal health[5] (e.g. tetanus toxoid), HIV/AIDS[6] (prevention of mother to child transmission) and malaria[7] (intermittent presumptive treatment) papers. Cost synergies between tetanus toxoid and the other interventions addressing maternal and neonatal health were included in that paper, but here we add synergies resulting from common delivery platforms across all the health MDGs.

The individual papers have eliminated a number of interventions from further consideration because they proved to be more costly, with lower health benefits than others. These are complementary feeding without growth monitoring and health promotion in both regions and alternative regimens to artemisinin-based combination therapy to treat malaria in Africa (WebTable 2 on bmj.com).

The remaining interventions were then classified in a way that is useful for setting priorities across multiple health conditions. We earlier argued that the uncertainty around estimates of costs and health gains, especially when information must be taken from a limited number of data points, precludes basing policy advice on the point estimates of cost-effectiveness.[254] For policy purposes, interventions should be compared in terms of order-of-magnitude cost-effectiveness bands. Interventions were deemed to be highly cost-

effective if they cost less than GDP per capita to avert each DALY, and cost-effective if each DALY could be averted at a cost of between one and three times GDP per capita. Other interventions are not cost-effective.[236] This incorporates an element of affordability. Regions and countries with lower levels of national income have lower cut-points, but within any band, individual decision-makers have a menu of interventions to choose from.

12. 3 Results

Recommended strategy changes, by MDG

In some cases, current strategies and plans were found to be essentially appropriate, while more opportunities to reallocate resources existed in others. The main modifications to strategies implied by the results of the earlier papers are reported in Box 12.2 and two illustrations are provided here. Significant reductions in maternal and neonatal mortality require, for example, increased access to clinical facility-based services providing basic and emergency obstetric and neonatal care, but also increased community-based prevention, including the encouragement of breastfeeding, support of low birth-weight babies, treatment of neonatal pneumonia and wider provision of tetanus toxoid. If no new resources are forthcoming and substantial resources currently support relatively high cost, low effect interventions such as antibiotics for premature rupture of membranes in both regions, or antenatal steroids for pre-term births in Sear-D, consideration could be given to reallocating current spending to the more cost-effective interventions described above.

Box 12. 2. Achieving the MDGs more rapidly: modifications to current strategies

Disease/ Condition	*Main Changes*
Maternal and neonatal	Higher priority should be given to increasing access to clinical facility-based services providing basic and emergency obstetric and neonatal care. At the same time, there is insufficient coverage of a set of highly cost-effective preventive interventions, including community-based support for breastfeeding mothers and low birth-weight babies, treatment of neonatal pneumonia, provision of tetanus toxoid, and screening mothers for syphilis, bacteruria and pre-eclampsia/pregnancy induced hypertension. Lower priority should be given to high cost, low effect interventions such as antibiotics for pre-term rupture of membranes, and antenatal steroids for pre-term births (in Sear-D) until resource availability increases.

Child health	Increased efforts to fortify processed food staples with multiple micronutrients; especially vitamin A and zinc. Current focus on personal interventions is appropriate - measles immunization, case management of pneumonia, oral rehydration therapy. If more resources are available, vitamin A and zinc supplementation could replace fortification. Where resources are very limited, these interventions should be given higher priority than higher cost, less effective alternatives e.g. food supplementation and nutrition advice. To combat malnutrition sustainable intersectoral action is critical but, in the meantime, research on health sector interventions that could contribute more cost-effectively to the problem is urgently needed.
HIV/AIDS	The MDGs focus on reducing transmission. Strategies based on treatment of sexually transmitted infections, educating sex workers and some types of mass media messages are highly cost effective ways to do this. There is considerable uncertainty around the effectiveness of school based education but for most assumptions it is not highly cost-effective in Sear-D. With recent changes in prices, treatment with first line antiretrovirals is at least as cost effective as some of the well known preventive interventions, such as voluntary counselling and testing, and is in the spirit of the MDGs.
Malaria	In most countries of sub-Saharan Africa serious consideration should now be given to improved case management with new combination therapies (ACTs). Successful long-term malaria control requires an integrated approach of case management with ACTs supplemented by use of impregnated mosquito nets or indoor residual spraying. Where these are being successfully implemented, intermittent presumptive treatment of pregnant women can bring an important additional health benefit. Much greater efforts should be given to increasing coverage of malaria interventions to ensure efficient malaria control.
Tuberculosis	Effective treatment of infectious (sputum smear-positive) cases is the first priority in TB control, including for patients co-infected with HIV. This is in line with current practice in almost all TB control programmes, but improving case finding and cure rates should now be given high priority. Once the essential elements of TB control are in place, treatment should be extended to patients who are less infectious (sputum smear-negative), and to those carrying multidrug resistant strains of TB. Anti-retroviral therapy should be offered in conjunction with TB treatment for those infected with HIV.

The HIV/AIDS target is to reduce transmission. While there is no strong evidence yet that antiretroviral therapy has an impact on transmission, it is as cost-effective in improving health as some preventive interventions that are currently undertaken, such as voluntary counselling and testing, and is in the

spirit of the MDGs. Some increase in resources available for the highly cost-effective preventive and treatment options could be found by reallocating from less cost-effective options such as school based education or even second line ART in Sear-D, although it is important to note the potential for the cost of second-line medicines to fall significantly in the near future.

Priority setting across health MDGs

Tables 12.1 and 12.2 classify interventions into the cost-effectiveness bands described above, for Afr-E and Sear-D respectively (See WebTables 3 and 4 on bmj.com for details of costs, effects and cost-effectiveness ratios). This can be visualized using Figures 12.1 and 12.2. The vertical and horizontal axes respectively depict total costs and DALYs averted per year in the region. The diagonal lines join points of equal cost-effectiveness. As the axes are in natural logarithms, each band between the diagonals represents one order of magnitude difference in cost-effectiveness. For example, interventions B1-B10 in Sear-D cost between I$1 and I$10 per DALY averted, while B21-B33 cost between I$100 and I$1000 per DALY averted. Although treatment of new smear-positive TB cases with 80% coverage (B6) has a very similar cost-effectiveness ratio to the expansion of community-based support for breastfeeding mothers at from 80-95% (B7), the former averts many more DALYs, but also costs more.

Table 1. Interventions to achieve health MDGs in Afr-E: highly cost-effective, cost-effective and cost-ineffective

No.	MDG Target Group	Intervention
Highly cost-effective interventions*		
A1	MNH	Community-based case management for neonatal pneumonia (95%)
A2	HIV/AIDS	Mass media campaign to promote safer sex (100%)
A3	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (50%)
A4	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (expanded to 80%)
A5	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (expanded to 95%)
A6	TB	Treatment of new smear-positive TB cases only under DOTS (50%)
A7	MNH	Community newborn package (95%): Support for breastfeeding mothers Support for low birth weight babies
A8	TB	Treatment of new smear-positive TB cases only under DOTS (expanded to 80%)
A9	Malaria	Case management of malaria with artemisinin-based combination therapy (95%)
A10	TB	Treatment of new smear-positive TB cases only under DOTS (expanded to 95%)
A11	Under-5	Vitamin A fortification of food staple (95%) Zinc fortification of food staple (95%)
A12	MNH	Tetanus toxoid (95%)
A13	HIV/AIDS	Prevention of mother-to-child transmission (ANC coverage)
A14	MNH	Screening & treatment of pregnancy induced hypertension Screening & treatment of asymptomatic bacteruria

No.	MDG Target Group	Intervention
		Screening & treatment of syphilis
A15	Under-5	Measles vaccination (80%)
A16	MNH	Normal delivery by skilled attendant
		Active management of the third stage of labour
		Initial management of post-partum haemorrhage
		Neonatal resuscitation
A17	MNH	Treatment of severe pre-eclampsia/eclampsia
A18	Malaria	Insecticide-treated bednets (95%)
A19	Under-5	Measles vaccination (expanded to 95%)
A20	MNH	Facility-based care of very low birth weight babies, severe neonatal infections, severe neonatal asphyxia and neonatal jaundice
A21	HIV/AIDS	Treatment of sexually transmitted infections (current coverage)
A22	Under-5	Case management for childhood pneumonia (80%)
A23	MNH	Management of obstructed labour, breech and fetal distress (95%)
A24	HIV/AIDS	Treatment of sexually transmitted infections (expanded to ANC coverage)
A25	Under-5	Vitamin A supplementation (80%, replaces fortification)
		Zinc supplementation (80%, replaces fortification)
A26	TB	Treatment of smear-negative TB cases under DOTS (95%)
A27	Under-5	Oral rehydration therapy for diarrhoea (80%)
A28	MNH	Antenatal steroids for pre-term births (95%)
A29	Malaria	Indoor residual spraying (95%)
A30	TB	Standardized second-line drug re-treatment of TB (95%)
A31	MNH	Management of maternal sepsis

No.	MDG Target Group	Intervention
A32	Malaria	Intermittent presumptive treatment during pregnancy (95%)
A33	MNH	Antibiotics for pre-term premature rupture of membranes (pPROM)
A34	HIV/AIDS	Voluntary counseling and testing (95%)
A35	MNH	Referral care for severe post-partum haemorrhage
A36	Under-5	Vitamin A Supplementation (expanded to 95%)
		Case management for childhood pneumonia (expanded to 95%)
		Zinc supplementation (expanded to 95%)
		Oral rehydration therapy for diarrhoea (expanded to 95%)
A37	HIV/AIDS	Treatment of sexually transmitted infections (expanded to 95%)
A38	HIV/AIDS	Antiretroviral therapy: no intensive monitoring, first-line drugs only (95%)
A39	HIV/AIDS	School-based education on safer sex (95%)
A40	HIV/AIDS	Antiretroviral therapy: intensive monitoring, first-line drugs only (95%)
Not cost-effective interventions*		
A41	HIV/AIDS	Antiretroviral therapy: intensive monitoring, first- and second-line drugs (95%)
A42	Under-5	Improved complementary feeding, growth monitoring and promotion (95%)

*Highly cost-effective if ratio is less than or equal to I$1,576.
Not cost-effective if ICER is greater than $4,728

Table 2. Interventions to achieve health MDGs in Sear-D: highly cost-effective, cost-effective and cost-ineffective.

No.	MDG Target Group	Intervention
Highly cost-effective interventions*		
B1	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (50%)
B2	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (expanded to 80%)
B3	HIV/AIDS	Peer education and treatment of sexually transmitted infections for sex workers (expanded to 95%)
B4	MNH	Community-based support for breastfeeding mothers (50%)
B5	MNH	Community-based support for breastfeeding mothers (expanded to 80%)
B6	TB	Treatment of new smear-positive TB cases only under DOTS (80%)
B7	MNH	Community-based support for breastfeeding mothers (expanded to 95%)
B8	MNH	Tetanus toxoid (80%)
B9	TB	Treatment of new smear-positive TB cases only under DOTS (expanded to 95%)
B10	MNH	Tetanus toxoid (expanded to 95%)
B11	Under-5	Zinc fortification of food staple (95%)
B12	MNH	Community-based support for low birth weight babies (95%)
B13	HIV/AIDS	Mass media campaign to promote safer sex (100%)
B14	TB	Treatment of smear-negative TB cases under DOTS (95%)
B15	Under-5	Vitamin A fortification of food staple (95%)
B16	Under-5	Case management for childhood pneumonia (80%)
B17	MNH	Normal delivery by skilled attendant (95%) Active management of third stage & initial treatment of post-partum haemorrhage (95%)

No.	MDG Target Group	Intervention
B18	Under-5	Case management for childhood pneumonia (expanded to 80%)
B19	Under-5	Measles vaccination (95%)
B20	MNH	Screening & treatment of pregnancy induced hypertension
		Screening & treatment of asymptomatic bacteruria
B21	HIV/AIDS	Treatment of sexually transmitted infections (95%)
B22	MNH	Community-based case management for neonatal pneumonia (95%)
B23	Under-5	Zinc supplementation (95%, replaces fortification)
		Oral rehydration therapy for diarrhoea (95%)
B24	MNH	Neonatal resuscitation (95%)
		Treatment of severe pre-eclampsia/eclampsia (95%)
B25	TB	Standardized second-line drug re-treatment of TB (95%)
B26	MNH	Referral care for severe post-partum haemorrhage (95%)
B27	MNH	Management of maternal sepsis (95%)
B28	HIV/AIDS	Voluntary counseling and testing (95%)
B29	Under-5	Vitamin A supplementation (95%, replaces fortification)
B30	HIV/AIDS	Prevention of mother-to-child transmission (ANC coverage)
B31	MNH	Facility-based care of very low birth weight babies, severe neonatal infections, severe neonatal asphyxia and neonatal jaundice (95%)
B32	HIV/AIDS	Screening & treatment of syphilis (95%)
B33	MNH	Antiretroviral therapy: no intensive monitoring, first-line drugs only (95%)
B34	HIV/AIDS	Antiretroviral therapy: intensive monitoring, first-line drugs only (95%)
Cost-effective interventions*		
B35	HIV/AIDS	School-based education (95%)

No.	MDG Target Group	Intervention
B36	MNH	Management of obstructed labour, breech and fetal distress (95%)
B37	MNH	Antibiotics for pre-term premature rupture of membranes (pPROM)
Not cost-effective interventions*		
B38	HIV/AIDS	Antiretroviral therapy: intensive monitoring, first- and second-line drugs (95%)
B39	MNH	Antenatal steroids for pre-term births (95%)
B40	Under-5	Improved complementary feeding, growth monitoring and promotion (95%)

*Highly cost-effective if ICER is less than or equal to Int$1,449
Cost-effective if ICER is greater Int$1,449 and less than or equal to Int$4,347
Not cost-effective if ICER is greater than Int$4,347

