

HPG Research Report

Measuring the impact of humanitarian aid

A review of current practice

Researched, written and published by the Humanitarian Policy Group at ODI

Charles-Antoine Hofmann, Les Roberts,
Jeremy Shoham and Paul Harvey

About HPG

The Humanitarian Policy Group at the Overseas Development Institute is dedicated to improving humanitarian policy and practice. It conducts independent research, provides specialist advice and promotes informed debate.



Britain's leading independent
think-tank on international development
and humanitarian issues

Overseas Development Institute
111 Westminster Bridge Road
London SE1 7JD
United Kingdom

Tel. +44 (0) 20 7922 0300
Fax. +44 (0) 20 7922 0399

E-mail: hpg@odi.org.uk
Websites: www.odi.org.uk/hpg
and www.odihpn.org

The authors

Charles-Antoine Hofmann is a Research Officer at the Overseas Development Institute.

Les Roberts is an independent consultant.

Jeremy Shoham is a nutritionist. He is co-editor of *Field Exchange* and technical consultant for the Emergency Nutrition Network.

Paul Harvey is a Research Fellow at the Overseas Development Institute.

Acknowledgements

ODI would like to thank the wide range of individuals and organisations that assisted and supported this study. Les Roberts and Jeremy Shoham wrote background papers for the study, co-authored this report and provided invaluable comment and support throughout. The Feinstein International Famine Center at Tufts University provided research assistance. In particular, Sue Lautze organised a series of meetings with academic staff and students in Boston and Gregory Lon Scarborough, a student from Tufts University, contributed to the literature review. ODI also wishes to thank Anisya Thomas from the Fritz Institute for her comments on earlier drafts of the report.

USAID, AusAID, WFP and The Fritz Institute provided particular support to the study, and a wide range of donors contributed through HPG's Integrated Programme.

Many thanks also to the peer review group which commented on a draft of the report: Lola Gostelow (Save the Children UK), Andre Griekspoor (WHO), Allan Jury (WFP), John Mitchell (ALNAP), Chris Roche (Oxfam), Egbert Sondorp (London School of Hygiene and Tropical Medicine), Paul Spiegel (UNHCR), Gerhard Wächter (ECHO), and Howard White (World Bank).

Thanks also to the many other people who contributed through commenting on earlier drafts, and giving up their time to be interviewed. Finally, thanks to colleagues in the Humanitarian Policy Group at ODI, particularly Joanna Macrae, for invaluable help in drafting and editing the report.

Humanitarian Policy Group
Overseas Development Institute
111 Westminster Bridge Road
London
SE1 7JD
United Kingdom

Tel: +44(0) 20 7922 0300
Fax: +44(0) 20 7922 0399
Website: www.odi.org.uk/hpg
Email: hpgadmin@odi.org.uk

ISBN: 0-85003-718-2

© Overseas Development Institute, 2004

Contents

Executive summary	1
Chapter 1 Introduction	5
1.1 Scope and methodology	6
Chapter 2 The impact of humanitarian aid: definitions, objectives and context	7
2.1 Definitions	7
2.1.1 Defining 'impact'	7
2.1.2 The origins and uses of impact assessment	8
2.2 The levels and objectives of impact analysis	8
2.2.1 Conceptual frameworks	8
2.2.2 The levels of impact analysis	9
2.2.3 The functions of impact analysis	10
2.2.4 Assessing negative impacts	11
2.3 Perspectives from new public management	12
2.3.1 Perspectives from the private sector	14
2.4 Chapter summary	14
Chapter 3 Measuring and analysing impact: methods, indicators and constraints	15
3.1 Methods for measuring and analysing impact	15
3.1.1 Participatory approaches	16
3.2 Indicators	17
3.2.1 Mortality	18
3.2.2 Surveillance systems	19
3.2.3 Surveys	19
3.3 Constraints to measuring impact	19
3.3.1 The emergency context and the nature of the system	19
3.3.2 Lack of consensus over the objectives of humanitarian aid	20
3.3.3 Lack of baseline data and seasonality issues	20
3.3.4 Control groups and regression analysis	21
3.3.5 Problems of attribution	22
3.4 Impact assessment beyond the project level	22
3.5 Measurement or analysis?	23
3.6 Chapter summary	24
Chapter 4 Impact assessment: current practice in the health and nutrition sectors	25
4.1 Impact assessment and the health sector	25
4.2 Impact assessment and the food and nutrition sector	26
4.2.1 General ration programmes	27
4.2.2 Selective feeding programmes	27
4.2.3 Interpreting nutritional status data	28
4.2.4 Micronutrient status	28
4.2.5 Livelihoods approaches	29
4.3 Chapter summary	30
Chapter 5 Conclusions and recommendations	31
5.1 Beyond the project level	31
5.2 Measuring, analysing or demonstrating impact?	31
5.3 The evidence-base of humanitarian aid	32
5.4 The new public management agenda	32
5.5 The way forward: approaches to impact assessment	32
References	35

Executive summary

This report investigates the current state of the art in measuring and analysing the impact of humanitarian assistance. It is concerned with questions around how impact can be measured, why this is increasingly being demanded, and whether it is possible to do it better. It also explores the benefits, dangers and costs that paying greater attention to impact might entail.

Although questioning the impact of humanitarian assistance is not new, it has moved up the humanitarian agenda in recent years. As the overall volume of humanitarian assistance has increased, so there has been greater scrutiny of how this money is spent, while reforms within the West's public sectors have seen the introduction of new management systems focusing on results (Macrae et al., 2002). Several UN agencies, donors and NGOs are developing results-based management systems, and investing considerable resources in them, partly in an effort to demonstrate impact more clearly.

This increasing pressure to show results has yet to translate into clear improvements in the measurement or analysis of impact. Assessment of impact is, in fact, consistently poor. There are, of course, many good reasons why it is difficult to measure the impact of humanitarian interventions, including difficult issues of causality and attribution and a lack of basic data, such as population figures. Relief interventions are often of short duration, capacity and resources are stretched, insecurity may limit access to populations and the space for analysis and research is constrained. Nor is the new emphasis on results without costs of its own: within the humanitarian sector, a focus on measurement could reduce operational effectiveness and lead to the neglect of issues such as protection and dignity because they are difficult to measure. Focusing on what is measurable risks reducing humanitarian aid to a technical question of delivery, rather than a principled endeavour in which the process as well as the outcome is important.

These difficulties and risks do not, however, mean that impact cannot be measured in some circumstances; where measurement in a scientific and quantifiable sense is not possible, impact can still be analysed and discussed. Indeed, the scientific tradition and more participatory and analytical approaches should not be seen as polar opposites, but as complementary approaches.

Definitions, objectives and context

The question of impact within the humanitarian sector has been addressed in three main ways:

- The analysis of likely impact before the start of a project, in order to anticipate the wider consequences of an intervention.
- Ongoing analysis of impact throughout a project or as part of management systems, in an attempt to adapt interventions or monitor performance.
- Analysis of the impact of interventions after the fact, as part of evaluations or research. Impact is used as a key criterion in the evaluation of humanitarian work, and most evaluations consider it.

Impact can be analysed at the level of individual projects, and at much broader organisational or country-wide levels. Attempts to measure impact can restrict their focus to the intended effects of interventions, or they can encompass broader indirect and unintended consequences.

There is no accepted definition of 'impact' within the humanitarian sector, and the definitions current within the development field, though adopted for use in the humanitarian sector, may not fully capture the particular nature of humanitarian work. In particular, the concept of change is central in developmental definitions of impact, but in humanitarian aid the aim is often to avert negative change (for example to prevent famine), rather than bring about a positive change. This may be harder to measure. This report, though using existing developmental definitions, uses them with due caution.

The humanitarian system's increasing interest in impact needs to be understood in the context of broader debates about accountability in humanitarian aid, and against the background of public management reforms within Western governments. A central element of this reform is the shift from an input-output management model towards a greater emphasis on results. Service providers not only report progress in implementing activities, but must also demonstrate that they generate some achievements. Experience from the introduction of results-based or performance management systems within Western governments suggests a need for caution in adopting these approaches uncritically. The analysis of impact should not, therefore, be seen purely as a narrow technical question about the effectiveness of individual projects; discussion about impact should not be confined to a sub-set of evaluation techniques.

Measuring and analysing impact

Analysing the impact of a humanitarian intervention is not straightforward. A number of generic methodological

constraints and factors particular to humanitarian action make impact measurement difficult. The difficulties of the operating environment, the need to act quickly in situations of immediate crisis, an organisational culture that values action over analysis and the fact that there is little consensus around the core objectives of humanitarian aid – all these issues make analysing impact difficult. This does not mean that progress is not possible; even where impact cannot be formally measured, it may be possible to generate useful analysis. In particular, lessons could be learned from participatory approaches in the development sector.

There are broadly three main approaches to impact assessment (Hallam, 1998):

- the *scientific approach*, which generates quantitative measures of impact;
- the *deductive/inductive approach*, which is more anthropological and socio-economic in its methods and approach; and
- *participatory approaches*, which gather the views of programme beneficiaries. Participatory approaches are widely recognised as a key component in understanding impact, but have rarely been used in the humanitarian sector.

Within these broad approaches, there is a huge array of tools for analysing impact, often divided between qualitative and quantitative. These include surveys; interviews, workshops and discussions; direct observation; participatory research; and case studies (Roche, 1999). The identification and use of relevant indicators is a crucial part of determining the impact of an intervention. Although the terminology sometimes varies, there are generally two types of indicator: those that relate to the implementation of programmes (input, process and output indicators); and those concerned with the effects of the programme (outcome and impact indicators). Humanitarian agencies tend to use a mix of indicators, depending on their own monitoring and reporting systems and the particular function of the indicators collected. Since the core of the humanitarian agenda is about saving lives, mortality rates seem to be a logical starting point for the analysis of impact. However, there is no standard accepted method for measuring mortality rates.

Taken as a whole, the humanitarian system is poor at measuring or analysing impact, and the introduction of results-based management systems in headquarters has yet to feed through into improved analysis of impact in the field. Yet the tools exist: the problem therefore seems to be that the system currently does not have the skills and the capacity to use them fully. This suggests that, if donors and agencies alike want to be able to demonstrate impact more robustly, there is a need for greater investment in the skills and capacities needed to do this. Given the large (and

rising) expenditures on humanitarian assistance, it is arguable that there has been significant under-investment in evaluation and impact analysis. Many of the changes identified in this study would have wider benefits beyond simply the practice of impact assessment: greater emphasis on the participation of the affected population, the need for clearer objectives for humanitarian aid, more robust assessments of risk and need and more research into what works and what does not would be to the advantage of the system as a whole.

Conclusions and recommendations

Reviewing the literature and practice in analysing impact can be an exercise in pessimism; the methodological and practical difficulties seem so great that it is tempting to conclude that expecting meaningful analysis is unrealistic, at least in the humanitarian sphere. It certainly seems undeniable that the humanitarian system has not to date been particularly good at analysing impact. However, that does not mean that improvement is not possible, though it is likely to require greater commitment on the part of both donors and agencies. This study concludes with a series of recommendations for how impact analysis can be improved.

