WHO Technical Brief No.3 - DRAFT 3, June 12th 2008

SCALING UP HEALTH SERVICES: CHALLENGES AND CHOICES

"Scaling up to meet the need is equivalent to when a large group of people must use a bus to undertake a crucial journey. If the bus is too small, or it goes too slowly, or it takes a wrong turn, or its mechanical problems are not fixed, or it is badly driven, it won't reach its destination in time. Simply pouring in more fuel won't resolve these problems. Government and other players in the countries involved must deal with all the issues if the journey is to succeed." (quoted by Rivers)

Main messages

- *Scaling up* in the health sector means "doing something in a big way to improve some aspect of a population's health". It can be applied to scaling up inputs; outputs (access, scope, quality, efficiency); outcomes (coverage, utilization) or impact *(*reducing morbidity or mortality).
- Three cross-cutting considerations are also relevant to all scale-ups – sustainability, equity and the effects of scaling up an intervention (or a package of interventions) on the rest of the health system (external consequences).
- In its current usage, *scaling up* is often intended to convey haste, urgency and the need for a "special effort" – this is qualitatively different from "doing a bit more, but in the style of business as usual".
- More resources alone are rarely enough to ensure successful scale-up. There are many other kinds of constraints to be tackled, including unsupportive laws, weak management systems or limited demand from clients.
- Scaling up involves working on several fronts at once and making a number of strategic choices, including about the nature and number of the interventions being scaled up; the roles of various partners; and the equity, financing, speed and sequencing of scaling up.
- Whilst there are few certainties in scaling up, three generalizations are fairly robust:
    - Scaling up generally involves a *partnership* of organizations working on service delivery, financing and/or stewardship (co-ordination, regulation etc.).
    - Scaling up generally requires a highly committed group of individuals to push it along.
    - Monitoring implementation of the scale-up is crucial for assessing progress relative to overall objectives and for identifying aspects of the scale-up which are not working well. In practice, this is often a neglected aspect of scaling up.

World Health Organization

PART 1 INTRODUCTION AND MEANING

Introduction

Scaling up has become a much-used term in the 21st century for a number of reasons, including:

- *globalization* – an increasing awareness of global inequalities and the number of people without access to essential health services
- the *Millennium Development Goals* – to achieve the health-related MDGs by 2015 requires scaling up in its senses of both "big" and "urgent"
- a renewed interest in *primary health care* as the most appropriate vehicle for scaling up towards universal access
- the emergence of well-funded *global health partnerships* (such as the Global Fund and GAVI) which aim to significantly increase access to a specific range of interventions.

Scaling up is clearly a complex topic which raises questions on many levels. Is it best to concentrate on scaling up one intervention or technical programme at a time? What happens when there is pressure to scale up several interventions or programmes simultaneously? Why are there so many stories of relatively well-financed scale-ups encountering basic bottlenecks such as the disbursement of money or enough staff to perform simple administrative tasks? Is scaling up just about reaching as many people as fast as possible, or are there trade-offs with issues such as equity and sustainability?

Obviously this Technical Brief cannot include a comprehensive discussion of all aspects of scaling up. Instead the Brief focuses on:

- the objectives of different types of scaling up
- the importance of identifying constraints to scaling up
- critical choices which have to be made when scaling up (a selection of topical issues, rather than a comprehensive overview)
- identifying existing frameworks and tools to structure discussions about scaling up.

….whilst emphasizing throughout the overall "bigness" of scaling up, which involves working on multiple fronts in a manner that goes beyond "business as usual".

For simplicity, the Brief concentrates on scaling up of **health services**. This is clearly an over-simplification. For example, we know that a comprehensive strategy to reduce child mortality would involve work in the water, education and economic sectors, as well as health. Some of the examples in this Brief give a flavour of the range of possible activities in other sectors – see for example Boxes 3, 9 and 15.

Definition – what does *scaling up* mean?

The term *scaling up* is used in the health sector in a wide range of contexts, including:

- Scaling up **inputs** (government expenditure, the health workforce or the supply of pharmaceuticals).
- Scaling up the **provision of services** - any form of services, from hospital- to home-based. The expansion can be either a new or existing service; it can be a geographical spread or involve a new client group. This version of the term is frequently used in the context of single programmes, notably HIV/AIDS – but it can equally apply to a multi-programme package of interventions.
- Using existing inputs **more efficiently** (e.g. providing more services in health centres by re-organizing the use of staff time).
- Scaling up in order to produce better **outcomes** – e.g. to achieve the health-related Millennium Development Goals (MDGs).
- Scaling up **from a small project** to a much larger client group.

Moreover in its current usage, *scaling up* is often intended to convey **haste**, **urgency** and the need for a "**special effort**" – this is qualitatively different from "doing a bit more, but in the style of business as usual".

The multiple uses of *scaling up* are summarized in Box 1. This Brief focuses on scaling up **services and support systems**, rather than on scaling up resources. This is because there is an growing body of experience, especially from global health financing institutions like GAVI and the Global Fund, that more funds alone is not enough. As we will see in Part 2, different forms of scale-up involve different objectives. This is turn has implications for the equity and sustainability of a particular scaling up exercise, as well as consequences for other parts of the health sector.

Box 1 What does *scaling up* mean?

Scaling up in the health sector means "doing something in a big way to improve some aspect of a population's health". Within this broad definition, scaling up can take many guises:

Inputs / resources:
- mobilizing more funds; more staff

Outputs
- providing more services (**access, range of services available**)
- performing better (**quality**, **efficiency**)

Outcomes
- reaching more people (**coverage**)
- attracting more clients (**utilization**)

Impact
- reducing morbidity or mortality

Two cross-cutting issues can be applied to any of the above – **equity** and **sustainability** (whether the benefits will persist on a lasting basis).

Scaling up – smooth, stepped or great leap?