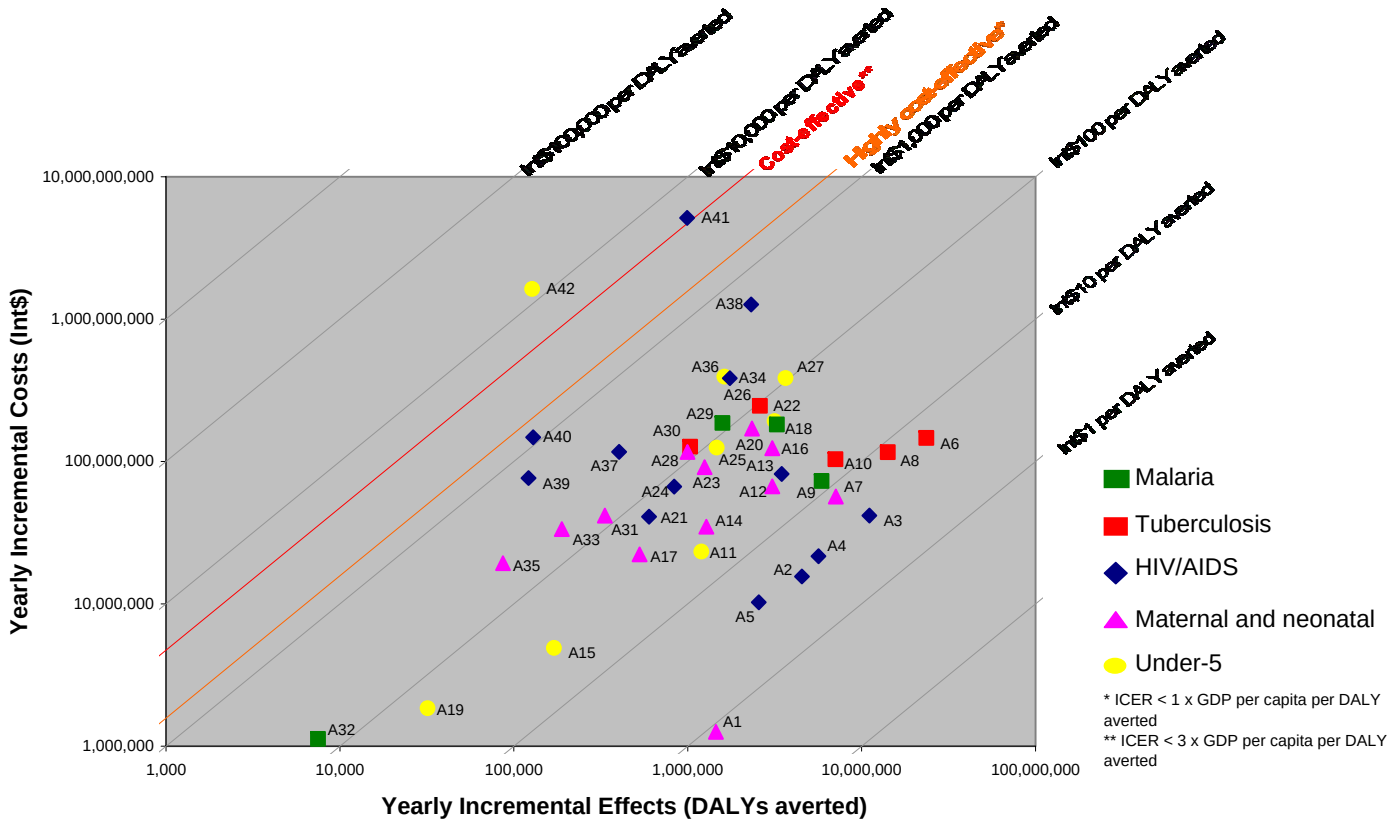


Figure 1. Interventions in Afr-E for the health MDGs: acceptability bands

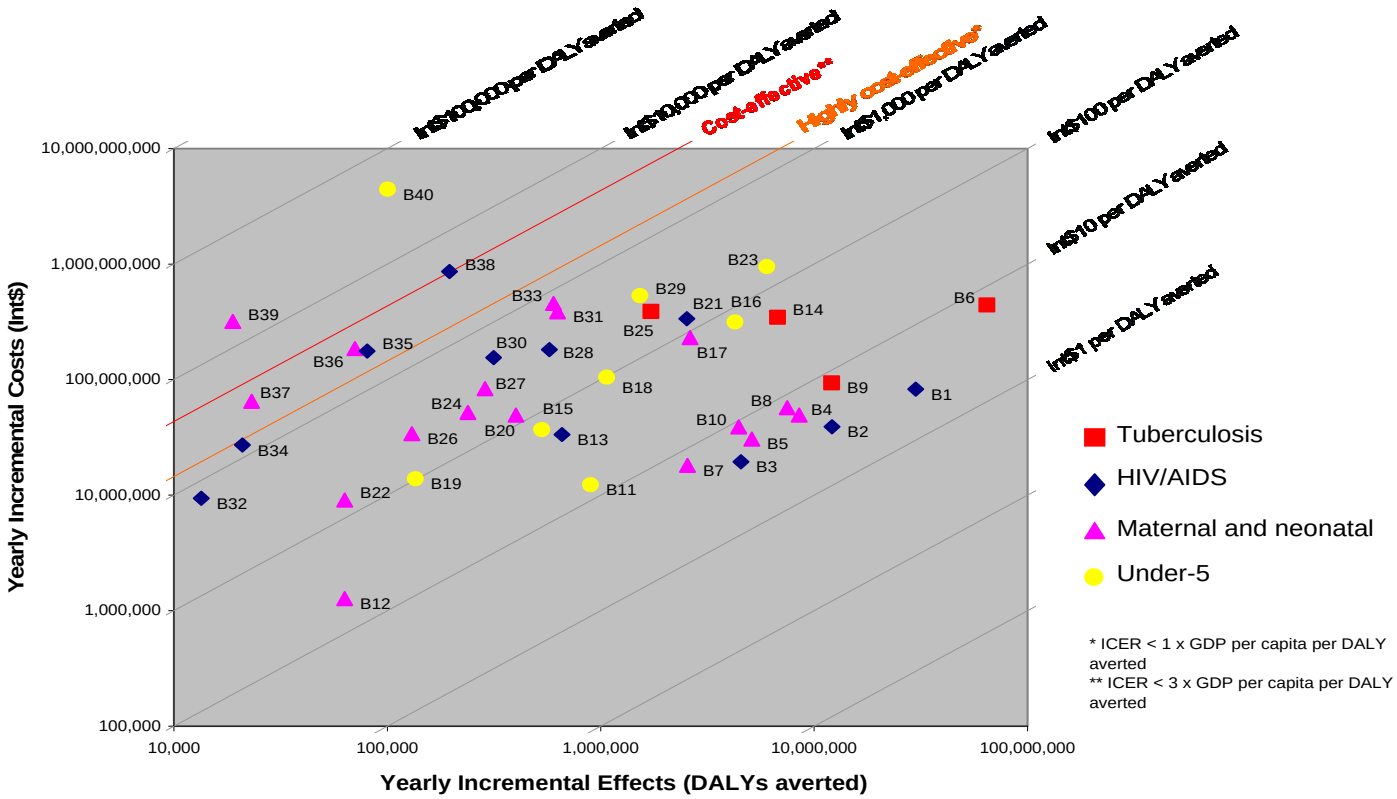


Figure 2. Interventions in Sear-D for the health MDGs: acceptability bands

12. 4 Discussion

In practice, resources are never allocated according to formulaic cost-effectiveness rules described in text books - e.g. choosing the most cost-effective intervention, then the next most cost-effective, until all resources are used. This would sometimes suggest that only prevention should take place, or only treatment, but in reality mixes of interventions are generally found. Our analysis suggests this is appropriate. The highly cost-effective group of interventions reported above includes a selection from each of the disease/condition groups in both regions, as well as mixes of treatment and preventive actions, and of population and individually focussed activities. This is true even if the threshold for highly cost-effective interventions is reduced to I$100 per DALY averted.

In both regions covered in this series, there is so much unmet need and so many underused interventions, that the opportunities for reallocating resources currently spent on the MDGs are limited. Purely on cost-effectiveness grounds, however, it is clear that priority should be given to the highly cost-effective group rather than activities such as second line antiretroviral therapy for AIDS (A41) and complementary feeding with growth monitoring and promotion (A42) in Afr-E. More could be achieved if these resources were reallocated to any of the under-used, highly cost-effective group.

A similar picture unfolds in Sear-D. Attention should be focused on scaling up interventions that are highly cost-effective rather than expanding second line antiretroviral therapy, antenatal steroids for pre-term births, and improved complementary feeding and growth monitoring and promotion (B38-B40), and even school-based education for HIV/AIDS (B35), management of obstructed labour, breech and fetal distress (B36) and antibiotics for pre-term premature rupture of membranes in pregnant women (B37), which fall into the cost-effective rather than the highly cost-effective band.

In both regions, a relatively large set of highly cost-effective interventions remain, offering considerable flexibility to adapt packages to particular contexts. The relative size of this group reflects the unmet needs, but also the fact that the MDGs were well chosen and need to be better funded. Many more interventions would fall outside the highly cost-effective group had the analysis included conditions outside the MDGs, and it is here that greater potential to reallocate resources toward the MDGs may well be found.

We accept that in practice, prioritization sometimes takes little account of cost-effectiveness information. While we think that public health decision makers have the responsibility to consider how best to improve population health for

the resources that are available, we also accept that many other considerations do, and should, influence decisions on resource allocation. There also continues to be important debates about the appropriateness of using cost-effectiveness analysis (CEA) to drive decisions in health. For example, the technique focuses only on the health gains associated with different uses of resources, so does not incorporate other effects of concern to society. This may be particularly relevant to antiretroviral treatment for HIV/AIDS, which keeps health workers and school teachers in their posts and which could, at the limit, prevent a possible break down of society.[259;260] These benefits cannot be captured in terms of DALYs. There are also a number of ethical issues implicit in CEA and its use, particularly the fact that equity is not explicitly incorporated.[260;261]

We recognize all of these concerns. Policy makers, however, cannot escape from the unfortunate fact that the resources available are insufficient even to undertake all the interventions designated in this paper as highly cost-effective, and it is not yet clear that the additional resources required to reach the MDGs will be found. In such cases, informed decisions about how to allocate the available resources require knowledge of the likely impact on population health of different courses of action. If not, decisions could be made to improve the health of a few people by a small amount at the expense of improving the health of more people by a larger amount, something that neither the proponents nor opponents of the use of CEA would want.

To implement one or more of the less cost-effective interventions identified in this paper might be justified on many grounds - perhaps an intervention, such as feeding malnourished infants or management of obstructed labour, targets a group in society with particularly poor levels of health. While this is perfectly legitimate, we argue that decision makers cannot make an informed decision without full information on the opportunities to improve population health that are foregone elsewhere.

The results presented in this series represent the best available evidence currently available. Costs and effects need to be re-evaluated from time to time, as new information becomes available or key parameters change. If, for example, prices of second-line antiretrovirals fall by the same proportion as recently observed for first-line medicines, they could also enter the highly cost-effective band.

13

Discussion

CHAPTER 13

Chapter 13. Discussion

The central theme of this thesis was to explore sources of variability in the methods used in costing studies and the implications for the validity and transferability of the results.

The first issue that was covered involved the reasons for variability in the methods used in costing and cost-effectiveness studies. These were examined and classified into avoidable and unavoidable sources of variations (Chapter 2). Secondly, a full cost-effectiveness analysis of the Integrated Management of Childhood Illness (IMCI) was presented to assess the practicality of rigorously applying the recommendations of guidelines when conducting costing studies (Chapter 3) and in using the results to provide policy-relevant information (Chapter 4).

Thirdly, a set of empirical studies was presented. The *first* explored the determinants of variation in hospital unit costs across hospital departments and countries (Chapter 5) and the possibility of using the results to extrapolate unit costs across different hospitals and settings; the *second* examined, and improved on, commonly used short-cut methods to estimate hospital costs (Chapter 6); the *third* took the analysis in the previous two chapters a step further by exploring the determinants of variation in unit costs across different hospital departments (Chapter 7); and the *fourth* explored the determinants of variation in staff time with the introduction of a new intervention (IMCI) in a facility-based setting (Chapter 8). All this empirical work was aimed at improving the transferability of cost results across within a country, and across geographical areas, when local data do not exist.

Finally, several applications of the results of these models were reported. One involved estimating the cost of the "3 by 5" strategy, a global health initiative to reduce mortality due to HIV/AIDS (Chapter 9). Another calculated costs in a consistent and comparable manner as part of an evaluation of the cost-effectiveness of interventions to achieve the Millennium Development Goal for maternal and neonatal health (Chapter 11). Both chapters related to specific diseases or disease problems. The approach was expanded to the sectoral perspective by comparing the cost-effectiveness of interventions across multiple disease areas (Chapters 10 and 12). The main findings and conclusions from the analysis presented in this thesis are summarized below.

In Chapter 2, a comparison between the nature of the recommendations of costing and cost-effectiveness guidelines and the methods used in published studies was presented. The results suggest that several, though not all, of the observed differences in the methods used in applied studies are avoidable and, therefore, could be reduced. This can be achieved, firstly, if studies comply with the recommendations of guidelines. In addition, reviewers of manuscripts for publication should ensure that only papers where it is clear that costs have been estimated appropriately get published, something that does not seem to happen. Secondly, variability in costing methods can be reduced if guidelines – or companion volumes to guidelines - provide more detail about how to follow their recommendations in practice, rather than discussing only principles. More empirical research is needed to guide this process, however, including research allowing different data collection and measurement methods to be compared. Particularly valuable would be efforts to validate relatively rapid, low cost data collection techniques compared to more expensive gold standard methods.

This thesis contributes to these methodological and practical issues in different ways. In Chapter 3, a costing study was undertaken in Tanzania as part of a multi-country economic evaluation of the Integrated Management of Childhood Illness (IMCI) — a new strategy to reduce under-five mortality. The study used a standardized costing methodology and data collection tools developed by the author for the evaluation sites in Brazil, Uganda and Bangladesh in addition to Tanzania.[79] The study used the ingredients approach to collect, analyse and report cost information. The analysis was undertaken from the societal perspective and included costs incurred at the national, district, hospital, health facility and household levels. Using standardized methods made it possible to compare the results of the different settings. The fact that costs were collected and analyzed in similar ways eliminated the possibility that differences in the findings and interpretation of the results may be due to different methods.