Moving beyond the project level

- Concern for the impact of humanitarian aid should not be narrowly restricted to the project level. There is a need for greater investment in system-wide evaluations that can ask difficult and important questions about the responsibility for humanitarian outcomes, and the broader political dimensions within which the humanitarian system operates.
- Project-based approaches that focus on determining the impact of a particular intervention through a causal pathway from inputs to impact should be complemented by approaches that start with changes in people's lives, and that situate change in the broader external environment.
- Questions of impact should not be limited to the evaluation process. In the humanitarian sphere, a concern with change in the short term implies a need for impact to be considered in ongoing monitoring processes, and through techniques such as real-time evaluation.

Results-based management: potential and dangers

- It is too early to say whether the introduction of results-based management in humanitarian organisations will significantly improve the measurement and analysis of impact. Experience from elsewhere suggests that there will be a need for caution; in particular, measurement may remain largely focused on outputs and not impact.
- An increased focus on results brings with it a risk that harder-to-measure aspects of humanitarian action, such as protection and the principles and ethics that

underpin the humanitarian endeavour, could be neglected.

- There may be room for humanitarian actors to explore further the potential for learning from experience in the private sector.

Measuring impact: skills, capacity and resources

- Impact in any context is difficult to measure and attribute; this difficulty is exacerbated in the dynamic and chaotic environments of complex emergencies. This does not mean, however, that it is impossible, and greater efforts could be made.
- The humanitarian system currently lacks the skills and capacity to successfully measure or analyse impact. Greater investment needs to be made in human resources and research and evaluation capacity if the desire to focus more on results is to be realised.

Approaches to impact: science, analysis and participation

- The humanitarian system has been consistently poor at ensuring the participation of affected populations. This is as true in impact analysis as in other aspects of the humanitarian response. Much could be learnt from

innovations in participatory approaches in the development sphere, and possibly from customer-focused approaches in the private sphere. The humanitarian system remains largely ignorant of affected people's views about the assistance being provided.

- There is a place for both art and science in impact measurement: scientific, analytical and participatory approaches can often be complementary.

Indicators and objectives

- Analysis of impact could be improved through greater clarity about the objectives of humanitarian assistance, and by more consistent assessment of needs.
- Process indicators can sometimes be used as proxies for impact when there is strong evidence of a link between the action being monitored and an expected impact. An example would be measles vaccinations, which are known to reduce mortality. There is a need for greater investment in strengthening the evidence base for how other activities, such as supplementary feeding or support to health clinics, relate to humanitarian outcomes such as reductions in mortality or malnutrition.

Chapter 1

Introduction

This report investigates the state of the art in measuring and analysing the impact of humanitarian assistance. It is concerned with questions around how impact can be measured, why this is increasingly being demanded and whether it is possible to do it better. It also explores the benefits, dangers and costs that paying greater attention to impact might entail. This subject is not new: both donors and the providers of humanitarian aid have always wanted to know about the impact of the assistance that they provide. It has, however, become increasingly urgent in recent years. ‘Results count,’ Andrew Natsios, the Administrator of USAID, told NGOs in 2003:

And if you cannot measure results, if you cannot show what you’ve done, other partners will be found. Why is that? Doing good is not enough. We have to show what kind of good we’re doing, in which sectors, in which communities, and whether the good has bad consequences, or bad side effects, that no one anticipated.

Good intentions, Natsios was saying, are no longer sufficient: agencies are being asked to demonstrate that they are achieving positive impacts.

This greater interest in impact analysis arises from a number of linked developments. As the overall volume of humanitarian assistance has increased, so there has been greater scrutiny of how this money is spent, and how effectively, while reforms within the West’s public sectors have seen the introduction of new management systems focusing on results (Macrae et al., 2002). This increasing pressure to demonstrate results has not, however, translated into clear improvements in the measurement or analysis of the impact of humanitarian aid. Aid agencies have long found impact difficult to measure, and have tended to focus on what are called ‘process’ or ‘output’ indicators – on what is provided, rather than on its impact in terms of the humanitarian outcome. Assessment of impact is poor, and consistently so:

Reports were so consistent in their criticism of agency monitoring and evaluation practices that a standard sentence could almost be inserted into all reports along the lines of: It was not possible to assess the impact of this intervention because of the lack of adequate indicators, clear objectives, baseline data and monitoring (ALNAP, 2003c:107).

There are, of course, many good reasons why it is difficult to measure the impact of humanitarian interventions. There are problems around causality and attribution, ethical dilemmas around the use of control groups and

often a lack of basic data, such as population figures. Even in relatively stable, developmental environments, measuring impact is difficult. In the humanitarian field, there are many additional challenges: interventions are often short term, capacity and resources are stretched, insecurity may limit access and the space for analysis and research is much more constrained.

The new emphasis on results and measurement in public sector management also has potential risks. In the UK, for example, the adoption of quantitative targets as a measure of performance in the public health service led to a focus on reducing waiting lists for care. This in turn has provided an incentive for doctors to treat simple cases before difficult ones (Chapman, 2003). In the humanitarian sector, a focus on measurement could, for example, reduce operational effectiveness, increase bureaucracy and lead to a neglect of areas that are by their nature difficult to ‘measure’, such as protection. Focusing on what is measurable also risks reducing the humanitarian project to a technical question of delivery, rather than a principled endeavour in which the process, as well as the outcome, is important.

These difficulties and risks do not, however, mean that the impact of humanitarian interventions cannot be measured in some circumstances and, where measurement in a scientific and quantifiable sense is not possible, impact can still be analysed and discussed. Indeed, the scientific tradition and more participatory and analytical approaches should not be seen as polar opposites, but as complementary approaches to impact assessment. The difficulties faced in measuring and attributing impact to particular interventions do, however, suggest a need for greater caution in the bold claims sometimes made by aid agencies about the impact of the assistance they provide. In the humanitarian response in southern Africa in 2002–2003, for example, it was often stated that food aid saved millions of lives:

Generous donor support allowed humanitarian organisations to respond quickly to the crisis, focusing on the immediate goals of saving lives with food assistance and stabilising the nutritional situation. As a result, famine was averted and mass starvation and death were avoided (UNOCHA, 2003:4).

Similar claims have been made in many recent emergencies, such as Afghanistan in 2001–2002 and Ethiopia in 2003, and there are clear institutional reasons for such justifications of humanitarian aid expenditure. Such assertions are often, however, made on the basis of very little evidence, and involve something of a leap of faith; the argument seems to run like this: large numbers of people are

in need, large amounts of assistance are provided and large numbers of people survive. Ergo, aid must have saved lives. In the absence of more rigorous analysis, scepticism on the part of donors and the general public is growing about the claims being made for the impact of humanitarian assistance. By helping to substantiate claims about the success of humanitarian aid in saving lives, greater attention to the art and science of analysing and measuring impact may, therefore, help to foster political and public support and trust in the humanitarian enterprise.

1.1 Scope and methodology

This study maps out the existing concepts, methods and practices of impact assessment used by humanitarian actors. It does not make judgements about humanitarian assistance and the impact it might have, either at the global level or in particular contexts; the focus is on the mechanics of measurement and analysis – and the implications of this for how the humanitarian system, broadly conceived, operates. The study is based on a review of published and grey literature within the humanitarian sector and more broadly, drawing lessons from experience with impact analysis in international development assistance and the private and public sectors. A limited number of interviews with key aid agency staff were also carried out. The study also draws on two pieces of work published as background papers. These are:

- a review of the role of nutrition and food security information in assessing the impact of humanitarian interventions (Shoham, 2004); and
- measuring the impact of health programmes in humanitarian assistance (Roberts, 2004).

Further background papers examining particular technical aspects of impact measurement and analysis will be published as part of a collaboration between HPG and the Feinstein International Famine Center at Tufts University in the US. A resource guide is also available (at www.odi.org.uk/hpg/impact.html).

The question of impact covers a large range of issues which are all potentially relevant for humanitarian action.

The main focus of this study is on the intended consequences of humanitarian aid, and the ways in which these can be measured and analysed. Three areas in particular were not addressed fully, but are clearly important and deserving of further study.

- The broader unintended consequences of humanitarian aid. These are discussed, but are not a primary focus. In the complex and chaotic environments which are faced in many emergencies, interventions are likely to have unexpected consequences. Equally, attributing impact to a particular intervention requires an appreciation of the relative contribution of aid to humanitarian outcomes. In that sense, the wider context within which aid is delivered is crucial in considering the intended effects of interventions.
- The question of the impact of wider initiatives aimed at improving humanitarian practice, such as the Sphere project, and the impact of advocacy campaigns by aid agencies (Van Dyke, 2004; Davies, 2001).
- Questions about the relative cost-effectiveness of different types of programming approaches are increasingly being discussed in the humanitarian sector, although rarely documented (Griekspoor, 1999). Clearly, the question of impact can be linked to that of cost-effectiveness in the sense of wanting to know about ‘bangs for the buck’.

An advisory group was constituted for this project, whose contribution and role have been crucial to the development of the analysis. Members of the group come from the donor community, UN agencies, NGOs and academia.

Chapter 2 examines current definitions of impact and their limitations, the different levels at which impact can be analysed, and the different objectives impact analysis can have. It also sets the increasing interest in impact within the context of recent trends in the humanitarian system. Chapter 3 explores the methodological issues involved in the analysis and measurement of impact, and Chapter 4 looks at the current practice of impact assessment within the humanitarian sector, with particular reference to health and food and nutrition.

Chapter 2

The impact of humanitarian aid: definitions, objectives and context

This chapter explores the concept of impact, looking at current definitions and their limitations, the different levels at which impact can be analysed, the different objectives impact analysis can have and the increased interest in impact analysis among key stakeholders within the humanitarian system. These are all important, and contested, questions. ‘Impact’ is not a value-free, neutral term. Beyond the technical difficulties involved in its definition, our understanding of impact also depends on who is assessing it, what they are assessing, and why.

2.1 Definitions

2.1.1 Defining ‘impact’

The most commonly-used definition of impact within the development sector is provided by the OECD/DAC. This describes impact as:

The positive and negative, primary and secondary, long-term effects produced by a development intervention, directly or indirectly, intended or unintended (OECD/DAC 2002:24).

This refers explicitly to development interventions, and applies only imperfectly to humanitarian assistance. In contrast to the emphasis on the long term in this definition, humanitarian interventions tend to have a short-term focus, and this is not captured here. More crucially, developmental definitions of impact tend to stress the concept of change. In humanitarian environments, a key distinction is often between the end condition (what happened), and *what would have happened* had the intervention not taken place. For humanitarian aid, the aim is often to avert negative change (for example preventing famine) rather than bringing about positive change. This may be harder to measure. (It should, though, be stressed that humanitarian action can also be concerned with trying to bring about positive change; the remedial aspect may be as important as the preventative.)