A useful analytic device is to "picture" scaling up. If scaling up is about radical change, it may not be enough to just do "more of the same". There may be institutional, legal or policy issues that need to be addressed before scaling up can proceed beyond a certain point.

The three graphs in Box 2 represent a useful mental device for thinking through how scaling up might develop in a particular context.

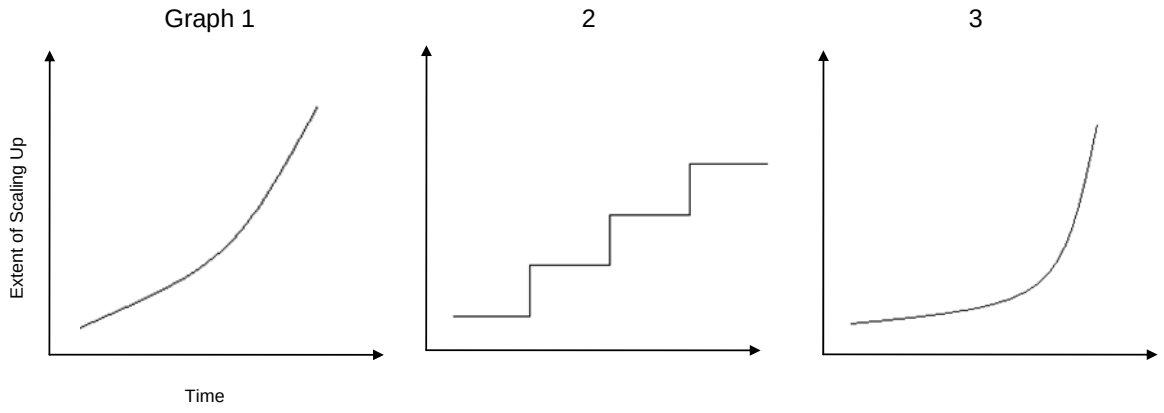
Graph 1 shows **smooth** scaling up – the underlying health system and the immediate environment are able to cope incrementally with more resources and more activities. Box 3 illustrates this with the reduction of maternal mortality in Sri Lanka.

Graph 2 shows a series of **steps** which have to be climbed if scaling up is to progress. Box 4 describes such a situation. Scaling up pneumonia treatment for children in Nepal encountered a number of delays as necessary steps were climbed.

Graph 3 – the **great leap** - shows a situation where a significant block needs to be surmounted. This is illustrated with the example of changes to the abortion law in South Africa (Box 5).

These graphs are obviously an over-simplification – and note that the three examples deal with very different lengths of time. Nevertheless, this mental device is a useful start to thinking through issues in scaling up – what blocks (or bottlenecks, or constraints) are there that could prevent a smooth scale-up?

Box 2 Scaling up – smooth, stepped and great leap



Box 3 Reducing maternal mortality in Sri Lanka – "smooth" scaling up

The maternal mortality ratio (MMR) in Sri Lanka was 2,136 per 100,000 live births in 1930; 486 in 1950; 121 in 1973 and 27 in 1996. Whilst Graph 1 is obviously not an exact depiction of this reduction, it is a stylized representation of the fact that the MMR decreased almost continuously over the period 1930-1996.

An analysis of MMR in Sri Lanka highlighted many "environmental" factors which facilitated the drop in MMR. These included a long-standing system for the civil registration of births and deaths; relatively high levels of female literacy; and a declining fertility rate (from the 1950s). This was complemented by a long history of training midwives in well-defined competences.

Broad service delivery strategies changed over time as the MMR dropped. To begin with, there was a focus on expanding access, especially in under-served areas. Later, the emphasis was on utilization and on removing financial and other barriers. More recently, quality of care has received close attention.

Key events and explicit "scaling up" phases do not feature prominently in the story of Sri Lanka's MMR. Rather it is a story of good health sector decision-making and implementation in a conducive environment. (Pathmanathan et al)

Box 4 The management of childhood pneumonia in Nepal: scaling up in steps

Access by Nepal's under-five population to community-based management of childhood pneumonia has increased through three phases since 1986. Research into effective interventions took place in one district in 1986-9, resulting in major publications in 1991. A programme in four districts tested the effectiveness of treatment by existing female community health volunteers from 1995 to 1997. In 1999, this became part of community-based IMCI (Integrated Management of Childhood Illness), meaning that community-based pneumonia case management could be part of routine annual programming for the first time. Moving from each phase to the next required formal Ministry of Health approval, which took time – this is scaling up in **steps**. (Dawson et al.)

Box 5 Abortion in South Africa – a legislative leap clears the way for a rapid increase in services

South Africa removed all restrictions limiting access to abortion during the first trimester by enacting legislation in 1997. This new law rapidly increased women's access to a broad range of options for the prevention and treatment of unwanted pregnancy, resulting in a substantial increase in abortion services throughout the country. The legislation played a significant role in the 91% decline in abortion-related deaths between 1994 and 2001. (Grimes et al.)

PART 2 SETTING OBJECTIVES AND IDENTIFYING CONSTRAINTS

Setting objectives

The way in which a particular scaling up exercise will be implemented depends on the objectives of the scale-up (see Box 1) and the perceived constraints. The objectives influence the nature of the scaling up activities. For example, a scaling up exercise with coverage as its prime objective

may be able to proceed faster than one focussing on quality. A scale-up aimed at improving health outcomes may have more of an incentive to reach high-risk clients than an exercise with coverage as its principal objective. And if sustainability is an explicit objective, there is added incentive to institutionalize the scale-up through measures such as supportive policies and budgets and including the intervention(s) in pre-service training curricula.

Objectives are about the internal ambitions of a particular scale-up exercise. Because scaling up is about significant change, it is also important to look at the effect on other parts of the health system. If one intervention or programme is significantly scaled-up, how does this affect other interventions or programmes and what is the impact on service delivery and overall health outcomes?