Chapter 4 combined the standardized costing approach and the subsequent results with the effectiveness of IMCI. This was measured in terms of the reduction in under-five mortality and compared intervention areas with IMCI with similar areas where IMCI had not been introduced. Not only did IMCI incur no extra costs compared to the way children under five were treated in its absence, but it also led to a reduction in under-five mortality. Evidence such as this, based on rigorous evaluation methods and standard tools, is of great importance to policy makers because there can be no suspicion that the results were due to inappropriate costing methods.

In Chapter 5, cost data from 1171 observations and around 60 countries were used to explore the determinants of variations in cost per bed day across

different hospital types and countries. The most important determinants of variation in the cost per bed day were the occupancy rate, an indicator of capacity utilization, hospital level as a proxy for case mix, and GDP per capita as an indicator of the availability of technology. A key result is that there is wide variation in the unit costs estimated from studies within a particular country. These differences are sometimes of an order of magnitude, and cannot only be attributed to different methods. This implies that analysts cannot simply take the cost estimates from a single study in a country to guide their assessment of the cost-effectiveness of interventions, or the costs of scaling-up activities. In some cases, they could be wrong by an order of magnitude. The model presented in this chapter fit with the data well in a statistical sense, and included a large number of countries from all regions of the world. It might, therefore, offer a way of obtaining unit costs for countries where a reasonable sample of observations on which to base unit cost estimates does not exist. It might also be preferable to extrapolating costs from another setting and using the official or purchasing power parity exchange rate, something that does not take into account the other determinants of costs explored this Chapter.

Chapter 6 addressed a different methodological issue - the use of simple rules of thumb to estimate hospital unit costs, something that has been used by some studies when data are lacking. For example, an assumed ratio of the cost of an inpatient bed day to that of an outpatient visit (commonly assumed to be 3:1) has sometimes been used to allocate total hospital costs between inpatient and outpatient departments where the only available information is total costs and the number of beds days and outpatient visits. In this chapter, I first determined if the simple rules of thumb that have been suggested are relatively accurate for the countries and hospitals for which costing studies were not available or feasible. Secondly, the determinants of the observed variations in the ratio of outpatient to inpatient unit costs were identified quantitatively, and then the estimated relationship was used to calculate unit costs for countries where data did not exist. The model was estimated using data from 832 hospitals in 28 countries.

The results suggest that no simple rule of thumb describes the available data. Even holding all other determinants except GDP per capita and hospital level constant, the ratio of the cost of an inpatient stay to that of an outpatient visit could be as low as two, and as high as 12. While this might not be totally surprising, there was also enormous variation in this ratio even for hospitals of the same type in the same setting. This means that even where some step down costing studies have been undertaken, analysts would be unwise to base policy advice on the results of only one or two studies. The hospital or hospitals for which data are available may well be atypical of other hospitals of the same type in that country. The implication is that a sufficiently large

random sample of hospitals is really required to provide accurate policy advice, further increasing the costs of doing such research.

The second conclusion is that a high proportion of the observed variation in the ratio observed across hospitals can be explained. All regression diagnostics suggested a good fit with the data and the signs of the coefficients have face validity. This suggests that the equations reported here could provide a viable alternative for estimating unit costs than rule of thumb if only a small number of observations from hospital costing studies is available. Knowledge of a few, readily obtainable explanatory variables could be used to estimate the ratio, and hence, the costs of inpatient stays and outpatient visits. Moreover, we argue that the approach is more likely to provide a reliable estimate of the average costs than relying on a single, or a few, cost studies in a given setting, even if those studies use appropriate methods. This is because of the great variation in unit costs observed for hospitals in the same country, described earlier.

Chapter 7 took this analysis a step further by exploring the determinants of variation in hospital unit costs across various inpatient departments, partly motivated by the methodological issues raised in the previous chapter. The analysis explored whether the ratio of inpatient to outpatient costs varies by type of inpatient department and sought to identify the determinants of any variation. This analysis was based on costs taken from 41 Chinese hospitals covering 13 provinces. A seemingly unrelated regression model for compositional data was used for the estimation process, where each dependent variable represented the share of a specific department's costs to total hospital costs. Compositional models take account of the key features of this type of data, namely that the share attributable to each fraction is bounded by zero and one, and that all of the fractions must sum to one. Seemingly unrelated regressions is not widely used in economic analyses, except from some examples where the determinants of demand for different types of consumption have been estimated and where the sum of the different types must add to total consumption or expenditure. As illustrated by Smith 2000, the approach leads to a significant improvement in efficiency of the estimation of compositional data because it models, rather than ignores, the correlated errors between equations.[193]

The model again explained a high proportion of the variation in the ratio of inpatient to outpatient costs. The average across the sample was close to the rules of thumb suggested by earlier studies (3:1 and 4:1) but observations ranged from 1.5:1 to 6:1. Much of the variation was due to differences in the number of staff and the proportion of department-specific bed-days to total bed-days, although some variability was also due to differences in occupancy rates. The model explained most of the variation between different hospital

departments (the adjusted R2 of the various equations ranged from 0.83 and 0.92.). This raises the question whether the relationship reported in this chapter could be used to estimate unit costs for any other hospital in China without having to under-take step down hospital costing studies. Further innovations in the future development of this work would be to use compositional models designed specifically for panel data or to explore the value of a multi-country analysis to explain variations in department-specific hospital costs across countries. If the estimated relationship could be used, it would be particularly valuable for cost-effectiveness studies and for estimates of the costs of scaling up interventions because there are many countries in the world where full step-down costing studies, the gold standard, do not exist.[4]

The decision to provide a new intervention or to modify an old one poses important questions regarding the resources required, one of which is staff time. At the margin, if health workers are currently fully occupied, it would not be possible to incorporate new activities that require additional time inputs unless new staff were employed or existing activities were eliminated or reduced. Information on how health workers currently spend their time can help programme managers determine whether it is possible to add new services within existing capacity constraints. Chapter 8 addressed this research question by measuring the amount of time required to provide IMCI, a new technology compared with existing care, and the time spent in other clinical and non-clinical activities using a time and motion study conducted in Northeast Brazil. The study also explored possible determinants of variations in consultation time across providers.

Multivariate regression analysis showed that, after controlling for other determinants, IMCI-trained providers spent significantly longer time on the average consultation with children than non-IMCI-trained providers. The size of the difference, however, was crucially determined by constraints on providers' time. The higher the workload, the lower the difference, to the extent that the difference was relatively small when the workload exceeded 50 consultations per provider per day. These findings have key policy and methodological implications. In the study sample, the mean number of consultations per provider per day was 34 and 95% of the providers had case loads lower than 54 patients per day. If this is representative of the rest of Brazil, it would be possible to introduce IMCI relatively easily throughout the country without encountering capacity constraints in terms of provider time. Interestingly, an unthinking cost analysis would suggest the opposite. Because the observed additional time is greater in areas of low workload, it would suggest that the financial costs of adding IMCI are higher where workload is low than where workload is high. On the contrary, these results are best interpreted as showing that IMCI can be introduced without

significant financial implications when capacity constraints on time are relatively low.

Applications of the empirical analysis presented in the previous chapters are presented in Chapters 9-12. Chapter 9 shows how the results of the country-specific cost estimation models in Chapter 5 and Chapter 6 can be used to inform policy makers about the cost of achieving the WHO initiative of covering three million HIV/AIDS patients with antiretroviral therapy (ARVs) by the year 2005, i.e., over a 2 year period from when the analysis was undertaken. Decision-makers need not only evidence on the costs of different interventions to guide decisions about how to allocate resources from the social or economic perspective, but also need to know the financial resources needed for planning and fund raising activities. The advantage of basing this type of analysis on the results of such models is the possibility of using country-specific estimates of unit costs and the ability to vary unit costs according to the changes in capacity utilization that would result from increased coverage levels. This takes into account both economies and diseconomies of scale.

In Chapters 10-12, a standardized methodology was used to evaluate the cost-effectiveness of interventions to achieve the health-related Millennium Development Goals (MDGs), see Chapter 10. They were first applied to a cluster of interventions for maternal and neonatal health (Chapter 11), then compared with the cost-effectiveness of the other interventions targeting the health-related MDGs, i.e., under-fives, Malaria, TB and HIV/AIDS estimated through the same methods of analysis (Chapter 12).[219;221-223] All interventions were first evaluated relative to "doing nothing" to allow the efficiency of the current use of resources to be assessed at the same time as indicating which interventions should be given priority if new resources become available. The next step was to trace out the implications for population health of adding all possible interventions and packages. The costs of each intervention scenario were then compared with the health gain to identify the most cost effective set of interventions at different levels of expenditure from this starting point. The analysis also takes interactions between interventions undertaken at the same time into account, something rarely found in cost-effectiveness studies but important for understanding what actually happens in countries (Chapter 10).

The most cost-effective mix of interventions for maternal and neonatal health was similar in the two regions that were studied in this analysis, both areas of high maternal and child mortality. These were the community-based newborn care package, followed by antenatal care, skilled attendance at birth offering basic maternal and neonatal care around childbirth and finally emergency neonatal and obstetric care around and after birth. There were some potentially important differences, however, for particular interventions.

Screening and treatment of maternal syphilis was relatively less cost-effective in the region located in South East Asia while community-based management of neonatal pneumonia was relatively more cost-effective in the African region. Although preventive interventions at the community and the primary care level are extremely cost-effective, the largest health benefits came from clinical interventions. Universal coverage of clinical services at primary and referral levels of care are, therefore, essential if the Millennium Development Goals for maternal and child health are to be met (Chapter 11).

In the comparative analysis across intervention clusters (Chapter 12), it is interesting to note that among the interventions costing less than $25 per DALY averted, at least one targets each of the health MDGs in both regions. They also include a mix of treatment and preventive actions, and a mix of population and individually focussed activities. A similar mix is found in the interventions with a high cost per DALY averted and which are not, therefore, recommended on efficiency grounds when there are severed shortages of resources. This confirms that generalizations implying that prevention is better than cure, or that population-based interventions are more efficient than those focusing on individuals, are simplistic. Each intervention needs to be evaluated on its merits, and in the light not just of local disease patterns, but also what other interventions are already being undertaken. These conclusions can only be made knowing that standard approaches to costing were used across all interventions, and that differences in total costs, and therefore in cost-effectiveness ratios, are due to true differences in the quantities or values of the resources used.

In summary, this thesis highlighted several methodological and practical issues related to the reasons for variability in costing methods and the implications on the generalizability and external validity of the results. It argued that some of the variability can be avoided if researchers increased their compliance with the recommendations of guidelines and if reviewers and editors of journals applied stricter rules in accepting manuscripts for publications. In addition, some areas where there are gaps in guidelines were identified and empirical studies were undertaken to address some of them. This included identifying alternative ways of extrapolating unit costs when this data does not exist, rather than just extrapolating the results of previous costing studies using traditional methods. The analysis shows that caution should be taken when using the results of a single hospital costing study as the basis of subsequent cost-effectiveness studies or in estimating of the costs of increasing coverage of particular interventions.[202] These studies might well use unit costs that are unrepresentative of the country as a whole if they do not assess whether the levels of capacity utilization, staff and bed days, shown here to influence hospital costs, are representative.[180]

Another important implication for the transferability of the results is reporting and adjusting for capacity utilization when estimating and presenting unit costs. General policy decisions should not be based on the results of costing studies that do not report capacity utilization. Given the way health workers adjust the time they spend in consultation with patients relative to increased case load, estimates of costs of scaling-up interventions will not be useful to policy-makers if they are only based on current costs of providing care.

Finally, rules of thumb have been used by some studies to estimate hospital costs due to the high costs and complexity of undertaking step down costing studies. It was shown that applying them across all hospitals is not appropriate as there is enough evidence that they mask important differences in hospital unit costs across hospitals and countries. Acknowledging the fact that the best alternative, step-down costing, may not always be feasible or affordable to undertake in a representative sample of facilities, the work presented in this thesis suggests that a less costly but valid alternative can be used where sufficient studies exist to allow the determinants of the ratio of inpatient to outpatient costs to be estimated econometrically. The unit costs of "out-of-sample" hospitals can be estimated using such a relationship if information on key determinants such as occupancy rates is available.

14

Reference List

Chapter 14. Reference list

1. Salah, H., Walsh, J., and Kumar, N. Cost analysis and efficiency indicators for health care: Report number 5. Summary output for 19 primary health care facilities in Alexandria, Bani Suef and Suez 1993-1994. 1997. USA, Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, School of Public Health.

2. Gold MR, Siegel JE, Russel LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.

3. Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1997.

4. Gutierrez, Juan Pablo, Johns, Benjamin, Adam, Taghreed, Bertozzi, Stefano, Edejer, Tessa Tan-Torres, Greener, Robert, Hankins, Catherine, and Evans, David B. Achieving the WHO/UNAIDS antiretroviral treatment 3 by 5 goal: what will it cost? The Lancet . 2004 Vol 364.

5. Finkler SA, Knickman JR, Hendrickson G, Lipkin M, Jr., Thompson WG. A comparison of work-sampling and time-and-motion techniques for studies in health services research. *Health Serv Res* 1993;**28**:577-97.

6. Wirth P, Kahn L, Perkoff GT. Comparability of two methods of time and motion study used in a clinical setting: work sampling and continuous observation. *Med Care* 1977;**15**:953-60.

7. Bratt JH, Foreit J, Chen PL, West C, Janowitz B, de Vargas T. A comparison of four approaches for measuring clinician time use. *Health Policy Plan* 1999;**14**:374-81.

8. Drummond MF,.Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275-83.

9. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ* 1994;**3**:95-104.

10. Creese A, Parker D. Cost analysis in primary health care: a training manual for programme managers. Geneva: World Health Organization, 1994.

11. Baltussen R, Adam T, Edejer TT-T, Hutubessy R, Acharya A, Evans DB *et al*. Methods for generalized cost-effectiveness analysis. In Edejer TT, Baltussen R, Adam T, Hutubessy R, Acharya A, Evans DB *et al*, eds. *Making choices in health: WHO guide to cost-effectiveness analysis*, Geneva: World Health Organization, 2003.