Moreover, if impact is defined as concerned only with lasting change then the idea of ‘short-term impact’ becomes a contradiction in terms. Oxfam accordingly defines impact as lasting or significant change in people’s lives, in recognition of the fact that, in humanitarian response, saving someone’s life is significant, even if the effect is not lasting, and that individual is again subject to life-threatening risk at some later point (Roche, 1999).

A variety of terms – such as ‘outcome’, ‘results’ or ‘effect’ – are also important here. The distinction between ‘impact’ and these other related terms is not clear, and they are sometimes used interchangeably. There is particular confusion between the terms ‘outcome’ and ‘impact’. In the OECD definition, outcome implies a focus on the short or medium term, rather than the long term. The OECD/DAC attempts to clarify these concepts and reduce terminological confusion; some of the key terms are shown in Box 1.

This report does not propose a separate definition of impact for humanitarian action, but follows the OECD definition given above – with the proviso that a consideration of impact in the humanitarian context may be more concerned with results in the short and medium term.

Box 1: A glossary of terms

Attribution: the ascription of a causal link between observed (or expected) changes and a specific intervention.

Counterfactual: The situation or condition which hypothetically may prevail for individuals, organisations or groups were there no development intervention.

Effect: Intended or unintended change due directly or indirectly to an intervention.

Impacts: Positive and negative, primary and secondary long-term effects, produced by a development intervention, directly or indirectly, intended or unintended.

Outcome: The likely or achieved short-term and medium-term effects of an intervention’s outputs.

Outputs: The products, capital goods or services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes.

Performance: The degree to which a development intervention or a development partner operates according to specific criteria/standards/guidelines, or achieves results in accordance with stated goals or plans.

Results: The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention.

Results chain: The causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives – beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts and feedback.

Source: OECD DAC, 2002.

2.1.2 The origins and uses of impact assessment

The question of impact in the development and humanitarian field has been addressed in three main ways:

- Analysis of likely impact before the start of a project, in order to anticipate the wider consequences of an intervention.
- Ongoing analysis of impact throughout a project or as part of management systems, in an attempt to adapt interventions or monitor performance.
- Analysis of the impact of interventions after the fact, as part of evaluations or research.

Impact assessment in the development sector dates back to the 1950s (Intrac, 2001; Roche, 1999). Generally, the aim is to judge the likely environmental, social and economic consequences of development projects. The most common approaches are environmental impact assessment (EIA), cumulative effect assessment, environmental health impact assessment (EHIA), risk assessment, social impact assessment (SIA), strategic environmental assessment (SEA), cost–benefit analysis (CBA) and social cost–benefit analysis (SCBA). These assessments tend to be done before the start of a project, in the appraisal phase, in order to approve, adjust or reject it (Donnelly, Dalal-Clayton and Hughes, 1998; Roche, 1999). The main concern is with the wider impact of interventions on the external environment, and with identifying and mitigating potential negative impacts. This is also the sense in which impact assessment is conventionally used in the private sector.

In the humanitarian sector, agencies are sometimes required to assess the possible environmental impacts of their work as part of the proposal process. Assessment tools such as the ‘do no harm’ approach can also in part be seen as an attempt to anticipate the possible negative impacts of interventions (Anderson, 1996). However, the assessment of impact has most often been considered as a sub-set of evaluation, and discussion of impact has tended to sit within evaluation departments. Impact is a key evaluation criteria, and most evaluations of humanitarian programmes consider it. In practice, however, assessment of impact within evaluations has often been poor (ALNAP, 2003b).

2.2 The levels and objectives of impact analysis

2.2.1 Conceptual frameworks

Questions around the impact of humanitarian assistance cannot be reduced to a technical discussion about how the impact of particular projects can best be measured. The wider environment in which aid is delivered, and the principles and ethics that underpin humanitarian action, may determine the humanitarian outcomes for populations as much, if not more than, the technical efficacy with which a particular project is delivered. The delivery of humanitarian aid in a manner that could be

construed as partial may jeopardise future humanitarian access, for example. Unprincipled aid could be technically effective in the short term, but could lead to longer-term negative impacts.

The wider elements that need to be taken into account in any assessment of impact can be related to the humanitarian agency itself (its level of resources, its technical competencies, the qualifications of its staff); to the humanitarian aid system (the degree of sectoral and inter-sectoral coordination); or more widely to the general environment. Figure 1 describes the wider dimensions that need to be considered in addressing the question of impact. This conceptual framework shows the complexity of the humanitarian aid system, and the multiplicity of factors that affect impact.

The concept of an ‘impact chain’ or ‘results chain’ (Roche, 1999) is often used to show causality between an action and its ultimate impact. Reduced to its simplest form, the impact chain looks like this:

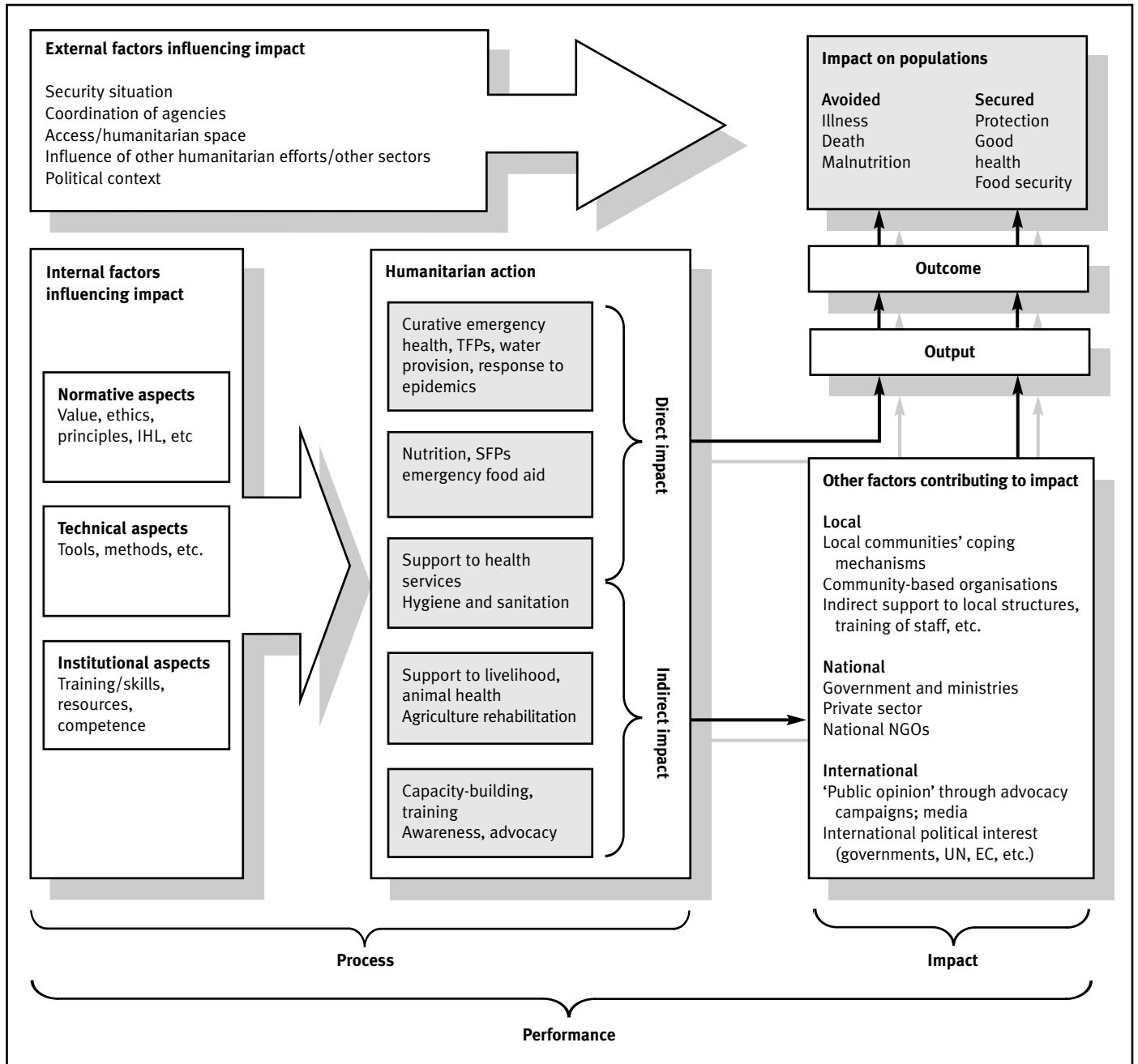
inputs → activities → outputs → outcomes → impact

Care is needed when applying this logic to humanitarian aid. First, there is an important distinction between long and short impact chains: the fewer the links in the chain, the easier it is to assess whether a given input achieves an impact (Roche, 1999). Thus, a therapeutic feeding programme or a response to a cholera outbreak will have a direct impact on mortality. In these cases, the impact chain is short. By contrast, there is a longer chain of causality between a seed distribution and any possible impact on mortality, which may make impact harder to demonstrate.

Second, impact chains can be seen as ‘organisation-out’ processes, which start with the inputs of an intervention, and move out along a causal pathway to impact. However, Roche (1999) argues that ‘much of the literature on impact assessment and evaluation presumes an important, if not predetermined, role for international agencies and actors in delivering humanitarian relief and in developing policy. There seems to be much less literature and evaluation experience about assessing the impact of local response, local preparedness, and local capacity-building’. This implies an additional need for ‘context-in’ approaches, which start with people and changes in their lives, and then work back towards causality. This can serve to better situate changes in a broader context, point to drivers of change other than the particular intervention, and allow for a triangulation based on different perspectives.

Third, there is a tendency to focus on the role of a given intervention at the expense of external factors. In the development context, Maxwell and Conway (2003) note that aid is generally not the central influence on change.

Figure 1: The wider dimensions of impact: a framework



Trends and shocks in the global economy and the capacity and will of the partner government and other non-aid actors are of key importance in explaining change, with aid generally influencing these processes only at the margin. This is often equally true in emergencies, where in most cases humanitarian aid meets only a portion – however vital – of total needs.

Activities such as advocacy, coordination or capacity-building, which have long impact chains, may be particularly difficult to measure. This study does not consider the question of the impact of advocacy in detail, but it is clearly a crucial issue. Advocacy is increasingly being seen as a key part of agency response in crises; a number of

agencies are developing particular methods for assessing the impact of their advocacy activities, and there is a developing literature on this (ActionAid, 2001; Lloyd Laney, 2003; Mayoux, 2003). Learning here may offer useful insights into how to assess complex change processes, which may also be applicable to humanitarian emergencies.

2.2.2 The levels of impact analysis

Impact can be analysed at many different levels, from the individual project, at the level of a sector such as health or nutrition, at a country level or at the level of an organisation and its global impact. Impact analysis may also have many different users, and many different objectives: from demonstrating success and informing

funding choices between agencies or projects to enabling wider judgements about the overall effectiveness of aid in a particular crisis. Table 1 outlines the various ways in which impact assessment may be used. It is intended to be illustrative, rather than exhaustive.