Boxes 14 and 16 demonstrate the importance of objectives. Box 14 describes the problems caused by neglecting the objective of sustainability in the context of universal child immunization. Box 16 illustrates how different objectives need to be balanced – in this case, balancing absolute numbers of clients receiving ART with the objective of equity.

Identifying constraints

Every scale-up has its own stories of constraints, or bottlenecks – areas which have slowed down progress. For example:

- The *disbursement of money*. Even when money is available, it can be difficult to set up as system of smooth-flowing disbursement to the districts, NGOs etc. which will eventually spend the money.
- Scaling up requires that a lot of people in a lot of places are well-informed about the relevant interventions. They may want to adapt the intervention(s) to suit their own local values or circumstances and need to know enough about the technical and financial aspects to be able to do this properly. This *communication* is often a challenge in scaling up (especially in federal and other decentralized states where local governments run health services and make significant resource allocation decisions).
- There may be limited *demand* to match the scaled-up supply. Immunization coverage cannot be scaled up if there is a widespread belief that immunizations cause unwanted side-effects; the use of interventions related to sexually-transmitted infections will be limited if there is a strong stigma attached to using such services.
- A policy or law may effectively block progress – for examples see Boxes 4 and 5.

An early stage in scaling up is to systematically identify such bottlenecks. Several frameworks and tools exist which may be helpful in the identification of bottlenecks and the related step of planning a scale-up. These frameworks and tools systematically specify the areas where potential bottlenecks lie – insufficient inputs; a lack of managerial or technical capacity; little political push or local ownership etc. Box 6 lists some potentially useful frameworks and tools, or it may be more appropriate to develop your own checklist.

Figure 1 illustrates one possible broad framework for the identification of both supply- and demand-side constraints. Questions which could be asked within the framework include:

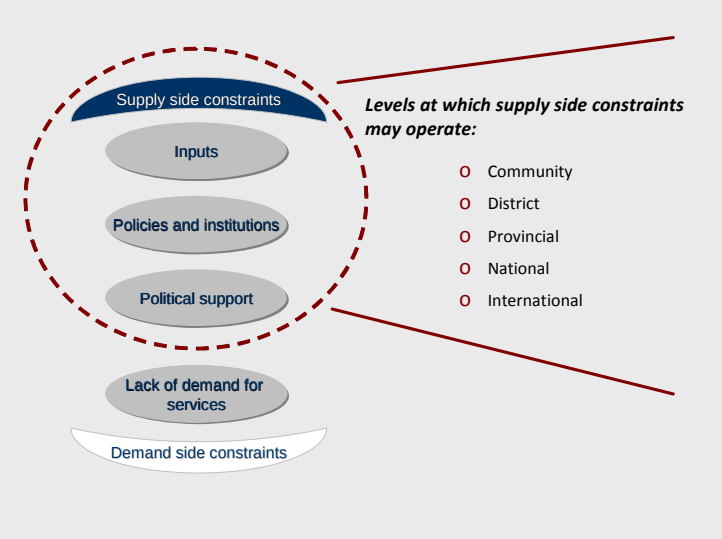
Inputs. What are the main inputs for the scale-up and in what quantities will they be required? (money, workforce, drugs etc.)

Policies and institutions. What public and/or private sector institutions will be involved in the scale-up? What laws, policies and rules do they operate under? Are these enforced? Do they have the requisite capacity and incentives to plan and implement? To whom are they accountable?

Political support Is there sufficient political support for the scale-up? In addition to politicians themselves, groups such as trade unions and women's associations may be important.

All these questions can be asked at the community, district, provincial, national and international levels. And underpinning all this, is there a demand for the intervention(s)?

Figure 1 Broad framework for identifying potential constraints to scaling up



Box 6 Frameworks and tools for thinking systematically about scaling up

There a various frameworks and tools that may help stakeholders to think through scaling up issues more systematically. No single tool or framework is perfect. Frameworks can, of course, be adapted to suit particular circumstances. Different tools address different aspects of scaling up.

Within the health sector, there are a frameworks that provide people with different ways of looking at constraints. One way is to look at the performance of each of a health system's core functions or 'building blocks', and their links to service outputs and outcomes - as can be done with WHO's health system framework. Other frameworks focus more on identifying constraints by level of the system, and the extent to which increased health sector funding can reduce different constraints, for example, the framework developed by Hanson et al. UNICEF's Marginal Budgeting for Bottlenecks tool is a more structured planning and budgeting tool designed to be used for maternal, newborn and child health interventions.

There are also tools from the more specific 'scaling up' literature. For example, the ExpandNet framework, which focuses on scaling up innovations, has developed a number of tools, including *Nine Steps for Developing a Scaling-Up Strategy.* The intervention complexity model of Gericke addresses scaling-up issues by looking at various dimensions of complexity – for example how complex are the human resource and management support requirements for a particular intervention?

There are some non health sector specific tools designed to help people more systematically navigate the political and institutional environment in which health systems operate, and mange change, which may also be useful. These include the Open Systems Model, Force Field Analysis and a number of Change Management tools. Tools such as the Everett Rogers model of diffusion of innovation can be used to think about potentially effective strategies. A useful overview of tools related to institutional development and change can be found in *Wilson et al.*

PART 3 DESIGN AND IMPLEMENTATION – MULTIPLE CHOICES

Scaling up generally involves changes in many parts of the health system. We saw in Box 6 that there are several frameworks and tools which may help in addressing the issues systematically. The exact nature of each challenge and choice obviously varies with circumstances and the objective(s) of a particular scale-up. This section discusses some of the most frequent and topical challenges and choices, related to the nature and number of the interventions being scaled up; the roles of various partners; and the equity, financing, speed and sequencing of scaling up. The Brief then ends with a reminder that scaling up entails working on many fronts at once, plus a section on the importance of monitoring and taking corrective actions.

The interventions to be scaled up

Single or multiple interventions?