12. Bryce J, Gouws E, Adam T, Black R, Schellenberg JA, Manzi F *et al*. Improving quality and efficiency of facility-based child health care through integrated management of childhood illness in Tanzania. *Health Policy Plan* 2005;**20**.

13. Kumaranayake L, Pepperall J, Goodman H, Mills A, Walker D. Costing guidelines for HIV/AIDS prevention strategies. A companion volume to: Cost analysis in primary health care: a training manual for programme managers. London: Health Economics and Financing Programme, London School of Hygiene and Tropical Medicine, 2000.

14. Barnum H, Kutzin J. Public hospitals in developing countries: resource use, cost, financing. Baltimore: Johns Hopkins University Press published for the World Bank, 1993.

15. Shepard DS, Hodgkin D, Anthony YE. Analysis of hospital costs: a manual for managers. Geneva: World Health Organization, 2000.

16. Lombard CJ, Stegman JC, Barnard A. Modelling net expenditure of hospitals in the Cape Province. *S Afr Med J* 1991;**80**:508-10.

17. Alban A, Gyldmark M, Pedersen AV, Sogaard J. The Danish approach to standards for economic evaluation methodologies. *PharmacoEconomics* 1997;**12**:627-36.

18. Canadian Coordinating Office for Health Technology Assessment. Guidelines for economic evaluation of pharmaceuticals. Ottawa: CCOHTA, 1997.

19. Commonwealth of Australia. Guidelines for pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Advisory Committee: including economic analyses. Canberra: Department of Health and Community Services, 1995.

20. Health-Care Economics Consensus Group. Health economic guidelines: the German perspective revisited. *Pharmacoeconomics & Outcomes News* 1996;**56**:4-6.

21. Langley PC. The November 1995 revised Australian guidelines for the economic evaluation of pharmaceuticals. *PharmacoEconomics* 1996;**9**:341-52.

22. Medicare Services Advisory Committee. Funding for new medical technologies and procedures: application and assessment guidelines. Canberra: Department of Health and Aged Care, 2000.

23. Murray CJ, Evans DB, Acharya A, Baltussen RM. Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Econ* 2000;**9**:235-51.

24. Ministry of Health. Ontario guidelines for economic analysis of pharmaceutical products. 1994. Ontario, Ministry of Health.

25. Oostenbrink JB, Koopmanschap MA, Rutten FFH. Standardisation of costs; the Dutch manual for costing in economic evaluations. *PharmacoEconomics* 2002;**20**:443-54.

26. Phillips, M., Mills, A., and Dye, C. Guidelines for cost-effectiveness analysis of vector control. 1993. Geneva, World Health Organization. PEEM Guidelines Series 3.

27. World Health Organization. Estimating costs for cost-effectiveness analysis: Guidelines for managers of diarrhoeal diseases control programmes. 1988. Geneva, World Health Organization. WHO/CDD/SER/88.3.

28. World Health Organization. Expanded Programme on Immunization: Costing guidelines. 1997. Geneva, World Health Organization. EPI/GEN/790/5.

29. World Health Organization. Safe blood and blood products: costing blood transfusion services. Geneva: World Health Organization, 1998.

30. World Health Organization. Generic protocols for cost and cost-effectiveness analysis of TB diagnosis and treatment services. Geneva: World Health Organization, 1999.

31. Jacobs P,.Bachynsky J. Costing methods in the Canadian literature on the economic evaluation of health care. A survey and assessment. *Int J Technol Assess Health Care* 1996;**12**:721-34.

32. Stone PW, Chapman RH, Sandberg EA, Liljas B, Neumann PJ. Measuring costs in cost-utility analyses. Variations in the literature. *Int J Technol Assess Health Care* 2000;**16**:111-24.

33. Evers SM, Van Wijk AS, Ament AJ. Economic evaluation of mental health care interventions. A review. *Health Econ* 1997;**6**:161-77.

34. Fryatt RJ. Review of published cost-effectiveness studies on tuberculosis treatment programmes. *Int J Tuberc Lung Dis* 1997;**1**:101-9.

35. Johannesson M,.Jonsson B. Cost-effectiveness analysis of hypertension treatment--a review of methodological issues. *Health Policy* 1991;**19**:55-77.

36. Mair FS, Haycox A, May C, Williams T. A review of telemedicine cost-effectiveness studies. *J Telemed Telecare* 2000;**6 Suppl 1**:S38-S40.

37. Petrou S, Henderson J, Roberts T, Martin MA. Recent economic evaluations of antenatal screening: a systematic review and critique. *J Med Screen* 2000;**7**:59-73.

38. Kortt MA,.Bootman JL. The economics of benign prostatic hyperplasia treatment: a literature review. *Clin Ther* 1996;**18**:1227-41.

39. Hutubessy RC, Bendib LM, Evans DB. Critical issues in the economic evaluation of interventions against communicable diseases. *Acta Trop* 2001;**78**:191-206.

40. Holloway RG, Benesch CG, Rahilly CR, Courtright CE. A systematic review of cost-effectiveness research of stroke evaluation and treatment. *Stroke* 1999;**30**:1340-9.

41. Walker D,.Fox-Rushby J. Economic evaluation of parasitic diseases: a critique of the internal and external validity of published studies. *Trop Med Int Health* 2000;**5**:237-49.

42. Schrappe M,.Lauterbach K. Systematic review on the cost-effectiveness of public health interventions for HIV prevention in industrialized countries. *AIDS* 1998;**12 Suppl A**:S231-S238.

43. Walker D,.Fox-Rushby JA. Economic evaluation of communicable disease interventions in developing countries: a critical review of the published literature. *Health Econ* 2000;**9**:681-98.

44. Brown J. Economic evaluation of cancer treatments: a review of the methods. *Clin Oncol (R Coll Radiol)* 1999;**11**:78-83.

45. Kotsopoulos IA, Evers SM, Ament AJ, de Krom MC. Estimating the costs of epilepsy: an international comparison of epilepsy cost studies. *Epilepsia* 2001;**42**:634-40.

46. Jacobs P, Bachynsky J, Baladi JF. A comparative review of pharmacoeconomic guidelines. *PharmacoEconomics* 1995;**8**:182-9.

47. Drummond M, Brandt A, Luce B, Rovira J. Standardizing methodologies for economic evaluation in health care. Practice, problems, and potential. *Int J Technol Assess Health Care* 1993;**9**:26-36.

48. Johnston K, Buxton MJ, Jones DR, Fitzpatrick R. Assessing the costs of healthcare technologies in clinical trials. *Health Technol Assess* 1999;**3**:1-76.

49. Walker D. How to do (or not to do)... Cost and cost-effectiveness guidelines: which ones to use? *Health Policy Plan* 2001;**16**:113-21.

50. Mullins CD,.Ogilvie S. Emerging standardization in pharmacoeconomics. *Clin Ther* 1998;**20**:1194-202.

51. Koopmanschap MA, Brouwer WBF. Controversial costs in guidelines. In Pinto JL, Lopez-Casasnovas G, Ortun V, eds. *Economic Evaluation. From theory to practice*, pp 129-43. Springer, 2001.

52. Torrance GW, Blaker D, Detsky A, Kennedy W, Schubert F, Menon D *et al*. Canadian guidelines for economic evaluation of pharmaceuticals. Canadian Collaborative Workshop for Pharmacoeconomics. *PharmacoEconomics* 1996;**9**:535-59.

53. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;**14**:171-89.

54. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. Reply to Johannesson's and Karlsson's comment. *J Health Econ* 1997;**16**:257-9.

55. Newell ML, Dabis F, Tolley K, Whynes D. Cost-effectiveness and cost-benefit in the prevention of mother-to-child transmission of HIV in developing countries. Ghent International Working Group on Mother-to-Child Transmission of HIV [editorial; comment]. *AIDS* 1998;**12**:1571-80.

56. Garattini L, Giuliani G, Pagano E. A model for calculating costs of hospital wards: an Italian experience. *J Manag Med* 1999;**13**:71-82.

57. Asian Development Bank. Guidelines for the economic analysis of projects. Manila: Asian Development Bank, 1997.

58. Dinwiddy C, Teal F. Principles of cost-benefit analysis for developing countries. Cambridge: Cambridge University Press, 1996.

59. Johannesson M,.Meltzer D. Some reflections on cost-effectiveness analysis [editorial; comment]. *Health Econ* 1998;**7**:1-7.

60. Cohen DJ, Breall JA, Ho KK, Weintraub RM, Kuntz RE, Weinstein MC *et al*. Economics of elective coronary revascularization. Comparison of costs and charges for conventional angioplasty, directional atherectomy, stenting and bypass surgery. *J Am Coll Cardiol* 1993;**22**:1052-9.

61. Johannesson M. On aggregating QALYs: a comment on Dolan [see comments]. *J Health Econ* 1999;**18**:381-6.

62. Brouwer WBF, van Exel NJA, Koopmanschap MA, Rutten FFH. The evaluation of informal care in economic appraisal: a consideration of individual choice and societal costs of time. *Int J Technol Assess Health Care* 1999;**15**:147-60.

63. Brouwer WBF, Koopmanschap MA, Rutten FFH. Patient and informal caregiver time in cost-effectiveness analysis. *Int J Technol Assess Health Care* 1998;**14**:505-13.

64. Garber AM,.Phelps CE. Economic foundations of cost-effectiveness analysis. *J Health Econ* 1997;**16**:1-31.

65. Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. *J Health Econ* 1997;**16**:33-64.

66. Weinstein MC,.Manning WG, Jr. Theoretical issues in cost-effectiveness analysis [editorial]. *J Health Econ* 1997;**16**:121-8.

67. Barnum H, Kutzin J. Public hospitals in developing countries: resource use, cost, financing. Baltimore: The Johns Hopkins University Press for the World Bank, 1993.

68. Gove S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. *Bull World Health Organ* 1997;**75 Suppl 1**:7-24.

69. Tulloch J. Integrated approach to child health in developing countries. *Lancet* 1999;**354 Suppl 2**:SII16-20.

70. Gouws, E., Bryce, J., Habicht, J. P., Amaral, J., Pariyo, G., Schellenberg, J. A., and Fontaine, O. Improving antimicrobial use among health workers in first-level facilities: results from the Multi-country Evaluation of the Integrated Management of Childhood Illness strategy. Bulletin of the World Health Organization 82(7). 2004.

71. Oluwole D, Mason E, Costello A. Management of childhood illness in Africa. Early evaluations show promising results. *BMJ* 2000;**320**:594-5.

72. Patwari AK,.Raina N. Integrated management of childhood illness (IMCI): a robust strategy. *Indian J Pediatr* 2002;**69**:41-8.

73. Tanzania IMCI multi-country evaluation health facility survey study group. Health care for under-fives in rural Tanzania: effect of integrated management of childhood illness on observed quality of care. *Health Policy Plan* 2004;**19**:1-10.

74. Kolstad PR, Burnham G, Kalter HD, Kenya-Mugisha N, Black RE. Potential implications of the integrated management of childhood illness (IMCI) for hospital referral and pharmaceutical usage in western Uganda. *Trop Med Int Health* 1998;**3**:691-9.

75. Khan MM, Ahmed S, Saha KK. Implementing IMCI in a developing country: estimating the need for additional health workers in Bangladesh. *Human Resources for Health Development Journal* 2000;**4**:73-82.

76. Victora CG, Wagstaff A, Armstrong-Schellenberg J, Gwatkin D, Claeson M, Habicht JP. Applying an equity lens to child health and mortality: more of the same is not enough. *Lancet* 2003;**362**:233-41.

77. Adam, T., Manzi, F., Kakundwa, C, Schellenberg, J. A., Mgalula, L., de Savigny, D., Mbuya, C., and Wilczynska, K. Analysis report on the costs of IMCI in Tanzania. 2004.

78. World Bank. World Development Indicators 2001. Washington, D.C: World Bank, 2001.

79. Adam, Taghreed, Bishai, David, Khan, Mahmud, and Evans, David. Methods for the costing component of the multi-country evaluation of IMCI. http://www.who.int/imci-mce . 2004. Geneva, World Health Organization.

80. Schellenberg JA, Victora CG, Mushi A, de Savigny D, Schellenberg D, Mshinda H *et al*. Inequities among the very poor: health care for children in rural southern Tanzania. *Lancet* 2003;**361**:561-6.

81. Alonso Gonzalez M, Menendez C, Font F, Kahigwa E, Kimario J, Mshinda H *et al*. Cost-effectiveness of iron supplementation and malaria chemoprophylaxis in the prevention of anaemia and malaria among Tanzanian infants. *Bull World Health Organ* 2000;**78**:97-107.

82. Barcelona paediatric inpatient surveillance system, 1994-2001. Tanzania: Ifakara Health Research and Development Centre, 2002.

83. Armstrong S, Jr., Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O *et al*. Effectiveness and cost of facility-based Integrated Management of Childhood Illness (IMCI) in Tanzania. *Lancet* 2004;**364**:1583-94.

84. Adam, Taghreed, Manzi, Fatuma, Kakundwa, Carolyn, Schellenberg, Joanna, Mgalula, Leslie, de Savigny, Don, Mbuya, Conrad, and Wilczynska, Katarzyna. Multi-country evaluation of the integrated management of childhood illnesses (IMCI). Analysis report on the costs of IMCI in Tanzania. http://www.who.int/imci-mce . 2003. Geneva, World Health Organization.

85. Adam, Taghreed, Aikins, Moses, and Evans, David. CostIt software. www.who.int/evidence/cea . 2003. World Health Organization.

86. Stata 8. Stata Statistical Software: Release 8. College Station, TX: Stata Corporation, 2003.

87. World Bank. World development report 1993: Investing in health. New York: Oxford University Press, 1993.

88. Wammanda RD, Ejembi CL, Iorliam T. Drug treatment costs: projected impact of using the integrated management of childhood illnesses. *Trop Doct* 2003;**33**:86-8.

89. Boulanger LL, Lee LA, Odhacha A. Treatment in Kenyan rural health facilities: projected drug costs using the WHO-UNICEF integrated management of childhood illness (IMCI) guidelines. *Bull World Health Organ* 1999;**77**:852-8.

90. Khan MM, Saha KK, Ahmed S. Adopting integrated management of childhood illness module at local level in Bangladesh: implications for recurrent costs. *J Health Popul Nutr* 2002;**20**:42-50.