2.2.3 The functions of impact analysis

The literature in this area has tended to identify two main functions for impact analysis: learning and accountability. There is no clear-cut distinction between learning and accountability, and the two are linked in important ways. The links are reciprocal: if an organisation is not learning from its experience, then its accountability is liable to be deficient; being accountable to diverse stakeholders, and managing that balance of accountabilities, can be an important means of learning.

Learning

Learning can apply at different levels, from the project to the sector to the system. At the project level, ongoing monitoring of impact informs changes to projects. However, the form of impact analysis best suited to the learning function may be located outside the routine monitoring of projects. It may require greater expertise and time, and it may follow more of an academic research

model. Impact data have to be collected and analysed across the same type of interventions over time.

Learning may also imply the use of impact assessment in the course of aid activities in order to reassess, reorient, or possibly close down activities: this could be called a corrective function. This is usually done via mid-term impact assessment, or through monitoring activities. Real-time evaluations are also increasingly being used for this purpose (UNHCR, 2001). This form of assessment may also support and document the impact of a reduction in the level of aid. For example, in order to make informed decisions about reducing levels of food aid and ultimately closing down a project, a manager needs to know current levels of need, as well as the impact of food aid on a population (see Box 2).

The learning function of impact analysis relates to a wider question around the evidence-base of humanitarian aid: to what extent are operational choices based on a repetition of previous approaches, or on strong supportive evidence that this is the best approach? New approaches to therapeutic feeding provide an example of the development of an evidence-base leading to a gradual shift in programme approaches (see Box 3). Analysis of impact

Table 1: Levels and objectives of impact analysis

Levels of analysis	Who wants to know and why?
The impact of particular humanitarian aid projects	<i>Aid agencies</i> , in order to improve their work, demonstrate impact and make choices between projects <i>Donors</i> , to choose what to fund and to develop policy <i>Agencies and donors</i> , to assess the impact of new approaches and innovations in programming <i>National governments</i> , to guide disaster preparedness, planning and response
The impact of particular humanitarian organisations	<i>Aid agencies</i> , to demonstrate success and raise money from the public and from donors <i>Donors</i> , to choose between competing agencies or to make choices about whether to use NGOs or private contractors <i>National governments</i> , to choose who to register and work with as partners
The impact of humanitarian aid at a sectoral level	<i>Aid agencies and donors</i> , to build up the evidence-base for what works; to develop sectoral policies and best practice protocols and guidelines <i>National governments</i> , to put protocols in place
The impact of humanitarian aid at a country level or for a particular crisis	<i>Donors</i> , to know how many lives were saved <i>Donor publics</i> , to know whether the money they donated made a difference <i>National governments</i> , to assess whether appealing for international aid was the right thing to do <i>Agencies</i> , to advocate for increases in levels of aid
Impact of international engagement in a crisis, including, but not limited to, humanitarian aid	<i>Governments</i> , to review their overall engagement with countries in crisis (diplomatic, political, military and aid) <i>Aid agencies and humanitarian donors</i> , to be clear about the role and scope of humanitarian aid <i>Agencies</i> , to advocate for greater political engagement in 'forgotten crises' <i>Governments</i> , to promote the coherence of political and humanitarian agendas <i>Aid agencies</i> , to maintain the neutrality and independence of assistance

Box 2: The corrective function of impact assessment: an example from Angola

As conditions improved in Malange, Angola, in 1995, the international humanitarian community began to explore ways to decrease the amount of food aid being provided. It was quickly noted that, to determine the optimal level of food aid requirements, reductions in the general ration were best carried out in conjunction with surveillance activities. This showed that a gradual reduction in the general ration did not necessarily have a negative impact on nutritional status. However, the near-total withdrawal of the ration in December 1995 resulted in an increase in the level of acute malnutrition among children under five years of age. This suggested that certain population groups within Malange were still at least partially dependent on external assistance. More in-depth qualitative studies revealed that a proportion of the population, especially those who were displaced from rural areas and had no access to land within the peri-urban secure boundaries of the city, were particularly vulnerable if the ration was withdrawn. The information allowed for more effective targeting of limited resources (Borrel and Salama, 1999).

Box 3: Community-based approaches to managing severe malnutrition

Therapeutic feeding centres (TFCs) are the accepted form of intervention for the treatment of severely malnourished children in emergency nutrition programmes (Collins, forthcoming 2004). Recent operational research has showed that, in some circumstances, a different approach to severe malnutrition could potentially have greater impact, and minimise some of the risks and problems associated with traditional TFC programmes. A concept sometimes labelled 'Community Therapeutic Care (CTC)' has been developed and field tested. In this approach, services are brought closer to the malnourished children's principal carers. Studies in Ethiopia, Malawi and other countries have suggested that outpatient care can meet or exceed internationally accepted minimum standards for recovery (Collins, 2001; Collins and Sadler, 2002; Emergency Nutrition Network, 2003b).

The main principles of CTC are:

- High coverage and good access
- Timeliness – because mortality often occurs before TFC interventions are up and running.
- Sectoral integration – nutrition interventions must be integrated with other programmes, including food security, health, water and sanitation.
- Capacity-building – an active commitment to building on existing structures through consultation, training and ongoing support.

Agencies are also testing similar models, such as the Home Base Treatment approach used by MSF and ACF (Emergency Nutrition Network, 2003a).

can be crucial in generating evidence regarding different forms of humanitarian programming, and in enabling informed choices between different programme options.

Accountability

A second function of impact assessment is related to the question of accountability. This is primarily an issue of upwards accountability, between donors and aid agencies; downwards accountability (of aid agencies to the beneficiaries of aid) is often talked about, but seldom practised. Donor governments want to know about impact, both because they have a responsibility to ensure that public funds are well spent, and because they need to choose where to allocate scarce public resources. Debates about accountability flourished in the 1990s, as humanitarian agencies faced growing levels of scrutiny and criticism, and the idea that humanitarian aid was unquestionably a good thing began to be questioned. There was a realisation that aid alone, even when well delivered, might have only negligible impacts in situations where other political, economic or social factors were far more important in determining humanitarian outcomes. Political economy approaches highlighted a series of difficult dilemmas around the delivery of aid in conflict, such as the risks of aid being diverted to warring parties (Cliffe and Luckham, 2000; Collinson, 2003; Le Billon, 2000). This focused attention on the possible negative and unintended consequences of humanitarian aid.

Linked to this were a number of initiatives aimed at increasing accountability, such as the Code of Conduct, Sphere, ALNAP, the Ombudsman project, the Humanitarian Accountability Project (now the Humanitarian Accountability Partnership International), and People in Aid. These all have in common a concern for the quality, performance and accountability of humanitarian aid. Two elements are particularly relevant here. First, greater efforts and attention are being put into humanitarian evaluation (ALNAP, 2003b). Second, there has been a focus on the development of standards and indicators through the Sphere process (Sphere, 2004). Potentially, these provide benchmarks against which humanitarian impact can be measured, though critics argue that Sphere is overly focused on the technical aspects of aid delivery. Several international NGOs have introduced impact assessment systems that aim to improve accountability at the organisational level. ActionAid's Accountability, Learning and Planning System (ALPS) and Save the Children (UK)'s Global Impact Monitoring (GIM; see Box 4) stress both upwards and downwards accountability (British Agencies Aid Group, 2002; SC UK, 2003a; Starling, 2003).

2.2.4 Assessing negative impacts

In the last decade, much has been written about the potentially negative impacts of humanitarian intervention on people's livelihoods. Research during the 1990s

Box 4: Save the Children (UK)'s Global Impact Monitoring framework

Save UK's Global Impact Monitoring (GIM) framework was developed in 2001, and pilot tested in a number of countries and regions in 2002 and 2003. Its core purpose is to shift the emphasis from a list of activities towards a greater focus on the impact of programmes: 'Our aim is to banish ever hearing 'we have this really excellent programme' when there is not evidence to back up this statement!' (Save the Children UK, 2004). Evidence should be collected throughout the lifecycle of an intervention (project or programme) through ongoing monitoring and periodic reviews and evaluations.

The GIM approach is largely participative and geared towards providing qualitative information. Country teams are encouraged to include qualitative elements in the analysis whenever possible and to collect quantitative as well as qualitative evidence through relevant indicators. Impact Review Meetings give an opportunity for stakeholders (staff, partner organisations, children, the local authorities, donors) to share their experience and analysis about the impact – positive or negative, intended or unintended – of Save's activities.

The GIM is a flexible approach, and does not provide a blueprint or standard format for assessing impact. It is rather a process that encourages analysis, discussion and consensus-building. The central feature in the analysis is the identification of the 'common dimensions of change'. These are:

- change in the lives of children and young people;
- change in policies and practices affecting children's and young people's lives;
- change in equity and non-discrimination of children and young people; and
- change in civil societies' and communities' capacities to support children's rights.

The five dimensions of change provide an analytical framework that can be applied in different contexts. However, the flexibility of the GIM process makes it difficult to compare impact across countries and to aggregate and consolidate information at higher levels. The GIM has been used mainly in relatively long-term programmes. It is more rarely used in humanitarian emergencies, mainly because less data is usually available, and it may be more difficult to bring stakeholders in for an Impact Review Meeting.

highlighted the fact that the provision of relief resources may influence the dynamics of conflict, and evidence emerged indicating the manipulation of relief supplies by warring parties (Macrae and Leader, 2000; De Waal, 1997; Duffield, 1994; Keen, 1994). Assessment tools were developed to analyse the relative benefits and potential harm of providing assistance in conflicts where diversion is a risk, and there were calls for aid agencies to improve

their capacity to analyse the political economy of conflicts (Anderson, 1996 and 1999; Collinson, 2003).

The availability of assistance in one area may encourage migration from elsewhere, or the arrival of food aid may force down local prices. There is also a belief that long-term food assistance creates dependency and undermines livelihoods. Some of these impacts, such as migration to feeding points, can be easily assessed; others, such as fostering dependency or driving down producer prices, are far harder to measure, let alone attribute. During an intervention, market analysis can help measure the consequences of food aid interventions, but this is not an exact science, especially given the type of data available in typical emergency environments.

2.3 Perspectives from new public management

Concern for the effectiveness, performance and impact of aid programmes has also been influenced by changes in the wider context of public management. Most Western governments have reformed their systems of public management with a view to improving service provision. A central element of this reform is the shift from an input–output management model towards a greater emphasis on results. Service providers not only report progress in implementing activities, but must also demonstrate that they generate some achievements. A culture of setting targets, measuring performance and assessing achievements in quantifiable terms has emerged (Wallace and Chapman, 2003). Performance management or results-based management has become the instrument of choice for monitoring performance and impact.

The OECD/DAC defines results-based management as 'a management strategy focusing on performance and achievements of outputs, outcomes and impacts' (OECD/DAC 2002). The key features of results-based management are:

- A focus on the recipient of the service.
- A concern for quality and performance.
- The consistent use of objectives and indicators.
- The participation of stakeholders.
- The reform of budget processes and financial management systems.