Scaling up can involve anything from a single intervention to a broad package of activities. Single intervention scale-up may be driven by an urgent or new health need, the mandate of a funder, a local political issue, or a new technology which can be added to existing services. Whole-package scale-up often occurs because of a change in a country's political and/or economic circumstances – e.g. a new government with a real commitment to improve service provision for the poor.

There is no right or wrong answer to the question of single or multiple interventions – they entail different advantages and risks. A broad package may potentially be more cost-effective, but scaling it up across a population can be complex and slow. Scaling up a single intervention (or several related interventions within one programme) may be simpler to implement, but may have consequences for other parts of the health system. This is because "narrow" scale-ups often use shared resources, which are also involved in delivering other interventions (e.g. health workers, health facilities, drug distribution systems). These external consequences (the effects that extend beyond the actual scale-up) may be negative (the financial incentives attached to dealing with the scaled up intervention may not reflect its local epidemiological importance) or positive (a general improvement in laboratory services across the board). Box 7 gives some examples of external consequences related to human resources.

The issue is complicated when several narrow scale-ups happen simultaneously. If the scale-up is being done in a relatively short timeframe (as is often the case), the tendency is for each separate scale-up group to focus inwards on its own needs and goals. Without strong government leadership, this can lead to an unregulated competition for scarce resources, such as the time of doctors and nurses. Co-ordination of issues such as these is the focus of the Paris Declaration on harmonization and alignment - scale-ups should be aligned with government priorities. Box 8 applies this point about simultaneous single-issue scale-ups to maternal, newborn and child health services.

Box 7 Examples of "external consequences" related to human resources

"External consequences" are the effects (positive or negative) on other parts of the health system – i.e. not the part directly being scaled up. Examples of external consequences related to human resources and HIV/AIDS work are:

- Work on mainstreaming HIV/AIDS may involve a large percentage of the public sector health workforce in training activities. This is often appreciated – for its own sake and for the perks (such as per diems) which it brings. So scaling up HIV/AIDS activities has a positive effect on the wider workforce in this example.
- HIV/AIDS workers may have access to more financial incentives than workers in other technical areas – this can cause resentment and problems for trade unions.
- Workers in other technical areas may feel dispirited because issues they care about are neglected. For example a hospital doctor in Ethiopia said: "With the scale up of ART services in hospitals, patients with chronic illnesses like diabetes, hypertension and other internal illnesses are not given attention."
- In Malawi, a number of health surveillance assistants took the opportunity to train in voluntary counselling and testing. For the individuals, this was effectively a promotion – but the service as a whole lost experienced health surveillance personnel.

Source**:** Global HIV/AIDS Initiatives Network

Box 8 "Narrow" scale-ups can threaten the continuum of care

The expanding evidence base on single interventions (related to maternal, newborn, child and reproductive health) has resulted in many strategies for scaling up being designed around such interventions. Their implementation has certainly improved the situation in some areas, for example, under various child survival initiatives. However the approach also has encountered limitations. The narrow focus on specific elements has allowed specific programme targets to be reached, but has prevented other country priorities being addressed adequately, or achievement of high intervention coverage levels across a continuum of care, as demonstrated in the example from Kerber et al below:

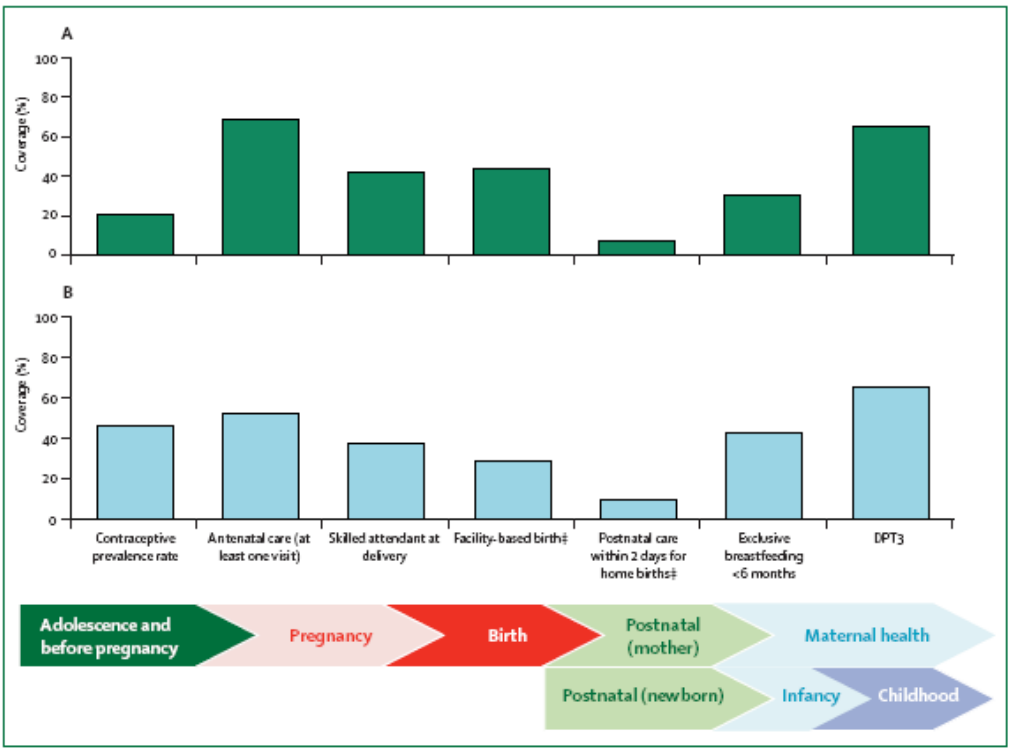


Figure 3: Coverage along the continuum of care in sub-Saharan Africa* (A) and South Asia† (B) between 2000 and 2006
Adapted from reference 5, which used data from Demographic and Health Surveys (DHS), 2000–2006,[15] with permission. *Sub- Saharan Africa includes Benin, Burkina Faso, Cameroon, Chad, Congo, Eritrea, Ethiopia, Gabon, Ghana, Guinea, Kenya, Lesotho, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Tanzania, Uganda, and Zambia; these countries have 74% of the region's annual births. †South Asia includes Bangladesh, India, and Nepal; these countries have 82% of the region's annual births. DPT3=three doses of diphtheria, pertussis, and tetanus. ‡DHS have assumed that all women who had a facility-based birth received postnatal care; therefore, only women whose most recent birth was outside a health facility were asked about a postnatal visit within 2 days.