91. Black RE, Morris SS, Bryce J. Where and why are 10 million children dying every year? *Lancet* 2003;**361**:2226-34.

92. Walker N, Schwartlander B, Bryce J. Meeting international goals in child survival and HIV/AIDS. *Lancet* 2002;**360**:284-9.

93. Jones G, Steketee R, Black RE, Bhutta ZA, Morris SS, and the Bellagio Child Survival Study Group. How many child deaths can we prevent this year? *Lancet* 2003;**362**:65-71.

94. Bryce J, Arifeen S, Pariyo G, Lanata CF, Gwatkin D, Habicht JP *et al*. Reducing child mortality: can public health deliver? *Lancet* 2003;**362**:159-64.

95. World Health Organization (Department of Child and Adolescent Health and Development). Global monitoring of IMCI implementation. http://www.who.int/child-adolescent-health . 2005. Geneva, World Health Organization.

96. Bryce J, Victora CG, Habicht J-P, Vaughan JP, Black RE. The multi-country evaluation of the integrated management of childhood illness strategy: lessons for the evaluation of public health interventions. *Am J Public Health* 2004;**94**:400-5.

97. National Bureau of Statistics (Tanzania) and MEASURE Evaluation. Tanzania reproductive and child health survey, 1999. 1999. Chapel Hill, N.C., University of North Carolina. Measure Evaluation Technical Report Series, No 7.

98. United Nations Development Programme. Human Development Report 2001. Making new technologies work for human development. New York: UNDP, 2001.

99. Mwageni E, Momburi D, Juma Z, Irema M, Masanja H, Tanzania Essential Health Interventions Project and Adult Morbidity and Mortality Project teams. Rufiji demographic surveillance system, Tanzania. In INDEPTH Network, ed. *Population and health in developing countries. Population, health, and survival at INDEPTH sites*, Ottawa, Canada: IDRC, 2002.

100. 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**:10-8.

101. Armstrong Schellenberg JRM, Victora CG, Mushi A, de Savigny D, Schellenberg D, Mshinda H *et al*. Inequities among the very poor: health care for children in rural southern Tanzania. *Lancet* 2003;**361**:561-6.

102. INDEPTH Network. Population and health in developing countries. Population, health, and survival at INDEPTH Sites. (1), 1-356. 2002. Ottawa, Canada, IDRC. INDEPTH Monograph Series. INDEPTH Network.

103. Tanzania Ministry of Health. National health accounts. Final draft report. 2001.

104. National Bureau of Statistics (Tanzania). Household budget surverys 2000/01. President's Office, Government of Tanzania. 1-115. 2002. Tanzania.

105. Cassels, A. Guide to sector-wide approaches for health development: concepts, issues and working arrangements. 1997. Geneva, World Health Organization. WHO/ARA/97.12.

106. de Savigny, D., Kasale, H., Mbuya, C., Munna, G., Mgalula, L., Mzige, A., and Reid, G. TEHIP Interventions: an overview. 2002. Tanzania, Ministry of Health. Tanzania Essential Health Interventions Project, Discussion Paper No. 2.

107. Bennett S, Radalowicz A, Vella V, Tomkins A. A computer simulation of household sampling schemes for health surveys in developing countries. *Int.J.Epidemiol* 1994;**23**:1282-91.

108. World Health Organization, Expert Committee on Nutrition. Physical status: uses and interpretation of anthropometry. WHO Technical Report Series, No. 854. 1995. Geneva, WHO.

109. Binka FN, Ngom P, Phillips JF, Adazu K, Cohen B, Macleod BB. Assessing population dynamics in a rural African society: the Navrongo Demographic Surveillance System. *J Biosoc Sci* 1999;**31**:375-91.

110. Hayes RJ,.Bennett S. Simple sample size calculation for cluster-randomized trials. *Int.J Epidemiol* 1999;**28**:319-26.

111. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am.J Public Health* 2004;**94**:400-5.

112. Pariyo GW, Gouws E, Bryce J, Uganda IMCI Impact Study Team. Improving facility-based care for sick children in Uganda: training is not enough. *Health Policy Plan* 2005;**(in press)**.

113. Font F, Quinto L, Masanja H, Nathan R, Ascaso C, Menendez C *et al*. Paediatric referrals in rural Tanzania: the Kilombero District Study - a case series. *BMC International Health and Human Rights* 2002;**2**:4.

114. Muela SH, Mushi AK, Ribera JM. The paradox of the cost and affordability of traditional and government health services in Tanzania. *Health Policy Plan* 2000;**15**:296-302.

115. Baltussen, R. M., Adam, T, Tan Torres, T, Hutubessy, R. C., Acharya, A., Evans, D. B., and Murray, C. J. Generalized cost-effectiveness analysis: a guide. http://www.who.int/evidence/cea . 2003. Geneva, World Health Organization, Global Programme on Evidence for Health Policy.

116. Hutubessy RCW, Baltussen RMPM, Tan Torres-Edejer T, Evans DB. Generalised cost-effectiveness analysis: an aid to decision making in health. *Applied Health Economics and Health Policy* 2002;**1**:89-95.

117. Breyer F. The specification of a hospital cost function. A comment on the recent literature. *J Health Econ* 1987;**6**:147-57.

118. Conrad RF,.Strauss RP. A multiple-output multiple-input model of the hospital industry in North Carolina. *Appl Econ* 1983;**15**:341-52.

119. Cremieux PY,.Ouellette P. Omitted variable bias and hospital costs. *J Health Econ* 2001;**20**:271-82.

120. Grannemann TW, Brown RS, Pauly MV. Estimating hospital costs. A multiple-output analysis. *J Health Econ* 1986;**5**:107-27.

121. Lave JR,.Lave LB. Hospital cost functions. *Annu Rev Public Health* 1984;**5**:193-213.

122. Vitaliano DF. On the estimation of hospital cost functions. *J Health Econ* 1987;**6**:305-18.

123. Li T,.Rosenman R. Estimating hospital costs with a generalized Leontief function. *Health Econ.* 2001;**10**:523-38.

124. Wagstaff, A. and Barnum, H. Hospital cost functions for developing countries. 1992. Washington, D.C, World Bank. WPS 1044.
Ref Type: Serial (Book,Monograph)

125. Lopez-Casasnovas G,.Saez M. The impact of teaching status on average costs in Spanish hospitals. *Health Econ.* 1999;**8**:641-51.

126. Liu X,.Hsiao WC. The cost escalation of social health insurance plans in China: its implication for public policy. *Soc Sci Med* 1995;**41**:1095-101.

127. Newhouse JP. Medical care costs: how much welfare loss? *J Econ.Perspect* 1992;**6**:3-21.

128. Peden EA,.Freeland MS. Insurance effects on US medical spending (1960-1993). *Health Econ.* 1998;**7**:671-87.

129. Kirigia JM, Snow RW, Fox-Rushby J, Mills A. The cost of treating paediatric malaria admissions and the potential impact of insecticide-treated mosquito nets on hospital expenditure. *Trop Med Int Health* 1998;**3**:145-50.

130. Mitchell M, Thomason J, Donaldson D, Garner P. The cost of rural health services in Papua New Guinea. *P N G Med J* 1991;**34**:276-84.

131. Patcharanarumol W. A study of unit cost of out-patient and in-patient service of Khon Kaen Hospital in the fiscal year 1996. Bangkok: Chulalongkorn University, 1997.

132. Huff-Rousselle M. Dzongkhag costing study for Tashigang Dzongkhag. Royal Government of Bhutan: Department of Health Services, Ministry of Social Services, 1992.

133. Rannan-Eliya, R. and Somanathan, A. Bangladesh facility efficiency survey. 1999. Dhaka, Health Economics Unit, Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh and Health Policy Programme, Institute of Policy Studies of Sri Lanka. Working Paper No. 16.

134. Musau, S., Kilonzo, M., and Newbrander, W. Development of a revised FIF user fee schedule. Report July 1996. Kenya Health Care Financing Project: Contract No. 623-0245-C-00-0040-00. 2000. Boston, Management Sciences for Health.

135. Shepard DS, Carrin G, Nyandagazi P. Household participation in financing of health care in government health centres in Rwanda. In Lee K, Mills A, eds. *Health economics research in developing countries*, pp 140-64. Oxford University Press, 1990.

136. Flessa S. The costs of hospital services: a case study of Evangelical Lutheran Church hospitals in Tanzania. *Health Policy Plan* 1998;**13**:397-407.

137. Ministry of Health The Gambia and World Health Organization. Cost analysis of the health care sector in The Gambia. Volume 1. 1995. Ministry of Health.

138. Puglisi, R. and Bicknell, W. J. Functional expenditure analysis. Vol. I Final Report for Queen Elizabeth II Hospital, Maseru, Lesotho. 1990. Boston University, Health Policy Institute.

139. Wong, H. Cost analysis of Niamey hospital. USAID Project No. 683-0254. 1989. Bethesda, MD, Abt Associates Inc.

140. Ojo, K., Foley, J., Renner, A., and Kamara, F. M. Cost analysis of health services in Sierra Leone. A case study of Connaught hospital and Waterloo Community Health Centre. Annex III. 1995. Sierra Leone, Ministry of Health.

141. De Virgilio G, Haile M, Lemma A, Mariani D. Technical and economic efficiency of the Asella regional hospital in Ethiopia. *La Medicina Tropicale nelle Cooperazione allo Sviluppo* 1990;**6**:1-7.

142. Omar AO, Komakech W, Hassan AH, Singh CH, Imoko J. Costs, resource utilisation and financing of public and private hospitals in Uganda. *East Afr Med J* 1995;**72**:591-8.

143. Hansen K, Chapman G, Chitsike I, Kasilo O, Mwaluko G. The costs of HIV/AIDS care at government hospitals in Zimbabwe. *Health Policy Plan* 2000;**15**:432-40.

144. Pepperall J, Garner P, Fox-Rushby J, Moji N, Harpham T. Hospital or health centre? A comparison of the costs and quality of urban outpatient services in Maseru, Lesotho. *Int J Health Plann Manage* 1995;**10**:59-71.

145. Carey K,.Burgess JF, Jr. On measuring the hospital cost/quality trade-off. *Health Econ* 1999;**8**:509-20.

146. Raymond, S. U., Lewis, B., Meissner, P., and Norris, J. Financing and costs of health services in Belize. HCFLAC Research Report no. 2. 1987. Stony Brook, State University of New York.

147. Lewis MA, La Forgia GM, Sulvetta MB. Measuring public hospital costs: empirical evidence from the Dominican Republic. *Soc Sci Med* 1996;**43**:221-34.

148. Olave, M. and Montano, Z. Unit cost and financial analysis for the hospital 12 de Abril in Bolivia. Small Applied Research Report No. 11. 1993. Bethesda, Abt Associates Inc.

149. Gill, L. and Percy, A. Hospital costing study Glendon Hospital - Montserrat. Report No.15. 1994. Organisation of Eastern Caribbean States, Health Policy & Management Unit.

150. Russell, S. S., Gwynne, G., and Trisolini, M. Health care financing in St Lucia and costs of Victoria Hospital. HCFLAC Research Report no. 5. 1988. Stony Brook, State University of New York.

151. Snow, J. Papua New Guinea: health sector financing study project. Final Report - volume II hospital cost study. Prepared for the Papua New Guinea Department of Health under contract with the Asian Development Bank, TA No. 1091-PNG. 1990.

152. Chan S. Unit cost estimation for outpatient and inpatient departments in Nakleoung District Hospital, Cambodia. Bangkok: Chulalongkorn University, 1997.

153. Banks DA, As-Sayaideh ASK, Shafei ARSH, Muhtash A. Implementing hospital autonomy in Jordan: an economic cost analysis of Princess Raya Hospital. Bethesda, MD: The Partners for Health Reform*plus* Project, Abt Associates Inc, 2002.

154. Department of Health. The new NHS - 2001 reference costs. London: Department of Health, 2001.

155. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, and School of Public Health. Cost analysis and efficiency indicators for health care: report number 1 summary output for Bani Suef General Hospital, 1994-1994. 1997. Boston, Harvard University Press.

156. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, and School of Public Health. Cost analysis and efficiency indicators for health care: report number 2 summary output for Suez General Hospital, 1993-1994. 1997. Boston, Harvard University Press.

157. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, and School of Public Health. Cost analysis and efficiency indicators for health care: report number 3 summary output for El Gamhuria General Hospital, 1993-1994. 1997. Boston, Harvard University Press.

158. Department of Planning, Ministry of Health and Population, Data for Decision Making, Harvard School of Public Health, University of California, Berkeley, and School of Public Health. Cost analysis and efficiency indicators for health care: report number 4 summary output for 19 primary health care facilities in Alexandria, Bani Suef and Suez, 1993-1994. 1997. Boston, Harvard University Press.

159. Waters H, Abdallah H, Santillan D. Application of activity-based costing (ABC) for a Peruvian NGO healthcare provider. *Int J Health Plann Manage* 2001;**16**:3-18.

160. Robertson RL, Barona B, Pabon R. Hospital cost accounting and analysis: the case of Candelaria. *J Community Health* 1977;**3**:61-79.

161. Jorbenadze A, Zoidze A, Gzirirshvili D, Gotsadze G. Health reform and hospital financing in Georgia. *Croat Med J* 1999;**40**:221-36.

162. Mills, A. J. The cost of the district hospital. A case study from Malawi. 1991. Washington,D.C, World Bank. WPS 742.

163. Stata 7. Stata Statistical Software: Release 7.0. College Station, TX: Stata Corporation, 2001.

164. World Bank. World Development Indicators 2000. Washington, DC: World Bank, 2000.

165. King G, Honaker J, Joseph A, Scheve K. List-wise deletion is evil: what to do about missing data in political science. Paper presented at the Annual Meeting of the American Political Science Association, Boston, 1998.

166. Allison PD. Multiple imputation for missing data. A cautionary tale. *Sociological Methods & Research* 2000;**28**:301-9.

167. Honaker J, Joseph A, King G, Scheve K, Singh N. Amelia: A Program for Missing Data (Windows Version). Cambridge, MA: Harvard University, 1999.