Results-based management covers diverse areas, such as waiting lists in hospitals, the performance of primary schools, the reliability of train services or suicide rates in prisons (Hailey and Sorgenfrei, 2003).

Donors such as USAID, DFID, AusAID, ECHO (see Box 5), CIDA and Danida have adopted results-based management approaches, and this has created a demand for analysis of impact; DFID's partnership agreements

with NGOs require agencies to report on impact at the global level (British Agencies Aid Group, 2002). Within the UN, WFP, UNHCR and UNICEF have established results-based management systems. In WFP, the division for results-based management reports directly to the Executive Director (WFP, 2003b). ICRC introduced a performance management system called Planning for Results in 1998 (see Box 6). Planning for Results aims to cover planning, budget construction and appeals, implementation, financial and human resource management and evaluation. Many of these initiatives, however, are at an early stage.

Results-based management approaches have been criticised. First, it is argued that they are simplistic, assuming that a specific intervention has a linear, causal effect, when the reality is in fact more complex. Second, it is argued that the imposition of quantitative targets increases administrative overheads, makes institutions more fragile, demotivates staff and disillusiones clients (Chapman, 2003). It can also lead to perverse incentives, as the need to meet targets encourages organisations to tackle the easier tasks at the expense of the more difficult. In the UK, it has been argued that the new 'accountability culture' has devalued professional responsibility and distorted professional practice (O'Neill, 2002a).

In the aid sector, as Maxwell (2003) argues, there can be little disagreement with the proposition that outcomes matter, or with the desire to move from measuring inputs and outputs to focusing more on results and the real needs of clients. There is nonetheless growing concern among aid

Box 5: ECHO's Activity Based Management

ECHO has made significant changes to its management and reporting systems in order to fulfil the Commission's requirements for Activity Based Management, introduced in 2001. This entails producing annual management plans and reports, and agreed output and, where possible, impact indicators (Commission of the European Communities, 2001).

These changes occurred at two levels. First, contractual arrangements with partners were reviewed in 2003, and now incorporate a 'results-based' approach. These agreements no longer focus on the control of inputs, but on the definition of clear objectives and indicators (ECHO, 2004). Partners are requested to provide information on results with systematic reporting on selected indicators. Second, ECHO has established methodologies and indicators that allow it to measure progress in its strategic priorities (these are intervention in areas of greatest need, a focus on forgotten crises and addressing four priorities: disaster preparedness, linking relief, rehabilitation and development (LRRD), child-related activities and water). Considerable technical changes are required to fulfil these new reporting requirements.

Box 6: ICRC's Planning for Results

'Planning for Results' (PFR) was progressively introduced within the ICRC in 1998 following a change management initiative called the Avenir Process. PFR is an integrated system used at every level of the institution, and at all phases of the project cycle; it covers not only planning but also other management functions like accounting, logistics, statistics, human resources, fundraising and donor reporting. The objectives of this new management system are to enhance the performance of ICRC's operations, to promote a results-oriented management culture and to evaluate results and impacts. The collection of indicators is a central dimension of PFR. Indicators are defined in relation to outputs, outcomes and, whenever possible, the impact of programmes.

agencies around the possible negative effects of the current vogue for results-based management. There is anecdotal evidence that results-based management puts higher reporting demands on agencies, increasing bureaucracy at the expense of operationality (Wallace and Chapman, 2003). In turn, these requirements may increase donors' influence over humanitarian agencies, and may encourage NGOs to address the perceived demands of donors at the expense of other considerations. The focus on results also risks inflating claims about what NGOs can achieve. This leads to a vicious cycle: 'the growing reporting requirements increase this pressure to show in a positive light everything that has been done. However, at the same time this increases the cynicism about development and their [NGOs'] effectiveness – and this leads to increasing pressure to tell good stories to counter that cynicism' (Wallace and Chapman, 2003). There is a tendency in the humanitarian sector to favour conservative approaches over innovative ones. Results-based management may reinforce this: humanitarian agencies may rule out new approaches where there is no guarantee that they will yield results.

Introducing criteria too focused on the results of humanitarian assistance may also lead agencies to neglect important but difficult-to-measure dimensions of humanitarian aid, such as protection and advocacy. ICRC, for example, explicitly regards its presence as having a positive impact in particular environments, and humanitarian agencies involved in advocacy consider that this work too can have an indirect beneficial impact. Quite how results-based management would capture these aspects of agencies' work is not self-evident. Data related to local culture and context may be squeezed out.

Finally, there is a question over the extent to which results-based management approaches actually lead to more knowledge about impact. A focus on results can be just as prone as previous systems to look at process and not impact. Problems with attribution mean that it can be very

difficult to link accountability to impact, as opposed to outputs. In practice, this often means that results-based management has the effect of holding people accountable for outputs and proxy indicators that may not be proven means of achieving the impact sought.

2.3.1 Perspectives from the private sector

Whereas humanitarian agencies are increasingly borrowing approaches from the new public management, initiatives drawing on experiences from the private sector have been relatively limited. In the business world, impact is ultimately measured by profitability, but management approaches may also have lessons for the analysis of impact in the humanitarian sector.

Performance management systems in the private sector include approaches such as total quality management, benchmarking, balanced scorecard and excellence model systems. These management approaches have been only sparingly used in the humanitarian sphere. Hailey and Sorgenfrei (2003) note that the logical framework, originally a planning tool for the military and introduced into international aid by USAID, has since become a common tool for many humanitarian agencies. 'Logframes' are not used in the private sector, and recent performance management systems may provide a less rigid approach.

Although not widely used, some management systems developed in the private sector have been applied in the humanitarian sphere. The NGO Medair is using a Quality Management System, and has introduced ISO 9001 quality standards. Medair sees some clear benefits in using this system, such as promoting downwards accountability, being sensitive to contexts and being flexible, while enhancing the quality of programmes. The agency also recognises that ISO 9001 can be misused, and can generate bureaucracy (Medair, 2002). Private sector experience is also being drawn upon in the areas of technology and logistics. The Fritz Institute, drawing on expertise from the private sector, has developed a logistics software for the IFRC that can accelerate relief delivery (IFRC, 2003; Refugee Studies Centre, Fritz Institute and Norwegian Refugee Council, 2003). Evaluation tools that draw on private sector techniques, such as customer focus groups and customer polling, are another area that may be worth exploring.

2.4 Chapter summary

This chapter has reviewed how impact is defined, the different levels at which it can be measured and the different objectives it can have. It has also explored the influence of wider changes in the culture and practice of public sector management in the West. Impact assessment can refer to attempts to anticipate the possible future impacts of projects, and to evaluate the impact of projects after the fact. Impact can be analysed at the level of individual projects, and at much broader organisational or country-wide levels. Attempts to measure impact can restrict their focus to the intended effects of interventions, or they can encompass broader indirect and unintended consequences. The humanitarian system's increasing interest in impact needs to be understood in the context of broader debates about accountability in humanitarian aid, and as part of public management reforms within Western governments. Experience from the introduction of results-based or performance management systems within Western governments suggests a need for caution in adopting these approaches uncritically. In particular, an increased focus on results may lead to the neglect of harder-to-measure aspects of humanitarian action, such as protection.

The analysis of impact should not be seen just as a narrow technical question about the effectiveness of individual projects. Project-based approaches that focus on determining the impact of a particular intervention should be complemented by approaches that start with changes in people's lives, and situate change within the broader external environment. Nor should discussion about impact be confined to a sub-set of evaluation techniques. In the humanitarian sphere, a concern with significant change in the short term implies a need for impact to be considered in ongoing monitoring processes, and through techniques such as real-time evaluation. There is a need for greater investment in approaches that aim to build up the evidence-base of what works through more detailed research; and for system-wide evaluations that ask difficult and important questions about responsibility for humanitarian outcomes and the broader political dimensions within which the humanitarian system operates.

Chapter 3

Measuring and analysing impact: methods, indicators and constraints

Analysing the impact of interventions, whether developmental or humanitarian, is not straightforward. There are a number of methodological constraints that make impact measurement difficult. The nature of the operating environment, the need to act quickly in situations of immediate crisis, an organisational culture that values action over analysis and the fact that there is little consensus around the core objectives of humanitarian aid all make impact analysis hard. This chapter sets out the technical requirements of measuring impact and the indicators and methods generally used. It then reviews the key factors constraining the measurement of impact. Given these constraints, measuring impact in a strict scientific sense may rarely be possible. However, this does not mean that a more discursive analysis will not be achievable.

3.1 Methods for measuring and analysing impact

Three main approaches to impact assessment can be identified within the literature. Hallam (1998) summarises them as:

- the scientific approach, which is favoured by those wishing to generate quantitative measures of impact;
- the deductive/inductive approach, which is more anthropological and socio-economic; this approach relies on interviews with key informants, and draws on other similar or comparable cases; and
- participatory approaches, which depend on obtaining the views of those benefiting from a programme.

Hulme (1997) similarly distinguishes between scientific approaches, the humanities tradition and participatory approaches. The humanities tradition does not try to prove impact in a statistical sense; instead, it seeks to provide:

an interpretation of the processes involved in intervention and of the impacts that have a high level of plausibility. It recognizes that there are usually different and often conflicting accounts of what has happened and has been achieved by a programme (Hulme, 1997).

The validity of this approach has to be judged based on the logical consistency of the arguments, the quality of the evidence and methodology, the degree of triangulation to cross-check evidence and the reputation of the researcher.

Within these broad approaches, there are a huge array of tools for analysing impact, often divided between qualitative

Box 7: Steps and conditions for measuring impact

While impact analysis should not be reduced to the purely technical, there is clearly a sense in which measuring impact is a technical question. In order for impact to be measured or demonstrated a number of key steps need to be taken, and technical conditions satisfied. Maxwell and Conway (2003), for example, describe three such steps:

- identify change;
- establish a causal connection between the change and the input; and
- measure the magnitude of the change.

Hill develops criteria for attributing the causation of a disease to exposure to a chemical or biological agent (Hill, 1965).

- The greater the strength of the association, the more likely it is to be causative.
- There is a dose-response relationship between the exposure and the health outcome.
- Exposure consistently induces the health consequence in different settings at different times.
- The exposure occurs before the health outcome.
- There is a biologically plausible explanation for the exposure resulting in the health outcome.
- There are no more plausible explanations for the health outcome.
- Experimental results add particular weight to the evidence.

Hill's criteria have a clear logic that can be more widely applied to assessing impact. Not all of these criteria need to apply in order to demonstrate causation. For example, not everyone exposed to a pathogen becomes ill and most people who smoke never develop lung cancer. Nonetheless, having several of these criteria met greatly strengthens the argument that a programme has had some impact. The idea that an intervention should be replicable and show impact in different settings and at different times is important in developing an evidence base about what works. There should not be a more plausible alternative explanation for the impact being claimed for a particular intervention. Programmes need to be evaluated with particular regard to the likelihood that the level of inputs provided could plausibly result in the outcome reported. For example, the number of clinic visits or the amount of food provided per child need to be sufficient to induce the effects observed.

and quantitative. Roche distinguishes several groups of tools and methods: surveys; interviews, workshops and discussions; direct observation; participatory research; and case studies (Roche, 1999). Most authors argue that a mix of methods and/or approaches is desirable to meet the broad objectives of impact assessment. As Chelimsky (1995) puts it, by using methods in complementary and critical ways (methodological triangulation), 'the strength of one can compensate for the limitations of others'. The distinction between scientific, analytical and participatory approaches should not be seen as hard and fast, and they should be understood not as polar opposites, but as complementary facets of impact assessment.