Scaling up can also take place by grafting additional interventions on to existing service provision. This is called diversification. For example, in recent years outreach immunization services have added other interventions, such as Vitamin A.

Do some characteristics of an intervention help with scaling up?
Obviously many characteristics of the intervention(s) to be scaled up are fixed. Nevertheless, analysing certain features of the intervention(s) can inform a scaling up strategy:

- **evidence** about the benefits of the intervention(s). Good evidence – the more localized the better – helps with both the initial decision to scale up and with implementation, especially if there is a significant cadre of doubters or agnostics about the value of scaling up.
- existing **demand** for the intervention(s) – at the political, provider and client levels. Some interventions require more demand creation efforts than others. Box 9 describes aspects of Brazil's multi-pronged demand generation work related to HIV/AIDS.
- **compatibility with the norms and values** of the population and its health services. If a radical change is required in the basic way of doing things (rather than just a quantitative increase), scaling up may be slower and require more intensive efforts. Scaling up

innovations which challenge existing norms (about sexual behavior, for example) face particular challenges.

- Scaling up is easier if people can **easily observe quick results** – this helps to spread demand for further scaling up.
- **complexity** (in terms of technical and managerial requirements). Simplifying and standardizing interventions facilitates scaling up. Some interventions are by their nature *relatively* simple – for example immunization and social marketing. (Janovsky et al, 2006)
- The larger the overall **cost** of the scale-up relative to existing expenditures, the more work may be necessary to mobilize the funds or to persuade local decision-makers that the scale-up should be a national priority. This applies to costs to consumers too – see Box 10.

The understanding gained by answering the above questions may be useful for informing the scaling up exercise. For example:

- HIV/AIDS programmes – globally and nationally – have emphasized the importance of **standardizing and simplifying** protocols and procedures for HIV testing and counselling, prevention, the management of opportunistic infections and antiretroviral therapy. (WHO 2006)
- Box 4 describes the growth of community-based management of pneumonia in Nepal. The Technical Working Group promoting the intervention was convinced of its effectiveness by 1993. One of the main aims of the **second, larger, research phase** was to provide evidence to government officials who doubted that female community health volunteers could effectively manage pneumonia. It is not unusual to have such a demand for a second round of research with a larger population in more "normal" circumstances – people want to be convinced about local practicalities.
- The Malawian NGO Banja La Mtsogolo (BLM) has worked on scaling up health care in prisons. Recognising the unpopularity of the cause, the scaling up plans included **political-level advocacy** from the start. A Steering Committee was set up both to monitor the initial project and to be involved in advocacy for prisoners' rights. BLM worked with parliamentarians from both the Health and Legal Affairs Committees.

Box 9 Scaling up demand *and* supply – HIV/AIDS in Brazil

Brazil is perhaps best known in the field of HIV/AIDS for its work in reducing ART drug costs through domestic production and international negotiation. However there is also a range of activities aimed at increasing demand for HIV-AIDS-related products, as lyrically described by Okie:

"In Brazil this past February, during the week before Carnaval…… citizens who ventured out to catch a bus, buy a beer, or mail a letter were likely to be reminded by their government to use condoms. Postal consumers received condom brochures along with their stamps. Public health officials contracted with Coca-Cola distributors to deliver condom posters to bars along with the soft drink. In a television commercial on the country's most popular soap-opera network, a famous comedian riffed on strategies for remembering not to leave home without a condom. In the north-eastern city of Recife, banners on buses proclaimed, "On or off the float, *camisinha"* (literally, "little shirt", the street term for condom). During one lunch hour in Recife's

business district, a prostitutes' organization working with the local health department on prevention of HIV infection staged a pro-*camisinha* demonstration, passing out free condom samples to spectators."

There have been other demand-creation activities in addition to those described above:

- free HIV testing, backed up by a national media campaign which featured popular athletes, entertainers and models
- the 1996 law providing free ARTs to all eligible citizens

The demand-side work in Brazil is by no means all top-down. Activist groups are heavily engaged in HIV/AIDS work, reminding government of its constitutional promise of health care as a right for all Brazilians and encouraging high risk groups to demand their entitlement.

Brazil's scaled-up response to HIV-AIDS worked on demand and supply, on prevention and treatment. From 1996-2002, mortality from AIDS fell by 50% and AIDS-related hospitalizations fell by 80%. (Okie)

Box 10 Scaling up TB control in India – reducing costs to patients

Many poor patients in India use the private sector. Adapting DOTS treatment so that it could be provided through a public-private partnership (private provider, public subsidy and technical support) has substantially lowered the costs incurred by patients and their attendants in the private sector. Outcomes have also improved.

Another way of classifying the characteristics of interventions is to distinguish between the expansion of existing services and scaling up from a pilot demonstration project. There are at least four different "routes" for scaling up – pilot only at outset; pilot plus scaling-up plan at outset; phased implementation without pilot; and national scale at outset. (Janovsky et al) These routes differ in the degree to which they use evidence-based learning. Such structured learning can help to ensure that the intervention(s) fit properly into new situations as the scale-up progresses and that the intervention is technically effective.