168. Lu K,.Tsiatis AA. Multiple imputation methods for estimating regression coefficients in the competing risks model with missing cause of failure. *Biometrics* 2001;**57**:1191-7.

169. Patrician PA. Multiple imputation for missing data. *Res Nurs Health* 2002;**25**:76-84.

170. King G, Tomz M, Wittenberg J. Making the most of statistical analyses: improving interpretation and presentation. *Am J Pol Sci* 2000;**44**:341-55.

171. King G, Honaker J, Joseph A, Scheve K. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *Am Polit Sci Rev* 2000.

172. Rubin DB. Multiple imputation for survey nonresponse. New York: Wiley, 1986.

173. Cowing TG, Holtmann AG, Powers S. Hospital cost analysis: a survey and evaluation of recent studies. *Adv Health Econ Health Serv Res* 1983;**4**:257-303.

174. Gujarati DN. Basic econometrics. New York: McGraw-Hill, Inc, 1995.

175. Duan N. Smearing estimate: a nonparametric retransformation method. *JASA* 1983;**78**:605-10.

176. King G. How not to lie with statistics: avoiding common mistakes in quantitative political science. *Am J Pol Sci* 1986;**30**:666-87.

177. US General Accounting Office. Military health care. Savings to CHAMPUS from using a prospective payment system. 1990. Washington D.C., United States General Accounting Office. Fact Sheet for Congressional Requesters GAO HRD-90-136FS.

178. Wouters A. The cost and efficiency of public and private health care facilities in Ogun State, Nigeria. *Health Econ* 1993;**2**:31-42.

179. Anderson DL. A statistical cost function study of public general hospitals in Kenya. *Journal of Developing Areas* 1980;**14**:223-35.

180. Adam T, Evans DB, Koopmanschap MA. Cost-effectiveness analysis: can we reduce variability in costing methods? *Int J Technol Assess Health Care* 2003;**19**:407-20.

181. Beech R, Rudd AG, Tilling K, Wolfe CD. Economic consequences of early inpatient discharge to community-based rehabilitation for stroke in an inner-London teaching hospital. *Stroke* 1999;**30**:729-35.

182. Samet JH, Libman H, Steger KA, Dhawan RK, Chen J, Shevitz AH *et al*. Compliance with zidovudine therapy in patients infected with human immunodeficiency virus, type 1: a cross-sectional study in a municipal hospital clinic. *Am J Med* 1992;**92**:495-502.

183. Sharp V, Bray D, Carey J. Implementing a directly observed therapy program in an urban community hospital. *J Public Health Manag Pract* 1995;**1**:35-9.

184. Tansuphasawadikul S, Poprawski DM, Pitisuttithum P, Phonrat B. Nonadherence in tuberculosis treatment among HIV patients attending Bamrasnaradura Hospital, Nonthaburi. *J Med Assoc Thai* 1998;**81**:964-9.

185. Mwikisa, Chris N. HIV/AIDS interventions in Zambia: financial implications. http://www.codesria.org/Links/Home/Abstracts%20GA%201-5/AIDS_Mwikisa2.htm . 2002.

186. Koné T, Silué A, Agness-Soumahoro J, Bail RN, Shepard DS. Expenditures on AIDS in Côte d'Ivoire. In Ainsworth M, Fransen L, Over M, eds. *Confronting AIDS: evidence from the developing world*, pp 255-62. Brussels: European Commission, 1998.

187. World Health Organization. The World Health Report 2005: Make every mother and child count. Geneva: World Health Organization, 2005.

188. Adam T, Evans DB, Murray CJL. Econometric estimation of country-specific hospital costs. *Cost-effectiveness and Resource Allocation* 2003;**1**:3.

189. Barten AP. Maximum likelihood estimation of a complete system of demand equations. *European Economic Review* 1969;**1**:7-73.

190. Aitchison J. The statistical analysis of compositional data. New York: John Wiley, 1986.

191. Katz JN,.King G. A statistical model for multiparty electoral data. *Am Polit Sci Rev* 1999;**93**:15-32.

192. Billheimer, Dean, Guttorp, Peter, and Fagan, William F. Statistical analysis and interpretation of discrete compositional data. 1998. Washington, D.C, National Research Center for Statistics and the Environment. NRCSE TRS No. 011.

193. Smith M,.Kohn R. Nonparametric seemingly unrelated regression. *J Econom* 2000;**98**:257-81.

194. Salomon JA,.Murray CJL. The epidemiologic transition revisited: compositional models for causes of death by age and sex. *Population and Development Review* 2002;**28**:205-28.

195. Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *JASA* 1962;348-68.

196. Greene WH. Econometric analysis, 3rd edition. Upper Saddle River, NJ: Prentice Hall, 1997.

197. World Bank. World Development Indicators 2002. Washington, D.C: World Bank, 2002.

198. Powell AA. Aitken estimators as a tool in allocating predetermined aggregates. *JASA* 1969;**64**:913-22.

199. Haupt, Harry and Oberhofer, Walter. Estimation of constrained singular seemingly unrelated regression models. http://www.econometricsociety.org/meetings/wc00/pdf/0398.pdf . 2000. Econometric Society World Congress 2000 contributed paper no. 0398.

200. Bewley R. Allocation models: specification, estimation and applications. Cambridge, Mass: Ballinger, 1986.

201. Greene WH. Econometric analysis, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2000.

202. Schwartlander B, Stover J, Walker N, Bollinger L, Gutierrez JP, McGreevey W *et al*. AIDS. Resource needs for HIV/AIDS. *Science* 2001;**292**:2434-6.

203. Murray C, Lopez A, Ahmad OB, Inoue M, Ferguson BD, Salomon JA *et al*. World mortality in 2000: life tables for 191 countries. Geneva: World Health Organization, 2002.

204. World Health Organization. The World Health Report 1999: Making a difference. Geneva: World Health Organization, 1999.

205. Amaral, J., Gouws, E., Bryce, J., Leite, A. J. M., Cunha, A. L. A., and Victora, C. G. Effect of integrated management of childhood illness (IMCI) on health worker performance in Northeast Brazil. Cadernos de Saude Publica . 2004.

206. Gilroy K, Winch PJ, Diawara A, Swedberg E, Thiero F, Kane M *et al*. Impact of IMCI training and language used by provider on quality of counseling provided to parents of sick children in Bougouni District, Mali. *Patient Education and Counseling* 2004;**54**:35-44.

207. Lambrechts T, Bryce J, Orinda V. Integrated management of childhood illness: a summary of first experiences. *Bull.World Health Organ* 1999;**77**:582-94.

208. Victora CG, César JA. Saúde materno-infantil no Brasil. Morbimortalidade e possíveis intervenções. In Rouquayrol MZ, Almeida-Filho N, eds. *Epidemiologia e saúde*, São Paulo: Editora Médico e Científica Ltda, 2003.

209. Cunha ALA, Silva MAF, Amaral J. A estratégia de "Atenção integrada às doenças prevalentes na infância - AIDPI" e sua implantação no Brasil. *Rev Pediatr Ceará* 2001;**2**:33-8.

210. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Applied regression analysis and other multivariable methods. Pacific Grove, CA: Duxbury Press, 1998.

211. Gujarati D. Essentials of econometrics. Singapore: McGraw-Hill, 1999.

212. Adam T, Manzi F, Armstrong-Schellenberg J, Mgalula L, de Savigny D, Evans D. Does the integrated management of childhood illness cost more than routine care? Results from Tanzania. *Bull.World Health Organ* 2005;**83**:369-55.

213. UNAIDS/WHO. Emergency scale-up of antiretroviral therapy in resource-limited settings: technical and operational recommendations to achieve 3 by 5. Report of the WHO/UNAIDS International Consensus Meeting on Technical and Operational Recommendations for Emergency Scaling-up of Antiretroviral Therapy in Resource-Limited Setting, 18-21 November 2003. Lusaka, Zambia. Geneva: World Health Organization, 2003.

214. World Health Organization. Scaling up antiretroviral therapy in resource-limited settings. Guidelines for a public health approach. Executive summary. Geneva: World Health Organization, 2002.

215. Johns B, Baltussen R, Hutubessy RCW. Programme costs in the economic evaluation of health interventions. *Cost Effectiveness and Resource Allocation* 2003;**1**:1.

216. UNICEF/UNAIDS/WHO/MSF. Sources and prices of selected medicines and diagnostics for people living with HIV/AIDS. Geneva: World Health Organization, 2003.

217. Grant AD, Kaplan JE, De Cock KM. Preventing opportunistic infections among human immunodeficiency virus-infected adults in African countries. *Am J Trop Med Hyg* 2001;**65**:810-21.

218. Evans DB, Adam T, Tan-Torres Edejer T, Lim SS, Cassels A, Evans TG *et al*. Achieving the Millennium Development Goals for Health: time to reassess strategies for improving health in developing countries? *BMJ* 2005;**In press**.

219. Tan-Torres Edejer T, Aikins M, Black R, Wolfson L, Hutubessy R, Evans DB. Achieving the Millennium Development Goals for Health: cost effectiveness of strategies for child health in developing countries. *BMJ* 2005;**In press**.

220. Adam T, Lim SS, Mehta S, Bhutta ZA, Fogstad H, Mathai M *et al*. Achieving the Millennium Development Goals for Health: cost effectiveness of strategies for maternal and neonatal health in developing countries. *BMJ* 2005;**In press**.

221. Morel C, Lauer J, Evans DB. Achieving the Millennium Development Goals for Health: cost effectiveness of strategies for malaria control in developing countries. *BMJ* 2005;**In press**.

222. Hogan D, Baltussen R, Hayashi C, Lauer J, Salomon J. Achieving the Millennium Development Goals for Health: cost effectiveness of strategies to combat HIV/AIDS in developing countries. *BMJ* 2005;**In press**.

223. Baltussen R, Floyd K, Dye C. Achieving the Millennium Development Goals for Health: cost effectiveness of strategies for tuberculosis control in developing countries. *BMJ* 2005;**In press**.

224. Jamison DT,.Mosley WH. Disease control priorities in developing countries: health policy responses to epidemiological change. *Am J Public Health* 1991;**81**:15-22.

225. Tengs TO, Adams ME, Pliskin JS, Safran DG, Siegel JE, Weinstein MC *et al*. Five-hundred life-saving interventions and their cost-effectiveness [see comments]. *Risk Anal* 1995;**15**:369-90.

226. World Health Organization. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization, 2003.

227. Olsen JA,.Richardson J. Production gains from health care: what should be included in cost-effectiveness analyses? *Soc Sci Med* 1999;**49**:17-26.

228. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization, 2003.

229. Bell C, Devarajan S, Gersbach H. The long-run economic costs of AIDS: theory and application to South Africa. Washington, DC: World Bank, 2003.

230. Johns, Benjamin, Adam, Taghreed, and Evans, David B. Enhancing the comparability of costing methods: Cross-country variability in the prices of non-traded inputs to health program. 2005. Geneva, World Health Organization. EIP Discussion Paper.

231. World Health Organization. Managing complications in pregnancy and childbirth: a guide for midwives and doctors. WHO document WHO/RHR/00.7. 2003. Geneva, World Health Organization.

232. World Health Organization. Pregnancy, childbirth, postpartum and newborn care: a guide for essential practice. 2003. Geneva, World Health Organization.

233. World Health Organization. Managing newborn problems: a guide for doctors, nurses and midwives. 2003. Geneva, World Health Organization.

234. World Health Organization. WHO antenatal care randomized trial: manual for the implementation of the new model. WHO document

235. Johns B,.Baltussen R. Accounting for the cost of scaling-up health interventions. *Health Econ* 2004;**13**:1117-24.

236. Commission on Macroeconomics and Health. Macroeconomics and health: investing in health for economic development. Report of the Commission on Macroeconomics and Health. Center for International Development at Harvard University. http://www.cid.harvard.edu/cidcmh/CMHReport.pdf . 2001. World Health Organization.

237. Devarajan, S., Miller, M., and Swanson, E. V. Goals for development: history, prospects, and costs. World Bank Policy Research Working Paper No. 2819. 2002. Washington, DC, World Bank.

238. UN Millennium Project. Investing in development: a practical plan to achieve the Millennium Development Goals. New York: Earthscan, 2005.

239. Lauer JA, Rohrich K, Wirth H, Charette C, Gribble S, Murray CJL. PopMod: a longitudinal population model with two interacting disease states. *Cost-effectiveness and Resource Allocation* 2003;**1**.

240. Murray CJL, Lopez AD, Jamison DT. The global burden of disease in 1990: summary results, sensitivity analysis and future directions. *Bull World Health Organ* 1994;**72**:495-509.

241. Murray, C., Lopez, A., Mathers, C., and Stein, C. The global burden of disease 2000 project: aims, methods and data sources. 2000. Geneva, World Health Organization. Global Programme on Evidence for Health Policy Discussion Paper No. 36.

242. Briggs AH, Goeree R, Blackhouse G, O'Brien BJ. Probabilistic Analysis of Cost-Effectiveness Models: Choosing between Treatment Strategies for Gastroesophageal Reflux Disease. *Med Decis Making* 2002;**22**:290-308.

243. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M *et al*. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Econ.* 2005;**14**:339-47.

244. World Health Organization (Department of Reproductive Health and Research). Maternal mortality in 2000: estimates developed by WHO, UNICEF, UNFPA. Geneva: World Health Organization, 2004.

245. Lawn JE, Cousens S, Zupan J. 4 million neonatal deaths: when? Where? Why? *Lancet* 2005;**365**:891-900.

246. Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L. Evidence-based, cost-effective interventions: how many newborn babies can we save? *Lancet* 2005;**365**:977-88.

247. Terris-Prestholt F, Watson-Jones D, Mugeye K, Kumaranayake L, Ndeki L, Weiss H *et al*. Is antenatal syphilis screening still cost effective in sub-Saharan Africa. *Sex Transm.Infect.* 2003;**79**:375-81.

248. Griffiths UK, Wolfson LJ, Quddus A, Younus M, Hafiz RA. Incremental cost-effectiveness of supplementary immunization activities to prevent neonatal tetanus in Pakistan. *Bull.World Health Organ* 2004;**82**:643-51.