This study cannot provide a comprehensive guide to the large number of quantitative and qualitative methods available for analysing impact. In the health sector alone, a host of guidelines are available for documenting health problems and health conditions in emergencies. WHO distributes a CD-ROM containing almost 200 guidelines and manuals for assisting workers in complex emergencies. There is a similar range of guidelines and manuals in other sectors, such as food security and nutrition. Instead, this study highlights key issues involved with some of the most commonly used methods, with a particular focus on participatory methods, surveillance systems, surveys and mortality data. More detailed discussion of some of the technical issues relating to methods of impact measurement in the nutrition and health sectors is in Chapter 4.

3.1.1 Participatory approaches

A vast development literature is dedicated to participation in general, and participatory monitoring and evaluation in particular (Estrella and Gaventa, 1998; Cornwall and Pratt, 2003; Chambers, 1997). A wide range of participatory tools and techniques is also available, including interviews, focus-group discussion, Venn diagrams, time-lines and historical profiles, ranking and impact flow charts. The most commonly-used participatory approach is Participatory Learning and Action (formerly Participatory Rural Appraisal). Yet participatory approaches, while widely recognised as a key component of understanding impact, have rarely been utilised in the humanitarian sector. Hallam (1998) notes that 'Humanitarian agencies are often poor at consulting or involving members of the affected population and beneficiaries of their assistance. Consequently there can often be considerable discrepancy between the agency's perception of its performance and the perceptions of the affected population and beneficiaries'. The ALNAP Global Study on participation shows that there are in practice very few examples of participatory approaches in project evaluations and impact analysis of humanitarian action (ALNAP, 2003a and b). This reflects a wider weakness in consultation with, and the participation of, affected populations in humanitarian response.

There have been attempts to adapt criteria used to assess the quality of conventional research, such as credibility, transferability, dependability and confirmability, and new techniques for participatory evaluation, such as 'most significant change approaches', may also have potential in emergency situations (Roche, 1999; Pretty 1993; Guba and Lincoln, 1989). This is a story-based technique where groups of stakeholders identify and discuss programme outcomes (Dart and Davies, forthcoming). Kaiser (2002) examines the potential for participatory approaches in the evaluation of humanitarian programmes. Whilst practice remains limited, UNHCR, for example, has used a lessons-learned workshop in Liberia to assess the impact of its voluntary repatriation and reintegration programme. The approach, which involved the participation of 200 stakeholders, was intended to move away from reviewing quantifiable outputs (food delivered, number of refugees transported to their homes) by asking stakeholders what difference UNHCR's programmes had made (UNHCR, 2003). The Disasters Emergency Committee (DEC) evaluation in Gujarat included a survey of beneficiary views (see Box 8) (Humanitarian Initiatives/DMI/Mango, 2001).

Box 8: A public opinion survey as part of the DEC Gujarat evaluation

The Disasters Emergency Committee (DEC), an umbrella organisation of 12 UK agencies, launches and coordinates national appeals in response to major disasters. Independent and public evaluations of the effectiveness of the response are conducted after each appeal. Beginning with the Central America Hurricane Mitch appeal in November 1998, evaluations include the collection of beneficiaries' perceptions, either through *ad hoc* or through systematic interviews, or with more elaborate beneficiary surveys.

One of the most detailed beneficiary surveys was conducted by the Disasters Mitigation Institute (DMI) in India as part of the independent evaluation of the DEC response to the earthquake in Gujarat in 2001 (Humanitarian Initiatives/DMI/Mango, 2001). The survey covered 50 villages, and over 2,300 people were interviewed. The survey gave useful insights on what the community felt about the targeting, timing, quality and quantity of the various interventions. It showed that the geographical distribution of relief was uneven, and that the timing was generally good for most relief items, with the exception of livelihood interventions and shelter. It highlighted some inconsistencies in the quantity, quality and appropriateness of relief, noting for instance that communities felt that the clothes distributed were not appropriate. The survey also provided information about the levels of community participation in the various interventions: people generally felt that they had not been involved enough in the assessment process or in the selection of beneficiaries.

3.2 Indicators

The identification and use of relevant indicators is a crucial part of determining the impact of an intervention. The OECD/DAC defines indicators as the ‘quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievements, to reflect the changes connected to an intervention, or to help assess the performance of a development actor’ (OECD/DAC 2002). Although the terminology varies, the literature generally distinguishes between two types of indicator: those that relate to the implementation of programmes (input, process and output indicators); and those concerned with the effects of programmes (outcome and impact indicators); these are described in Table 2.

Table 2: Types of indicator: example of measles immunisation programmes

Implementation of the programme			Effect of the programme	
Input indicator	Process indicator	Output indicator	Outcome indicator	Impact indicator
No. of vaccines administered	No. of people trained	Percentage vaccinated	Measles cases decrease	Mortality decreases

Humanitarian agencies tend to use a mix of indicators, depending on their own monitoring and reporting systems and the particular function of the indicators collected. Documenting the impact of a programme is only one of many reasons why indicators are collected; others include monitoring the implementation of activities, determining when aspects of a programme are off-track, or to inform decision-making. Both types of indicator – process and impact – are important. A manager needs to know that their activities are being carried out according to plan, and whether they are having any impact. Similarly, a donor wants to know whether the project being funded has been implemented, and whether the decision to fund this particular approach was the correct one. That said, the background papers for this study suggest that agencies tend to collect process rather than impact indicators (Roberts, 2004; Shoham, 2004). Roberts (2004) found that many organisations use process indicators (such as drug doses supplied, clinics supported or staff trained) or outcome indicators (such as clinic attendance) to justify general health programmes designed to reduce mortality. Similarly, the Sphere project, probably the most comprehensive attempt to define standards and indicators for most areas of humanitarian aid (Sphere, 2004), focuses largely on process, and its indicators are not designed to show impact.

There are several reasons why process/output indicators, rather than impact indicators, tend to be collected. Despite

the introduction of results-based management systems, donors tend to favour process/output indicators, and funding proposals and reporting formats are not necessarily geared towards a concern for impact. The collection of impact indicators is sometimes seen as too difficult; mortality indicators, for instance, are known to be hard to gather. It is easier for humanitarian agencies to monitor their own activities than to monitor or assess the effect these activities have on the populations they are helping. This is not a justification in itself, and it is arguable that greater incentives and/or training may encourage the collection of this information in a more systematic way.

Initiatives such as the inter-agency Standardised Monitoring and Assessment of Relief and Transition (SMART) may lead to more routine collection of impact indicators. SMART uses two measures – Crude Mortality Rate (CMR) and the nutritional status of children under five – as basic indicators for assessing the severity of population stress, and for monitoring the overall effort of the humanitarian community. SMART has its origins in North American legislation requiring public sector bodies to demonstrate performance results. The initiative has focused on developing standard methodologies for food security, vulnerability and livelihoods analysis (led by UNICEF); training implementing partners (led by Tulane University); and creating a related database (led by the Centre for Research of the Epidemiology of Disasters (CRED) at the Université Catholique de Louvain, Belgium). The intention is that the two indicators will be used both to measure the severity of a situation at a given time, and to measure changes attributable to relief interventions (i.e., as a gauge of impact). Of course, malnutrition and mortality figures will not necessarily enable impact to be attributed to particular projects or agencies, and may only be able to demonstrate the extent to which the relief system as a whole is meeting the needs of a population (SMART, 2002a). Discussions are being held with implementing partners about other indicators, including morbidity, and the US government’s Bureau of Population, Refugees and Migration (BPRM) is exploring the development of protection indicators.

Given the apparent emphasis on process/output indicators, can such indicators provide insights into the impact of an intervention? Process/output indicators can be used as a proxy for impact where there is strong evidence of causality between the action being monitored and the related impact. For example, immunising children against measles is known to have a direct effect on reducing mortality. With interventions that have a vast literature documenting the attributable benefits (for example measles vaccination or Vitamin A supplements), the need to show ‘proof’ that the intervention produced a health benefit may be small. For medical and public practices that have been well studied in non-emergency

situations, there is little reason to question the logic of the intervention, and evaluations tend to focus on the relative success of implementation. In a case such as this, there is considerable evidence and experience from similar past situations (Garner, 1997; Spiegel et al., 2001a; Toole, 1999), and the use of impact indicators may not be necessary to confidently show benefits from a particular intervention. Garner (1997:722) argues:

Reliable information on effectiveness comes from research, particularly randomised controlled trials, and summaries of evidence help inform policy. In areas of public health where trials are feasible, impact on morbidity and mortality are assessed, along with measures of success of implementation, such as coverage and compliance levels. Then, if the intervention is shown to be beneficial, the planner has an indicator (coverage for example) which can be assessed routinely; and has evidence from research of the coverage level at which an effect has been demonstrated on health status.

There are, however, few proxy indicators in humanitarian assistance that provide sufficient evidence of a link with impact. The link is often assumed, but rarely demonstrated. Spiegel et al. argue that more research is needed on the type and thresholds of indicators that lead to improved health outcomes in populations affected by complex emergencies. For many emergency interventions, such as HIV prevention via education, there may be little evidence that such programmes produce any health benefits, making the importance of documenting any benefits great (Spiegel et al., 2001a). The same argument could be expanded to other humanitarian sectors, in terms of the need to build up through research an evidence base of interventions that can reliably demonstrate impact through the collection of proxy indicators.

3.2.1 Mortality

The core of the humanitarian agenda is about saving lives, and so mortality rates seem to be a logical starting point for the analysis of impact (as reflected in their adoption in SMART). However, mortality is an extremely late indicator; clearly, the objective of humanitarian aid is often to prevent mortality from rising in the first place. Moreover, there is no standard, accepted method for measuring the mortality rate. Surveillance systems for monitoring mortality, such as monitoring burial places, routine reports from street leaders in refugee camps or reports of deaths in hospital, usually require a reasonably stable situation and reliable population estimates. They also take considerable time to establish, and need to run for some time before data can be meaningfully analysed. These factors make them unsuitable for estimating mortality in emergency assessments. Some of the methodological difficulties with collecting mortality data are described below, and more detail is provided in the background papers to this report (Shoham, 2004; Roberts, 2004).