Partnerships for scaling up - the roles of various partners

Whilst there are few certainties in scaling up, two generalizations about "who" are fairly robust:

- Scaling up generally involves a *partnership* of organizations working on service delivery, financing and/or stewardship (co-ordination, regulation etc.). Responsibility for these functions can be shared out in many different ways – it is easy to see how scaling up often involves quite complicated coalitions of government, private providers, national and international funders and agencies providing technical support. The ultimate responsibility for co-ordination and oversight obviously rests with national governments.
- Scaling up generally requires a highly committed group of individuals to push it along – a *driving team*.

This section concentrates on the role of the driving team and on the range of service providers which can be involved in scaling up.

The driving team

A driving team consists of institutions and individuals committed to the scaling up. The team may be more or less formally constituted and more or less formally working to a scaling up plan – whatever the exact situation, these are the champions and leaders of initiating and implementing scaling up. On paper, the list of organizations involved in the driving team and in delivery may look very similar. The key point is that driving and delivery are very different *functions*, which may or may not both occur in the same organization.

The driving team requires a variety of skills – the ability to win over local support and forge effective coalitions; competence in technical areas, management and training; and a talent for resource mobilization. To promote effective collaboration among scale-up partners, it may be beneficial for all the major ones to be represented on the driving team. The driving team associated with the increase in pneumonia management in Nepal (Box 4) came in the form of a technical working group which included government employees, local specialists and international agencies - UNICEF, USAID and WHO. Other driving teams may have little or no international involvement: in Thailand, the core coalition involved technical experts and government bureaucrats working with civil society groups and political parties to influence the adoption and implementation of reforms linked to universal coverage. (Mills)

Collaboration can be difficult, with institutional frictions and cumbersome processes to be navigated. Individuals on the driving team are often instrumental in smoothing over difficulties, as illustrated in Box 11.

Box 11 Smallpox eradication in India

In his historical account of the smallpox eradication programme in India in the 1960s and 1970s, Bhattacharya describes the roles of the health ministries of the central and state governments, WHO headquarters in Geneva and the WHO South East Asia Regional Office, all of which had special “eradication units”. He describes the complex interplay and changing coalitions amongst these organizations. Key WHO officials – in effect members of the **driving team** – expended huge energy on diplomatic contacts with the prime minister’s office, the states’ chief ministers and the federal and state health ministries. At times the tactic of using central government assistance to bring state employees into line in relation to smallpox targets was successful; at other times it backfired. Sometimes Prime Minister Indira Gandhi was approached directly – in violation of diplomatic protocol – because her support was deemed vital to progress with eradication.

Which delivery organization(s)?

'Delivery organizations' are those organizations which actually deliver the scaled up activities. They may be any combination of central and local government, private providers, social marketing organizations etc.

If scaling up is about radical growth, it should use all the suitable delivery organizations available. The choice varies with context – the private sector may have good logistics and be responsive to clients; NGOs may be appropriate when an intervention requires considerable local participation

and adaptation; starting with some local governments which have shown a particular interest has advantages; effective central government has a wide reach and responsibility for national policies. Box 12 illustrates how this notion of **comparative advantage** can be applied.

Box 12 ITNs: a role for the commercial and public sectors

The WHO Position Statement on ITNs (insecticide-treated mosquito nets) sees a vital role for the public financing of ITNs, but recognises that the type of organization with the best distribution system varies from country to country: "Where strong commercial markets exist or are developing, they should be encouraged: they can provide important benefits, ensuring longer-term access and enhancing management of logistics and education efforts".

This **comparative advantage** analysis has been put into practice in Tanzania, where there has been a targeted subsidy, plus support to the commercial sector from a social marketing programme. (Mulligan)

How the various types of organization affect each other is also relevant. For example, a significant increase in the subsidized distribution of an item through the public sector may cause a decline in private sector profits. This may or may not be a good thing, depending on the value of what the private sector was providing (fake drugs versus good-quality ITNs).

Can scaling up be equitable?

The hardest-to-reach populations are generally the last to benefit from scaling up exercises, unless deliberate efforts are made to hasten their inclusion. (Populations can be hard-to-reach for geographical, economic and/or social reasons.) These extra efforts may be justified on epidemiological grounds (hard-to-reach groups which are also high-risk) or simply on the grounds of fairness. Sometimes a scale-up may deliberately exclude equity from its short-term objectives, arguing in favour of "as many people as possible, as quickly and as cheaply as possible".

In immunization, the Reaching District (RED) approach tackles this very issue of scale up-with-equity. RED aims to improve immunization systems in districts with low coverage. With the RED approach, countries use coverage data to analyse the distribution of unimmunized infants and prioritize districts with poor coverage. Districts are encouraged to make micro-plans to identify local problems and adopt corrective solutions. In 2005, an evaluation of 5 countries in Africa that had implemented RED found that the proportion of districts with DTP3 coverage above 80% had more than doubled. (Vandelaer et al). Box 16 describes some equity issues which were raised in the evaluation of the "3 by 5" Initiative.

How can scaling up be financed?

Not all scale-ups require significant financial outlays – at the extreme, fiscal policy changes to increase the price of tobacco or alcohol involve no net government expenditure. But most scaling up in the health sector is expensive. Moreover, because scaling up is by definition not routine, some dedicated resources are necessary until implementation becomes standard practice and costs are financed through routine budgets.

Significant scaling up is generally financed through taxation, social health insurance (SHI), public-private infusions of money or international aid (including global health partnerships). Each brings its own challenges. Tax- or SHI-financed scaling up requires a favorable mix of circumstances in the national economic, labor and political scenes, but generally comes with ready-made sustainability and institutionalization. External financing raises issues of local ownership, sustainability and (if there are multiple funders) aid effectiveness. Some external financing may be earmarked in a way which effectively dictates what can be scaled up.

How Fast Can Scaling Up Happen?