249. Nakhaee N, Mirahmadizadeh AR, Gorji HA, Mohammadi M. Assessing the cost-effectiveness of contraceptive methods in Shiraz, Islamic Republic of Iran. *East Mediterr.Health J.* 2002;**8**:55-63.

250. Dziekan G, Chisholm D, Johns B, Rovira J, Hutin YJF. The cost-effectiveness of policies for the safe and appropriate use of injection in healthcare settings. *Bull World Health Organ* 2003;**81**:277-85.

251. Murray CJL, Lauer JA, Hutubessy RC, Niessen L, Tomijima N, Rodgers A *et al*. Effectiveness and costs of interventions to lower systolic blood pressure and cholesterol: a global and regional analysis on reduction of cardiovascular-disease risk. *Lancet* 2003;**361**:717-25.

252. Shibuya K, Ciecierski C, Guindon E, Bettcher DW, Evans DB, Murray CJ. WHO Framework Convention on Tobacco Control: development of an evidence based global public health treaty. *BMJ* 2003;**327**:154-7.

253. Evans D, Lim S, Adam T, Tan-Torres Edejer T, and the WHO-CHOICE MDG team. Achieving the Millennium Development Goals for Health: Evaluation of current strategies and future priorities for improving health in developing countries. *BMJ* 2005;**In press**.

254. Evans D, Tan-Torres Edejer T, Adam T, Lim S, and the WHO-CHOICE MDG team. Achieving the Millennium Development Goals for Health: methods to assess the costs and effects of interventions for improving health in developing countries. *BMJ* 2005;**In press**.

255. Stevens TP, Blennow M, Soll RF. Early surfactant administration with brief ventilation vs selective surfactant and continued mechanical ventilation for preterm infants with or at risk for respiratory distress syndrome. *Cochrane Database Syst Rev* 2004;**3**:CD003063.

256. Kulier R, Gumodoka AM, Hofmeyr GJ, Cheng LN, Campana A. Medical methods for first trimester abortion. *Cochrane Database Syst Rev* 2004;**2**:CD002855.

257. Nanda, G., Switlick, K., and Lule, E. Accelerating progress towards achieving the MDG to improve maternal health: a collection of promising approaches. Health, Nutrition and Population (HNP) Discussion Paper. 2005. Washington, DC, The International Bank for Reconstruction and Development and the World Bank. HNP Discussion Paper.

258. Baltussen R, Knai C, Sharan M. Iron fortification and iron supplementation are cost-effective interventions to reduce iron deficiency in four subregions of the world. *J.Nutr.* 2004;**134**:2678-84.

259. Evans David B. Communicable diseases: alternative perspectives. Perspective Paper 2.1. In Lomborg B, ed. *Global crises, global solutions*, pp 115-23. Cambridge, UK: Cambridge University Press, 2004.

260. Yong Kim J, Shakow A, Mate K, Vanderwarker C, Gupta R, Farmer P. Limited good and limited vision: multidrug-resistant tuberculosis and global health policy. *Social Science & Medicine* 2005;**61**:847-59.

261. Brock DW. Priority to the worst off in health care resource prioritization. In Battin M, Rhodes R, Silvers A, eds. *Health care and social justice*, New York: Oxford University Press, 2002.

English Summary

English summary

To ensure that policy makers are provided with consistent and valid evidence, it is important that costing studies use comparable and appropriate methods. Otherwise it is not possible to compare the efficiency of the various competing alternatives or be sure that the interventions claimed to be cost-effective have been analyzed in an appropriate manner.

A further complication relates to the fact that thorough economic evaluation of health interventions requires skilled economists and can be relatively expensive to undertake. In many countries, the necessary financial and human resources are not available and policy makers must draw on the results of studies undertaken in other settings and try to apply them to their own. For this reason, it is not only important to use comparable methods but also to report results in a way that allows policy-makers to assess the replicability and transferability of results between settings.

The need for consistency and standardization of methods for economic appraisals has been recognized for some time and has led to the development of several sets of guidelines for economic evaluation and for costs. Despite this, considerable diversity is still apparent in applied studies. Some of this diversity might be defensible, and some might not be.

The overall objective of the thesis is to explore the reasons for variability in costing methods used in costing and cost-effectiveness studies; to test the practicality of applying the recommendations of guidelines when conducting costing studies and in using the results to provide policy-relevant information; to understand the determinants of variation in unit costs; and finally to explore valid methods of extrapolating cost information to other settings, when local data do not exist.

The goal is to provide cost data in a way that allows the costs and cost-effectiveness of many different types of interventions to be compared, and provides the maximum assistance to researchers in one setting to generalize from the results of studies undertaken in another setting. Only then will the results of costing exercises be of widespread practical value to policy-makers and researchers.

The Chapters of this thesis are organized as follows. First, Chapter 2 explores the sources of variations in costing methods used in applied studies, discusses the nature of these variations relative to the recommendations of costing guidelines and the possibility of reducing some of them. Next, a full cost-

effectiveness analysis of the Integrated Management of Childhood Illness (IMCI) is presented to assess the practicality of rigorously applying the recommendations of guidelines when conducting costing studies (Chapter 3) and in using the results to provide policy-relevant information (Chapter 4). Then, a set of empirical studies is presented. The *first* explores the determinants of variations in hospital unit costs across hospital departments and countries (Chapter 5) and the possibility of using the results in extrapolating unit costs across different hospitals and settings; the *second* examines, and improves on, commonly used short-cut methods to estimate hospital costs (Chapter 6); the *third* takes the analysis in the previous two chapters a step further by exploring the determinants of variation in unit costs across different hospital departments (Chapter 7); and the *fourth* explores the determinants of staff time spent in providing IMCI in a facility-based setting (Chapter 8), all providing potentially useful and more valid alternatives to extrapolating cost information when local data do not exist. Finally, several applications of the results presented in these models are presented, e.g., to estimate the cost of the "3 by 5" strategy, a global health initiative to reduce mortality due to HIV/AIDS (Chapter 9), and to evaluate the cost-effectiveness of interventions to achieve the health-related Millennium Development Goals in a consistent and comparable manner both from a disease/condition perspective, i.e., interventions for maternal and neonatal health (Chapter 11) and from a sectoral perspective comparing the cost-effectiveness of interventions across multiple disease areas (Chapters 10 and 12).

This thesis highlighted several methodological and practical issues related to the reasons for variability in costing methods and the implications on the generalizability and external validity of the results. It argued that some of the variability can be avoided if researchers increased their compliance with the recommendations of guidelines and if reviewers and editors of journals applied stricter rules in accepting manuscripts for publications. In addition, some areas where there are gaps in guidelines were identified and empirical studies were undertaken to address some of them. This included using econometric methods to estimate unit costs in settings where no such data exist, rather than just extrapolating the results of previous costing studies using traditional methods. The analysis shows that caution should be taken when using the results of a single hospital costing study as the basis of subsequent cost-effectiveness studies or in estimating of the costs of increasing coverage of particular interventions. These studies might well use unit costs that are unrepresentative of the country as a whole if they do not assess whether the levels of capacity utilization, staff and bed days, shown here to influence hospital costs, are representative.

Another important implication for the transferability of the results is reporting and adjusting for capacity utilization when estimating and presenting unit costs. General policy decisions should not be based on the results of costing studies that do not report capacity utilization. Given the way health workers adjust the time they spend in consultation with patients relative to increased case load, estimates of costs of scaling-up interventions will not be useful to policy-makers if they are only based on current costs of providing care.

Finally, rules of thumb have been used by some studies to estimate hospital costs due to the high costs and complexity of undertaking step down costing studies. It was shown that applying them across all hospitals is not appropriate as there is enough evidence that they mask important differences in hospital unit costs across hospitals and countries. Acknowledging the fact that the best alternative, step-down costing, may not always be feasible or affordable, the work presented in this thesis suggests that a less costly but valid alternative can be used where sufficient studies exist to allow the determinants of the ratio of inpatient to outpatient costs to be estimated econometrically. The unit costs of "out-of-sample" hospitals can then be estimated using such a relationship if information on key determinants such as occupancy rates is available.

Samenvatting

Summary in Dutch

Samenvatting : Summary in Dutch

Bronnen van variabiliteit in berekeningsmethoden van kosten

Gevolgen voor het overzetten van resultaten van kosten-effectiviteitsanalyses van het ene land naar het andere land

Om er voor te zorgen dat beleidsmakers voorzien worden van consistente en geldige gegevens, is het van groot belang dat men bij kostenstudies gebruik maakt van vergelijkbare en toepasbare berekeningsmethoden. Doet men dit niet, dan is het onmogelijk om de doelmatigheid van verschillende met elkaar concurrerende zorgprogramma's te vergelijken, of kan men er niet zeker van zijn dat de zogenaamde 'kosteneffectieve interventies' op een juiste manier geanalyseerd zijn.

Een gerelateerd probleem is dat grondige economische evaluaties van gezondheidsinterventies relatief duur kunnen zijn, en dat hier ervaren economen voor nodig zijn. In veel landen zijn hiervoor noch de financiele noch de personele middelen voor beschikbaar en zijn beleidsmakers genoodzaakt de studieresultaten van andere landen toe te passen op de lokale situatie. Om deze reden is het niet alleen belangrijk om in kostenstudies vergelijkbare berekeningsmethoden te gebruiken, maar ook om de verkregen resultaten zodanig te rapporteren dat beleidsmakers de betrouwbaarheid en de overdraagbaarheid van deze resultaten naar andere landen kunnen beoordelen

De behoefte aan consistente en gestandaardiseerde berekeningsmethoden voor economische studies bestaat al enige tijd en heeft geleid tot de ontwikkeling van verschillende richtlijnen voor economische evaluaties en voor kosten. Ondanks deze ontwikkeling ziet men nog steeds een aanzienlijke diversiteit in dit soort toegepaste studies. Sommige van deze verschillen zouden verdedigbaar kunnen zijn, andere wellicht niet.

De doelstellingen van dit proefschrift zijn het verkennen van de redenen voor variabiliteit in kostenmethoden zoals die gebruikt worden in kosten- en kosten-effectiviteitsstudies; het nagaan van de toepasbaarheid van de aanbevelingen uit richtlijnen voor kostenstudies en of de verkregen resultaten beleidsrelevante informatie opleveren; inzicht verkrijgen in de determinanten van variatie in kostprijzen; en tenslotte het bepalen van valide berekeningsmethoden voor het extrapoleren van kosteninformatie naar andere landen in het geval dat lokale data ontbreken.

Het uiteindelijke doel is ervoor te zorgen dat kostengegevens zodanig worden aangeleverd dat de kosten en kosten-effectiviteit van verschillende typen interventies met elkaar vergeleken kunnen worden, en dat het onderzoekers ondersteunt om de resultaten van dit soort onderzoek te generaliseren naar andere landen. Alleen dan zullen de resultaten van deze kostenstudies praktisch relevant zijn voor beleidsmakers en onderzoekers.

De opzet van van dit proefschrift is alsvolgt. Allereest beschrijft hoofdstuk 2 de bronnen van variabiliteit in kostenmethoden in toegepaste studies, en bediscussieert deze variabiliteit in het licht van aanbevelingen uit kostenrichtlijnen en de mogelijkheid deze variatie te reduceren. Ten tweede wordt er een volledige kosten-effectiviteitanalyse gepresenteerd van het zogenoemde *Integrated Management of Childhood Illness (IMCI)* om de toepasbaarheid in de praktijk te kunnen beoordelen van aanbevelingen uit richtlijnen bij het uitvoeren van kostenstudies (hoofdstuk 3), en voor het gebruik van de resultaten voor het aanleveren van beleidsrelevante informatie (hoofdstuk 4). Vervolgens wordt er een aantal empirische studies gepresenteerd. De *eerste* hiervan verkent de determinanten van variabiliteit in de kostprijs van ziekenhuizen tussen ziekenhuisafdelingen en landen (hoofdstuk 5), en de mogelijkheid om de resultaten te gebruiken voor het extrapoleren van kostprijzen naar verschillende ziekenhuizen en landen. De *tweede* studie onderzoekt veel gebruikte verkorte methoden om ziekenhuiskosten te bepalen en tracht deze te verbeteren (hoofdstuk 6); de *derde* studie gaat door op de analyse van de twee voorgaande hoofdstukken en bekijkt de determinanten van variabiliteit in de kostprijs tussen verschillende ziekenhuisafdelingen (hoofdstuk 7); de *vierde* studie gaat na welke determinanten de arbeidstijd beïnvloeden bij het aanbieden van IMCI in een ziekenhuis omgeving (hoofdstuk 8), waar potentieel bruikbare en meer valide alternatieven gegeven worden om kosteninformatie te extrapoleren wanneer lokale gegevens niet voorhanden zijn. Tot slot worden verscheidene toepassingen van de resultaten uit deze modellen gepresenteerd, bijvoorbeeld om de kosten te schatten van de "3 by 5"strategie, een wereldwijd gezondheidsinitiatief van de Wereldgezondheids Organizatie (WHO) om de sterfte als gevolg van HIV/AIDS te verminderen (hoofdstuk 9), en om op een consistente en vergelijkbare manier de kosten-effectiviteit te bepalen van interventies voor het bereiken van de gezondheidsgerelateerde Millennium Ontwikkelings Doeleinden. Dit laatste gebeurt zowel vanuit het perspectief van de ziekte/aandoening, bijvoorbeeld voor interventies voor kraamzorg en neonatale zorg (hoofdstuk 11), als vanuit een sectoraal perspectief waarbij de kosten-effectiviteit van interventies over verschillende ziektegebieden met elkaar vergeleken wordt (hoofdstukken 10 en 12).