Mortality data can be used retrospectively to help demonstrate impact (as in Box 9). Aid agencies regularly use one of three methods for conducting retrospective mortality surveys: the past household census method; the current household census method; and the children ever born method (Woodruff, 2002). It is possible to estimate cumulative incidence retrospectively using a cross-sectional survey. This is currently the recommended method for estimating mortality in emergencies. There are, however, a number of problems with this approach. The first is manipulation: the number of deaths may be exaggerated in order to secure more aid; or they may be under-estimated for fear that an accurate report of the number of deaths will lead to lower rations. Second, in some cultures death is a taboo subject, and this can lead people to under-report mortality. The third problem is that methodological mistakes can make figures unreliable. The most common blunder is to 'nest' the mortality survey within a nutrition survey, thereby excluding households in which all children under five years of age had died. In general, the methods used lack standardised procedures for defining households, enumerating household members and selecting the principal informant. It may be difficult to ascertain whether identified household members were living at home during the survey period, there may be a failure to define live births and there may be no standardised question set. Fourth, there may be difficulties in estimating the size of the denominator. Household census methods require the tracking of a potentially large number of individuals over time, some of whom may move in and out of the household during the recall period. Fifth, guidance on sample size calculations and data analysis procedures may be missing, as current editions of handbooks on emergency assessment do not provide details on how sample sizes should be calculated. Finally, background or underlying mortality rates may vary widely between populations.

Box 9: Using mortality data in Ethiopia

A study in Gode, Ethiopia, used a two-stage cluster survey involving 595 households to examine mortality retrospectively over an eight-month period. UN agencies were claiming that widespread famine had been averted in 2000 due to the humanitarian response. However, the survey established that, from December 1999 to July 2000, the CMR in Gode was approximately six times higher than the pre-famine baseline, and three times higher than the accepted point of definition for a complex emergency. The study concluded that most deaths were associated with wasting and major communicable diseases, and occurred before the humanitarian intervention began. The intervention may in fact have increased disease transmission and mortality by attracting non-immune malnourished people to central feeding locations (Salama et al., 2001).

3.2.2 Surveillance systems

Surveillance is the systematic collection of information over time. Surveillance systems are often part of general monitoring systems, and have been used for analysing impact in both health and nutrition programmes. Aid agencies sometimes evaluate health programmes by establishing a surveillance system at the beginning of a project, and comparing data on morbidity levels at the beginning and end of the project. This is valid if a) all of the events of interest are captured by the surveillance network; or b) the data from within the system is representative of the health conditions of the entire population, and remains consistently so over the course of the project. Often, however, these conditions are not met in clinic-based surveillance systems, which have an in-built bias because they only record people who come to the clinic. If the majority of a population does not have access to formal health care, then a clinic or hospital-based surveillance system will be able to tell very little about the health conditions of the broader population.

Not all surveillance systems are linked to utilisation of formal health services. Sentinel site surveillance systems for nutrition monitoring, for instance, monitor selected communities in order to detect changes in context, programme and outcome variables. Surveillance systems can be cheaper than surveys and may reveal more detailed information on the causes of malnutrition. Their main weaknesses are that, depending on selection, population groups that may be of interest are excluded, and the data collected may be biased, and cannot be extrapolated since it may not be representative of the wider population. Surveillance systems may also be appropriate for mortality monitoring, depending on the extent and phase of the crisis. After the immediate response, it is often possible to establish mortality surveillance systems (Spiegel et al., 2001b).

3.2.3 Surveys

Surveys may take different forms. WHO and others have produced manuals specifically to guide health workers in conducting specific kinds of survey, with nutritional anthropometry and EPI (Expanded Program on Childhood Immunizations) coverage methodologies among the most succinctly described. Aid workers often do not have sufficient skills to take a valid sample and to analyse the results of a survey. This is why many initiatives to improve the quality of relief programmes have emphasised the importance of training relief workers in survey methodologies. There is also an issue around skills. To analyse impact properly, aid workers need knowledge in areas such as survey approaches, sampling techniques and statistical analysis. A review by the US Centers for Disease Control and Prevention (CDC) of data used in the Horn of Africa over the last decade showed that the majority of nutrition surveys had methodological problems, and established best practice

had not been followed. There are numerous other examples of poorly conducted nutrition surveys (Garfield, 2001; Shoham et al., 2001; SMART, 2002b).

Uncertainty about population figures (on which see below) creates particular difficulties in constructing sampling frames for use in surveys. Census data may be many years old, while the crisis may have had a dramatic impact on demographics and population numbers due to out-migration and high mortality. Although cluster surveys are a compromise measure, in many situations (especially in conflict environments or where terrain poses problems) it may be difficult to gain adequate access. Nomadic groups may also be difficult to sample (SMART, 2002b; SMART, 2002a; World Vision, 1999).

3.3 Constraints to measuring impact

3.3.1 The emergency context and the nature of the system

The context in which humanitarian agencies operate clearly creates difficulties for the analysis of impact. These include the difficulties of the operating environment, the need to act quickly in situations of immediate crisis, an organisational culture that often values action over analysis, and a lack of consensus about the objectives of humanitarian aid. ALNAP provides a useful summary of some of the key constraints, reproduced in Box 10.

Box 10: Constraints to impact assessment in the humanitarian sector

- The conditions under which the sector operates are invariably arduous, dynamic and often dangerous.
- Humanitarian operations involve combinations of organisations from different countries. In the initial phases, the sector operates under significant time pressures and often under the intense gaze of the media. In any particular context, there may be significantly differing assessments of needs, differing views as to how these are best provided for and divergent opinions about each agency's role therein.
- Resources flow from the top down, and beneficiaries exert little or no influence on the way assistance and protection services are provided.
- Incentive structures in agencies promote defensive behaviour and a culture of blame.
- Short-term funding mechanisms militate against a learning environment for field staff.
- There are very high rates of staff turnover.
- There is a lack of clarity as to the objectives and desired outcomes of interventions.
- Training is not properly linked to learning processes.
- Mechanisms for cross-organisational learning are poorly developed.

Source: ALNAP 2002 and 2003.

Systemic factors may inhibit impact analysis. The imperative or incentive for agencies implementing interventions to conduct impact assessment is reduced by the lack of flexibility in the international humanitarian aid system. Following an initial emergency assessment, and approval by donors for resources for a particular intervention (e.g. food aid, livestock off-take, cash transfers), agencies are aware that, if they have got it wrong and the intervention is inappropriate or inadequate, donors are unlikely to be sufficiently flexible to allow for the reallocation or redirection of resources.

3.3.2 Lack of consensus over the objectives of humanitarian aid

In order to assess impact, a degree of clarity over the objectives of humanitarian aid is required, whether at project level or more globally. The impact of a food aid programme is significantly different if its objective is to save lives or to sustain livelihoods; in each case, different indicators will be required. Hallam (1998) argues that a lack of consensus on what constitutes a humanitarian outcome is one of the principal challenges to measuring impact.

A 2003 study conducted by the Humanitarian Policy Group on needs assessment and decision-making in the humanitarian sector (Darcy and Hofmann, 2003) highlighted that the aims and objectives of humanitarian action are often disputed. The study suggested a simple, core definition: 'the primary goal of humanitarian action is to protect human life where this is threatened on a wide scale'. This incorporates both traditional relief assistance, such as food, health, water or shelter interventions, and activities aimed at protecting civilians from violence, coercion and deliberate deprivation. In that sense, it includes the protection of lives and dignity.

The concepts of dignity and protection are important values, although they are hard to translate into 'measurable outcomes'. Whilst technical knowledge has significantly progressed in the traditional assistance sectors of food, health, shelter, water and sanitation, there remains relatively little experience in estimating the impact of interventions on the protection of civilians. Standards and assessment methodologies do not exist, and there is no shared understanding of what is involved. Humanitarian protection is not susceptible to the commodity-based approach that tends to characterise humanitarian assistance, nor to the kind of quantitative analysis that may underpin it (Darcy and Hofmann, 2003).

There are initiatives that aim to define indicators of protection, which can be used for assessing the impact of humanitarian aid. For instance, ALNAP (2004) has published a guidance booklet on protection which deals with the problems of translating values into measurable outcomes. Leaning and Arie (2001) have developed a

framework that combines material and psychosocial needs in the concept of human security. They propose to add to the provision of basic material supports, which is essential but not sufficient, three basic psychosocial dimensions: (a) a sense of home and safety; (b) constructive family and social supports; and (c) acceptance of the past and a positive outlook on the future. This combines the need to ensure human survival with the need to sustain and develop a core psychological coping capacity in populations under stress.

There is also disagreement over the extent to which the humanitarian agenda should include notions of sustainability or capacity-building. Attempts to analyse the impact of development interventions often include a wider set of criteria, such as financial and institutional sustainability, replicability, scaling up and impact in terms of strengthening civil society. However, it remains unclear how far these wider criteria should form part of impact analysis in the humanitarian sector (Kruse, 2003).

3.3.3 Lack of baseline data and seasonality issues

Attempts to analyse the impact of humanitarian interventions are often handicapped by a lack of baseline data and knowledge about regular seasonal variations in key indicators. It is difficult to show that a humanitarian programme has had an impact without knowing the rate at which something was occurring before the intervention began, and after it was implemented. Likewise, when people are arriving in a new location or are returning home, it is often impossible to determine the baseline before their arrival. In those cases, established norms can be applied as an assumed baseline, or as a threshold above or below which the indicator should not fall. Programmes which keep mortality low or keep water and food provision high may be successful in meeting their objectives, but if they do not have a baseline rate or a comparison group they may not be able to quantify the impact of the intervention.

The lack of baseline data is a recurrent problem in humanitarian programmes. Baseline CMR are often not known. Countrywide figures may be unreliable, out of date or inappropriate. A related problem is the lack of reliable population statistics. Humanitarian programmes rely on scarce and often inaccurate information on population figures. The lack of baseline data for nutrition has created particular interpretive difficulties. Although an intervention may be associated with reduced levels of wasting, it will not be possible to know whether the intervention has led to 'normal' levels of wasting. In areas of endemic HIV infection there is growing evidence of unusually high levels of severe wasting, but without baseline data it may be difficult to interpret the findings of post-intervention nutrition surveys. Similarly, a lack of baseline data on seasonal variation in nutritional status

makes it difficult to interpret repeated nutritional surveys/surveillance during a project cycle (Young and Jaspars, 1995).

3.3.4 Control groups and regression analysis

Control groups are a commonly used research tool in the social sciences, where changes over time between those affected by a project and those outside a project can be compared. However, the technique has rarely been used for analysing the impact of humanitarian aid (Roberts et al., 2001; Tomashek et al., 2001). For ethical reasons, it is difficult to deliberately exclude a group from access to life-saving relief. It may happen that some particular groups do not receive relief, due to problems with access or lack of resources; but, as Hallam (1998) warns, comparisons between people who received assistance and those who did not need to be used very carefully. For example, mortality in the two groups may be equal, but the group not receiving aid may have had to sell important assets to ensure its continued survival. Aid, and particularly emergency aid, is highly fungible, and is shared between those who are targeted and those that are excluded, making comparisons between those who received aid from an aid agency and those that did not difficult.