We have seen how the current usage of *scaling up* often implies a fast pace of change. Impressive improvements can be achieved in a short time. In Rwanda, for example, a particular form of performance-based funding was introduced in 23 districts in 2006. At the end of two years, there had been a 173% increase in family planning users, a 90% increase in second dose malaria medications in ante-natal care and a 35% increase in utilization of curative care services. In contrast, Box 13 describes how long it took various countries to reach the point of legislation for universal coverage and some of the factors which determined the pace of change. This is a reminder that in some circumstances scaling up is a long-term undertaking.

There is clearly a balance to be struck between the achievements of rapid or explosive scaling up and the advantages of gradual scale-up. The current challenge is to use the policy windows afforded by the MDGs and the rise of new public-private financing mechanisms such as the Global Fund, but without responding so fast that quality or sustainability are unacceptably compromised. Problems arise when different parts of the scale-up are moving at different speeds. This is illustrated in Box 14, which describes the challenge of *maintaining* high levels of coverage reached during a period of planned scale-up.

Box 13 Scaling up towards universal coverage – a range of timescales

Countries have differed greatly in how long it took from the first health insurance law until legislation was passed to implement universal coverage (which is the ultimate scale-up). In Germany - a pioneer in health insurance in an era with very much less communications technology to support scale-up - it took 127 years, in Japan 36 years. 26 years elapsed in the Republic of Korea; 20 in Costa Rica. A number of factors were associated with a faster speed of change:

- Higher per capita income
- A larger percentage of the workforce in the formal sector
- More urbanized populations and higher overall population densities.
- Good national administrative capacity
- Supportive social values
- Good stewardship from government, including open debate about policies relevant to the scaling up. This fostered the population's trust in government and other agencies involved in the scale-up. (Carrin)

Box 14 Was Universal Child Immunization achieved too quickly?

History provides some lessons about the pace of scale-up. A 1996 review of the lessons learnt from Universal Child Immunization emphasized the importance of the *maintenance* of high immunization rates and described how this could be made difficult by inappropriately fast scaling up: "When targets have been set at unrealistically high levels, they have led to the development of unsustainable immunization strategies and to the manipulation of data. When their political use has taken precedence over their use in managing programmes, they have deterred effective critical assessment. Immunization targets should be ambitious but attainable".

What do we know about sequencing scale-ups?

A "frequently asked question" in relation to scaling up is – *Is there a correct sequence of scaling up activities?* Should systems be strengthened and then interventions scaled up through these systems, or are the interventions themselves the starting point for scale-up?"

There is, of course, no one answer to this question – as ever with scaling up, it depends on the circumstances. Scaling up in health generally relies on some blend of effective systems (public or private) and "special efforts" related to the activity which is being scaled up. (This is the bus and the fuel of the quotation at the beginning of the Brief.)

A number of questions can be applied to help think systematically about the issue of sequencing:

- What parts of the overall system does this scale-up absolutely need and what is the quickest way to establish/ strengthen them?
- Which bottlenecks can be solved in the short term?
- What else can happen at the same time as addressing medium-term bottlenecks?
- What opportunities or entry points are there? Sometimes a timely response to an emerging opportunity can provide a significant boost to a scale-up.

Scaling up is not a matter of systems first, then interventions – or vice versa. It is about developing a practical timetable of activities, then and adjusting if needed as one goes along.

The scaling up strategy

Having identified constraints and considered key choices, a scaling up strategy can be developed. This is an overall description of how scaling up will be implemented. How will it be financed and organized? How will challenges be overcome and opportunities exploited? How fast will it be, with what role for evidence and monitoring?

As we have seen, scaling up generally involves working on several fronts at once – with a variety of organizations, often spending money from several sources and with a range of activities (legislation, policy, advocacy, institutional change, procurement, training etc.) The scaling up strategy is an opportunity to ensure that work is planned in all the relevant areas and directions.

In particular, it is important to think both about "doing more" (usually expanding the availability of an intervention geographically or to new client groups) and about "institutional scale-up" (changes related to laws, policies, budget lines, regulation etc.).[1]

Box 15 describes work in Mbeya, Tanzania, where a particular point was made of scaling up the *scope* of activities to include all the interventions specified in the National Multisectoral Strategic Framework for HIV/AIDS.

Box 15 HIV/AIDS in Mbeya, Tanzania – a multi-pronged scale-up

The Mbeya Regional AIDS Control Programme is described as part of UNAIDS' *Best Practice Collection*. As well as concentrating on access and coverage, the Programme concentrated on continually expanding the range of activities related to HIV/AIDS, in order to be as comprehensive as possible. The Programme encouraged activities which included ART expansion, STI management, peer education in schools, workplace programmes in the private and public sectors, home-based care supported by government and NGOs, economic support for affected families, political advocacy and tackling stigma.

Monitoring progress

Monitoring implementation of the scale-up is crucial for assessing progress relative to overall objectives and for identifying aspects of the scale-up which are not working well. Scaling up in practice is a continuous stream of decision-making about how to deal with chronic and new constraints to further scale-up.

Exactly what is monitored depends on the objectives of a particular scale-up. This link between objectives and indicators is important. Box 16 describes how assessing whether or not a scale-up is equitable requires specific indicators.

As we saw above, scaling up can have external consequences – effects beyond the activities which are themselves being scaled up. Monitoring the effects of scaling up is thus not just an internal responsibility – it is also part of the wider stewardship function of governments and co-ordinated development partners.

Despite its importance, monitoring is often a neglected aspect of scaling up. A review of innovative strategies in 12 countries noted how rarely routine monitoring was used to inform scaling up. This absence of monitoring was even more extreme for strategies that cut across programmes where monitoring was organized around those programmes. (Janovsky et al)

The situation is similar for research, particularly operations research. Whilst some scaling up starts with a well-researched pilot phase, it is much rarer to see a national scale-up which is regularly informed by operations research.