In dit proefschrift zijn verschillende methodologische en praktische problemen aan het licht gekomen die te maken hebben met de oorzaken van variabiliteit in kostenmethoden, maar ook de gevolgen hiervan voor de generaliseerbaarheid en de externe validiteit van de verkregen resultaten. Beargumenteerd werd dat een deel van deze variabiliteit vermeden kan worden wanneer onderzoekers zich beter aan de aanbevelingen van richtlijnen zouden houden en wanneer redacties van tijdschriften strengere regels zouden toepassen bij het accepteren van manuscripten voor publicatie. Tevens is er een aantal gebieden geïdentificeerd waar nog richtlijnen ontbraken, en zijn er empirische studies verricht om sommige van deze onder de aandacht te brengen. Een voorbeeld hiervan was het gebruik van econometrische methoden om kostprijzen te bepalen in landen waar deze data niet voorhanden zijn, in plaats van het slechts op traditionele wijze extrapoleren van resultaten van voorgaande kostenstudies. De analyse toont aan dat voorzichtigheid is geboden bij het gebruik van de resultaten van een kostenstudie van slechts één ziekenhuis als basis voor kosten-effectiviteitsstudies of voor het bepalen van de kosten van het opschalen van interventies. In deze studies zouden heel goed kostprijzen gebruikt kunnen worden die niet representatief zijn voor het land als geheel, wanneer niet is nagegaan of het niveau van capaciteitsgebruik, de inzet van personeel en ligdagen, waarvan duidelijk is dat deze de ziekenhuiskosten beïnvloeden, representatief is.

Een ander belangrijk gevolg voor de overdraagbaarheid van resultaten is het bijhouden van capaciteitsgebruik en het corrigeren hiervoor wanneer kostprijzen geschat en gepresenteerd worden. Algemene beleidsbeslissingen zouden niet gebaseerd moeten zijn op de resultaten van kostenstudies waar geen rekening is gehouden met capaciteitsgebruik. Gezien de manier waarop gezondheidswerkers hun tijd die ze besteden aan het consulteren van patiënten aanpassen aan verhoogde werkdruk, zijn schattingen van kosten om interventies op te schalen niet bruikbaar voor beleidsmakers als ze alleen gebaseerd zijn op de huidige kosten van de zorgverlening.

Tenslotte zijn er bij bepaalde studies om ziekenhuiskosten te bepalen een aantal vuistregels gebruikt vanwege de hoge kosten en de complexiteit van 'step down' kostenstudies. Duidelijk werd dat het niet wenselijk is deze regels toe te passen op alle ziekenhuizen, omdat er voldoende bewijs is dat ze belangrijke verschillen in kostprijzen over ziekenhuizen en landen verdoezelen. Het feit erkennend dat het beste alternatief, namelijk 'step down' kostenberekening, niet altijd praktisch of betaalbaar is, suggereert dit proefschrift dat een minder kostbaar maar valide alternatief gebruikt kan worden, in het geval dat er voldoende studies bestaan waarbij de determinanten van de ratio van klinische ten opzichte van poliklinische kostprijzen econometrisch bepaald worden. De kostprijs van ziekenhuizen

"buiten-de-steekproef" kunnen vervolgens geschat worden door zo'n verhouding te gebruiken wanneer informatie over sleuteldeterminanten als bezettingsgraad aanwezig is.

A Word of Thanks

A word of thanks

There are lots of people I would like to thank for a variety of reasons.

Firstly, I would like to thank my supervisor and promotor, Professor Dr Frans Rutten. I could not have imagined having a better advisor and mentor for my PhD. His immediate attention, technical guidance, warm support and belief in my capacities were a tremendous help during the course of working on my thesis. I would also like to thank Dr Marc Koopmanschap, my co-promotor, for his inspiring and challenging comments and guidance throughout.

Thank-you to my examiners, Professors Eddy van Doorslaer, Dik Habbema and Julia Fox-Rushby for agreeing to assess my thesis and for the positive, generous and timely feedback.

David Evans, I had the privilege to work with you on several projects and I have learnt a lot from the interactions. Your trust in my abilities has given me the opportunity to develop myself and become more confident and independent. Many thanks for the very fruitful and inspiring discussions and the tireless and timely feedback.

Tessa Tan-Torres, many thanks for making it possible to accomplish all this work within my responsibilities in the CHOICE team in WHO. I valued the technical and personal interactions very much. Also thanks for the nice and intellectually stimulating working atmosphere.

I would also like to thank my co-authors for a very fruitful and enjoyable experience in each paper we published together. I particularly like to thank Chris Murray for his inspiring ideas and comments, Joanna Schellenberg for her constructive feedback and Stephen Lim for a very enjoyable and intellectually fulfilling working relationship and an exemplary collaborative attitude.

There are several colleagues who contributed to my statistical and analytical skills. Sincere thanks to all of you, particularly Xu Ke, Ajay Tandon and Amani Siyam, for the generous knowledge transfer and the time you spared to answer my questions in my search to get the analysis right.

I am very grateful to Jennifer Bryce who gave me the opportunity to join the Multi-Country Evaluation of IMCI and for her generous technical and administrative support throughout the years. This ensured that I had all I needed to continue learning and producing high quality products. I have learnt a lot from my interactions with the technical advisors and the country teams. I particularly like to thank Cesar Victora, Jean-Pierre Habicht, Bob

Black and Don de Savigny for their invaluable feedback and support while I was the coordinator of the cost component of the study. I also valued and enjoyed working with the country teams. Special thanks are due to David Bishai, Débora Amorim, Fatuma Manzi, George Pariyo, Shams El Arifeen, Shakil Ahmad, João Amaral and Luis Huicho.

There are so many current and former work colleagues to whom I am very grateful. I would like to mention especially Rob Baltussen, Dan Chisholm, Stephen Lim, Jeremy Lauer, Ben Johns, Raymond Hutubessy and Monica Ortegon. Thank you for all the stimulating discussion, support, chats and laughs on so many occasions! With Monica I could share the stress of the final stages of the thesis, for which I am very grateful. I am also grateful to the moral support and encouragements from my friends in WHO. Special thanks to Xu Ke, Sonia, Charu, Alice, Angelica, Mylena, Anna and Stanislava for being such good and caring friends.

I would also like to thank my friends all over the world for their warm support. Wessam, Azza, Akmal, Marwa, Sahar, Jean-Jacques, Nanda, Gary, Linda, Ahsan, Marlies and Uli are especially thanked for their loving care and attention.

Monica, many thanks for agreeing to stand beside me during the public defence of my thesis. I have learnt a lot about your very fine and responsible character while we were preparing for that day. Many thanks for being such a good friend and for your generous and warm help and support. Raymond, I am very pleased that you accepted to be the second person to stand beside me during the defence. You have introduced me to this PhD programme 6 years ago and have been always supportive and encouraging all the way through. I am pleased you will be there to share with me the most important moment of this work.

Diane, Paul and Raymond, a special thanks to all of you for your generous and prompt help with the Dutch translation of the summary of my thesis.

My father was always proud of every little success I had. I know you would have felt very happy to see this book and the outcome of this work. Thanks to all the support, understanding and love you gave me through out the years.

My mother has always supported my dreams and aspirations, and if I may say so myself, I think she did a fine job raising me ☺. I would like to thank her for her caring support, love and all she has done for me.

Finally, the loving care and understanding from my sister and her family makes me feel very lucky. Many thanks for your support and for being there for me whenever I need you.

Curriculum Vitae

Curriculum Vitae

Taghreed Adam studied Medicine at Cairo University, Egypt. As a paediatrician, she practiced medicine for three years in various private and public clinics in Cairo. She then joined the Health Sector Reform Project in Egypt where she participated in several projects, one of which was the development and costing of a basic benefits package for the universal health insurance of the Egyptian population. This experience highlighted issues of resource scarcity and was the stimulus for her to learn more about health economics, a much needed expertise in Egypt and many developing countries. In 2000, she started to work for her PhD thesis at Erasmus University, the Netherlands, and in 2003 she completed course work and obtained her diploma in health Economics at Curtin University, Australia.

Since 1999, she has worked in the World Health Organization, where she is contributing to both methodological and analytical developments in the economic evaluation of health interventions. She also helps to build evidence-based databases for the cost-effectiveness of a wide range of health interventions to assist in regional and country level priority setting for the allocation of scarce health resources.

Publications

Publications

1. Adam T, Evans DB and Koopmanschap MA. Cost-effectiveness Analysis: Can We Reduce Variability in Costing Methods? *International Journal of Technology Assessment in Health Care, 2003 19:2.*

2. Adam T, Manzi F, Armstrong Schellenberg J, Mgalula L, de Savigny D and Evans DB. Does the Integrated Management of Childhood Illness Cost more than Routine Care? Results from Tanzania. *WHO Bulletin, May 2005, 83(5) 369-77.*

3. Armstrong Schellenberg JRM, Adam T, Mshinda H, Masanja H, Kabadi G, Mukasa O, John T, Charles S, Nathan R, Wilczynska K, Mgalula L, Mbuya C, Mswia R, Manzi F, de Savigny D, Schellenberg D and Victora C. Effectiveness and Costs of Facility-based IMCI in Tanzania. *Lancet, 2004 (9445);364:1583-94.*

4. Adam T , Evans DB and Murray CJL, Econometric Estimation of Country-Specific Hospital Costs. *Cost Effectiveness and Resource Allocation Journal (CERA) 2003 1:3. Available online at:* http://www.resource-allocation.com/content/1/1/3

5. Baltussen RBM, Adam T, Tan Torres T, Evans DB and Murray CJL, Methods for Cost-Effectiveness Analysis; In: Making Choices in Health: WHO Guide to Cost-effectiveness Analysis 2003. Available online at: www.who.int/choice

6. Gutierrez JP, Johns B, Adam T, Bertozzi S , Tan-Torres Edejer T, Greener R, Hankins C and Evans DB. Achieving the WHO/UNAIDS Antiretroviral Treatment "3 by 5" Goal: What Will it Cost? *Lancet, Vol 364, July 3, 2004.*

7. Darmstadt GL, Bhutta ZA, Cousens S, Adam T, Walker N, de Bernis L; Lancet Neonatal Survival Steering Team. Evidence-based, Cost-Effective Interventions: How Many Newborn Babies Can We Save? *Lancet. 2005 Mar 12;365(9463):977-88.*

8. Adam T, Lim SS, Mehta SS, Bhutta ZA, Fogstad H, Mathai M, Zupan J and Darmstadt GL. Achieving the Millennium Development Goals for Health: Cost-effectiveness Analysis of Strategies for Maternal and Neonatal Health in Developing Countries. 2005 *BMJ*

9. Evans DB, Tan-Torres Edejer T, Adam T, Lim SS, and the WHO-CHOICE MDG Team. Achieving Millennium Development Goals for Health: Methods to Assess the Costs and Effects of Interventions for Improving Health in Developing Countries. *2005 BMJ*

10. Evans DB, Adam T, Tan-Torres Edejer T , Lim SS, Cassels A, Evans TG and the WHO-CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Time to Reassess Strategies for Improving Health in Developing Countries? 2005 *BMJ*

11. Evans DB, Lim SS, Adam T, Tan-Torres Edejer T for the WHO CHOICE MDG Team. Achieving the Millennium Development Goals for Health: Evaluation of Current Strategies and Future Priorities for Improving Health in Developing Countries. 2005 BMJ

12. Adam T, Amorim DG, Edwards SJ, Amaral JA and Evans DB. Capacity Constraints to the Adoption of New Interventions: Consultation Time and the Integrated Management of Childhood Illness in Brazil. *Health Policy and Planning 2005; 20 (Suppl. 1).*

13. Bryce J, Gouws E, Adam T, Black R, Schellenberg JA, Manzi F, Victora CG and Habicht JP. Improving Quality and Efficiency of Facility-based Child Health Care through Integrated Management of Childhood Illness in Tanzania. *Health Policy and Planning 2005; 20 (Suppl. 1).*

14. Manzi F, Armstrong Schellenberg JRM, Adam T, Mshinda H, Victora CG and Bryce J. Out of Pocket Payments for Under-five Hhealth Care in Rural Southern Tanzania. *Health Policy and Planning 2005; 20 (Suppl. 1).*

15. Bishai D and Adam T. Economics of Public Health Interventions for Children in Developing Countries. In Jones AM, ed. *The Elgar companion to health economics*, Cheltenham, UK & Northampton, Mass., USA: Edward Elgar Publishing, 2006.

16. Victora CG, Adam T, Bryce J and Evans DB. Integrated Management of the Sick Child. In: *Disease Control Priorities Project, 2006*

17. Adam T, Bishai D, Khan M and Evans DB. Methods for the Costing Component of the Multi-country Evaluation of IMCI. WHO Publication, 2004.

18. Adam T, Aikins M and Evans DB, CostIt software 2002. World Health Organization, www.who.int/choice .

19. Gaumer G, Adam T and El-Beih W. Rationalization Plan for Hospital Beds in Egypt. PHR report. Abt Associates.Inc, 1998.

20. NandaKumar, Frère JJ and Adam T. Costing the Basic Benefits Package in Egypt. PHR report. Abt Associates.Inc, 1998.

Propositions

Propositions

1. Variability in costing methods makes it very difficult to compare the results of different studies, diminishing their usefulness particularly outside their own setting and context (this thesis).
2. The wide variation in the unit costs estimated from studies within a particular country implies that analysts cannot simply take the cost estimates from a single study in a country to guide their assessment of the cost-effectiveness of interventions, or the costs of scaling-up activities (this thesis).
3. While the ideal method for estimating unit costs would be to conduct studies in a representative sample of facilities or activities, the econometric method presented in this thesis is more likely to provide a more accurate estimate of the average unit costs than relying on a single, or a few, cost studies in a given setting (this thesis).
4. An important implication for the transferability of the results is accounting for variations in capacity utilization in cost calculations and therefore these should be routinely reported in all published study (this thesis).
5. Estimates of the costs of scaling-up interventions will not be accurate, so not useful to policy-makers, if they are only based on the current costs of providing care as these are a function of current output (e.g., number of patient visits), and consultation time changes with changes in output (this thesis).
6. "Not everything that can be counted counts, and not everything that counts can be counted." Albert Einstein (1879-1955).
7. "Tell me and I'll forget; show me and I may remember; involve me and I'll understand." Chinese Proverb.
8. Great ideas are useless if they are not communicated effectively.
9. Evidence from cost-effectiveness analysis is only one element of the priority making process.
10. International work is a great opportunity to learn other ways of doing things.
11. Saying I do not know shows confidence, pretending you know shows insecurity.