[1] The scaling up literature calls these "horizontal" (expansion) and vertical (institutionalizing) scaling up. This is a different meaning to *vertical* than in the term *vertical programme.*

Box 16 The "3 by 5" Initiative: a lack of disaggregated data to monitor equity

An evaluation took place in 2006 of "3 by 5", the plan to treat 3 million people living with HIV/AIDS in low- and middle-income countries by 2005. The evaluation noted that a significant scale-up had occurred – the number of people on antiretroviral therapy in these countries nearly doubled in 2005 alone.

However the evaluation also reported many inequities. Equity was not a routine consideration when service delivery was being planned – indeed many countries did not collect the information which would be required to monitor equity. The evaluation stressed the importance of dis-aggregated sub-targets – disaggregated by gender, age, region etc. The evaluation acknowledged the balance which needs to be struck between equity and absolute numbers:

"It is important to emphasize that too much focus on equity in the beginning of a scale-up may drastically limit the expansion and number of people who have access. Starting from well-equipped urban centers with trained personnel is important in order to gain experience and develop contextualized treatment systems, and should not be criticized for lack of equitable coverage and distribution. It is, however, important for services to be designed in such a way that they can be rapidly rolled out to low-resourced rural settings."

The main messages are presented at the start of this Brief.

References and bibliographic note

Banja La Mtsogolo (2007) Annual Report 2006. And website accessed 18 May 2008, www.banja.org.mw

Battistella Nemes MI et al. (2006) *Evaluation of WHO's contribution to 3 by 5: main report.* WHO.

Bhattacharya S. (2004) *Uncertain advances – a review of the final phases of the smallpox eradication program in India, 1960-1980.* American Journal of Public Health 94 (11): 1875-1883.

Carrin, G and James, C. (2004) *Reaching Universal Coverage via Social Health Insurance*, *key design features in the transition period.* EIP Discussion Paper Number 2. WHO.

Dawson et al. (2008) *From research to national expansion: 20 years' experience of community-based management of childhood pneumonia in Nepal.* Bulletin of the World Health Organization 86:339–343.

ExpandNet/WHO. (2007) *Nine steps for developing a scaling up strategy.* ExpandNet/WHO. Draft. http://www.expandnet.net/tools.htm

Floyd K et al. (2005) *Cost and cost-effectiveness of PPM-DOTS for tuberculosis control: evidence form India.* Bulletin of the World Health Organization 84:437–445.

Gericke, CA et al. (2005) *Intervention Complexity – a Conceptual Framework to Inform Priority-Setting in Health.* Bulletin of the World Health Organization, 83 (4).

Global HIV/AIDS Initiatives Network. (2008) *Briefing Sheet 2: Human Resources.*

Grimes, DA et al. (2006). *Unsafe abortion: the preventable pandemic.* The Lancet, 368: 1908-19.

Hanson K et al. (2003) *Expanding access to priority health interventions: a framework for understanding the constraints to scaling-up.* Journal of International Development 15: 1-14.

Janovsky K and Peters D (2006) *Improving health services and strengthening health systems: adopting and implementing innovative strategies – An exploratory review in 12 countries.* Making Health Systems Work: Working Paper No. 5. WHO/EIP/health systems/2006.2

Mills, A. (2007) *Strategies to achieve universal coverage: are there lessons from middle income countries?* Health Economics and Financing Programme, London School of Hygiene and Tropical Medicine.

Mulligan J-A et al. (2008) *Costs and effects of the Tanzanian national voucher scheme for insecticide-treated nets.* Malaria Journal 7:32.

Okie, S. (2006) *Fighting HIV – Lessons from Brazil.* 354:19, 1977-81.

Pathmanathan, I et al. (2003) *Investing in Maternal Health - Learning from Malaysia and Sri Lanka.* World Bank.

Rivers, B. (2008). Scaling Up to Meet the Need: Overcoming barriers to the development of bold Global Fund-financed programs. Aidspan White Paper

Simmons, R, Fajans P and Ghiron L (2007). *Scaling up health service delivery from pilot innovations to policies and programmes.* WHO/ExpandNet.

Steering Committee for Lessons Learned on Sustainability for Child Immunization. (1996) *Sustainability of achievements: lessons learned from Universal Child Immunization.* UNICEF.

Travis P et al. (2004) *Overcoming health-systems constraints to achieve the Millennium Development Goals.* Lancet; 364: 900-6.

UNICEF (2008). *The State of the World's Children 2008: Child Survival.*

Vandelaer, J et al. (2008) *Reaching Every District (RED) approach: a way to improve immunization performance.* Bulletin of the WHO, 86 (3)

Vogel, UF. (2007) *Towards universal access to prevention, treatment and care: experiences and challenges from the Mbeya region in Tanzania – a case study.* UNAIDS (Best Practice Collection).

WHO (2006). *Towards universal access by 2010: How WHO is working with countries to scale-up HIV prevention, treatment, care and support.*

WHO (2007). *Insecticide-treated mosquito nets: a WHO Position Statement.*

WHO (2007) *Everybody's business. Strengthening health systems to improve health outcomes. WHO's Framework for Action.*

Wilson D and Beaton L. (2003) *Promoting Institutional and Organisational Development: A Source Book of Tools and Techniques.* DFID. www.dfid.gov.uk/pubs/files/prominstdevsourcebook.pdf

A note on the scaling up literature

The scaling up literature comes in many guises, including historical accounts, readings on universal coverage, evaluations of global scaling up exercises and the "innovations" literature, which looks at how to encourage the adoption and dissemination of effective innovations. Other sectors – notably agriculture and rural development - also have rich literatures on scaling up. The work of ExpandNet is of particular relevance to this Brief - see http://www.expandnet.net. ExpandNet is a network of public health professionals and scientists which focuses on scaling up health service innovations which have been tested in experimental, pilot and demonstration projects. The website includes guidance tools to assist countries with scaling up.