This does not necessarily mean that opportunities to compare people that are targeted in an intervention with those that do not receive aid cannot provide useful insights if carefully handled. For example, in a situation where only a percentage of the population is being targeted, or where one district receives aid and a neighbouring area does not, comparisons might be possible. Another imperfect but possibly useful way of looking at more 'ethical' approaches to control groups is to use new populations joining an intervention, to act as a form of control for populations already in a scheme. The broader point is that such comparisons are crucial to impact analysis; while it may not be possible to construct rigorous control groups, some sort of comparison may often be possible. There are, after all, few occasions where humanitarian aid is delivered to entire populations.

There may also be situations where a new type of programme design, such as clinic-based therapeutic feeding, discussed in Chapter 1, can be compared with the impact on a control group assisted with a traditional type of programme. Box 11 summarises an economic impact assessment carried out by Oxfam in Wajir, Kenya. This suggests an example of the use of control groups in practice (Odhiambo, Holden and Ackello-Ogutu, 1998). There may also be opportunities to learn from technical innovations in development evaluations, such as propensity score matching (White, 2003; Ravallion, 2002). The objective here is to match each person in the treatment group with a person in a control group, where matched pairs are as similar to one another as possible.

Box 11: Oxfam's Wajir Pastoral Development Project: an economic impact assessment

Oxfam carried out an impact assessment of a pastoral development project in Kenya's Wajir District in the late 1990s. The project developed Pastoral Associations, supported local NGOs, established a private sector delivery mechanism for veterinary and medical products through community-based health workers, restocked destitute pastoral families, supported women's groups in savings and credit schemes and provided inputs for water resources and rural schools.

Data on impact was collected through a two-week survey of pastoral households. Focus group discussions and interviews with participants in the restocking and savings and credit programmes were also conducted. The assessment measured the direct economic benefits of the project as:

- A lower livestock mortality rate in project sites compared to non-project sites due to improved animal health services and more reliable water supplies. The reduction in livestock mortality resulted in an estimated annual saving of £490,000 to pastoral communities over the lifetime of the project.
- An increase in livestock capital (over and above the original donated animals) and milk production valued at £101,000 from the restocking project.
- Returns of £44,000 on the credit made available to women's groups.
- A reduction in livestock theft valued at £30,000 through the activities of the Wajir Peace and Development Committee.
- A modest reduction in drug expenditure of some £9,726 due to cheaper drugs.

Additional benefits of the project included the empowerment of women and pastoral households to undertake collective action. Overall, these benefits led to a 19% reduction in demand for food aid among project beneficiaries compared to non-project households, a three-fold increase in milk consumption among poor households and a rise in confidence in their ability to survive future droughts, and thus continue with pastoralism as a way of life.

Source: Odhiambo, Holden and Ackello-Ogutu, 1998.

Where it is not possible to create some sort of control group comparison, economists and other social scientists have used a statistical analysis of determinants, which usually means a regression-based approach. There is a large literature in the development sector drawing on statistical techniques such as regression analysis to examine issues as varied as the effectiveness of aid in stimulating economic growth or the determinants of child survival (White, 2001). However, the data demands of such an approach mean that

no examples of regression techniques being used to examine impact in the humanitarian sector were found in this review. The emergency context often means that limited data is available, many of the variables influencing humanitarian outcomes such as mortality are difficult to measure and expertise in quantitative and statistical analysis is limited within the humanitarian sector (Hallam, 1998).

3.3.5 Problems of attribution

While impact may well be measurable (a particular change has occurred in a given context, such as a decrease in mortality, or an improvement of the nutritional status of the population), it is not certain that it can be attributed to a particular intervention. Problems of attribution are due in part to the difficulties of isolating the particular impact of humanitarian aid relative to other factors, such as changes in security, market conditions locally and globally and climatic factors. Have lives been saved because of food aid, a health intervention, a water and sanitation programme or local coping mechanisms? The answer may well be a combination of all of these, but is it possible to determine the relative significance of these different elements of response? The multitude of actors at different levels within the humanitarian system makes the problem of attribution even more complex. The further away from the actual implementation of activities, the harder it becomes to attribute impact.

De Waal (1997) argues that 'evaluation should be concerned with the question how humanitarian aid fits into and complements people's coping and livelihood practices'; relief assistance is, he contends, 'often a relatively small contributor to people's survival in emergencies'. Other studies focusing on the capacity of populations to survive in crisis situations show that humanitarian aid's contribution is only part of the picture (Hansch et al., 1994; Lautze, 1997). The signature of a peace agreement, a favourable harvest and local coping mechanisms may have as great an impact on food security, for instance, as a food aid intervention. This does not diminish the importance of food aid, but it does show that impact is difficult to attribute to a single intervention.

Difficulties of attribution are perhaps particularly acute in the nutrition sector. Due to the multi-causal nature of malnutrition, it is impossible to attribute the impact of a nutrition or food security intervention at a population level without controlling or accounting for other sectoral interventions, such as health, support for caring practices, water and sanitation or income support. It would also be necessary to account and control for other food security factors not related to the intervention, such as changing access to markets.

In rare cases where only one agency intervenes in a specific area and where external influences are relatively limited in

scope, it may be possible to attribute impact or results to a particular intervention. In the case of an MSF-H trypanosomiasis control programme in northern Uganda, for example, it was possible to attribute reductions in disease prevalence to the programme because nothing was done before it began, and no other control activities were implemented in the region during the five years MSF was active:

It is fair to say that the evaluation was not only successful because of what we did and how we did it but also because it was a vertical control programme for which simple impact indicators could be determined. Furthermore, because of the characteristics of the disease, morbidity and mortality figures measured by the evaluation could easily be associated with the MSF intervention (MSF Holland 1996:59).

3.4 Impact assessment beyond the project level

A large proportion of impact assessments look at the impact of specific projects carried out by a single agency. However, there is increasing interest in higher levels of impact analysis: several initiatives and mechanisms, by donors and humanitarian agencies alike, are attempting to move beyond the project level to consider the sectoral, multi-sectoral or system-wide impacts of aid. New contractual arrangements between donors and humanitarian agencies increasingly incorporate institutional partnerships such as framework agreements (Macrae et al., 2002). In turn, accountability mechanisms are shifting away from projects, and humanitarian agencies are increasingly being required to report results at a higher level.

Impact analysis at the sectoral level has been rare; humanitarian agencies have tended to concentrate on sectoral coordination mechanisms and sectoral needs assessments. An initiative to establish inter-agency health programme reviews in humanitarian situations has been developed, and a number of health reviews are under way (Inter Agency Health Programme Review, 2003).

Impact assessment at a system-wide level is generally found in large evaluations, such as the evaluation of the 1994 Rwanda response. These evaluations provide a multi-sectoral analysis, and can give useful information about non-aid factors and the wider context of aid. One of the objectives of the SMART initiative is to enable judgements about the overall impact of the humanitarian effort: crude mortality rates and malnutrition rates can be seen as critical indicators to assess the effectiveness of relief programmes considered globally.

The last type of wider impact assessment concerns a single agency or donor assessing its impact beyond the project level. Donors have attempted to measure the impact of

their humanitarian spending across a range of contexts (Danida, 1999; USAID, 2000), and some agencies are attempting to measure the global impact of their interventions. A number of British NGOs have set up systems to monitor global impact, mainly in development work though this may increasingly cover humanitarian aid programmes.

Assessing impact at the organisational level raises major difficulties around the question of aggregation and the comparability of data from different contexts. Organisations have found it difficult to report on impact at the global level through these systems (British Agencies Aid Group, 2002; Save the Children UK, 2004). Clearly, a donor looking at the overall effectiveness of its aid over a number of years needs far more aggregation than would be needed in the evaluation of the impact of a single project conducted by a single agency. In situations with a multiplicity of actors, different forms of programme and an equally large number of reporting systems, this is a challenge. According to an OECD/DAC review of donor-commissioned studies of the impact of NGO development projects: 'one should be careful about comparing the performance of NGO development projects in different geographical locations, and in different socio-economic and political settings, and even in different periods' (OECD/DAC, 1997).

Wider levels of impact assessment also generate problems to do with responsibility and coverage. Who is responsible for the collective impact of a number of individual humanitarian projects? Who will account for the overall success or failure (if this can be measured) of the humanitarian enterprise? As for coverage, while the impact of an individual project may be satisfactory, the overall impact at the sector or country level may be less so since some part of overall need will not be covered. Despite coordination efforts and mechanisms, notably through the CAP/CHAP process, a sense of collective responsibility remains relatively weak.

3.5 Measurement or analysis?

The methodological constraints to measuring impact suggest that proving impact in a scientific sense, in the demanding environments in which humanitarian aid operates, is a particularly difficult proposition. There may be some situations in which analysing rather than measuring impact is more appropriate. Maxwell and Conway (2003) suggest that 'if impact cannot be established scientifically or precisely (i.e. cannot be measured), it can at least be estimated and described by postulating and testing logical linkages between aid activities and observed changes'. Often, observation and judgement are useful complementary elements to 'measurement'.

Table 3: Mayne's 'contribution analysis' applied to the southern Africa crisis

Mayne's criteria for analysing the contribution of an intervention	A 'contribution analysis' applied to the 2002 southern Africa crisis
Acknowledge the problem	There are clearly myriad factors determining how people survived across six disparate countries in southern Africa in 2002 and 2003. Assessing the relative contribution of the humanitarian aid effort is very difficult
Analyse and present the logic of the problem	Why is it logically plausible that, given the needs identified (14 million people at risk), the aid delivered would have contributed to the eventual outcomes (few peaks in mortality or malnutrition)?
Explore and discuss alternative explanations	A possible alternative is that few people died because: <ul style="list-style-type: none"> • Needs were wrongly identified • Coping strategies were under-estimated • Commercial food imports or government efforts were greater than expected
Defer to the need for an evaluation	There have been several evaluations of project, sector and agency performance in southern Africa. Few, however, have addressed broad questions of impact in any detail. A well resourced system-wide evaluation would probably be needed to answer the questions whether lives were really at risk, and the international aid effort helped to save them

Source: Adapted from Mayne, 1999

Mayne (1999) argues that, in most cases of any complexity, it may not be possible to determine definitively the extent to which a programme contributes to a particular outcome; instead, it is more appropriate to think in terms of reducing uncertainty about the contribution of a programme. Mayne uses the term 'contribution analysis' to argue that, in dealing with attribution using performance measurement information, the need is to explore the issue in a systematic way, and present a credible performance story for attribution within the available evidence. This would entail;

- Well-articulated presentation of the context of the programme and its general aims.
- Presentation of the plausible programme theory leading to the overall aims.
- Highlighting the contribution analysis indicating that there is an association between what the programme has done and the outcomes observed.
- Pointing out that the main alternative explanations for the outcomes occurring, such as other related programmes or external factors, have been ruled out or clearly have had only a limited influence.
- If this is not enough and there are too many gaps in the story, one ought to admit it and accept the need for an evaluation to provide better understanding of the contribution of the programme.

A hypothetical and broad-brush application of contribution analysis to the 2002/2003 southern Africa crisis is shown in Table 3 (on previous page). This demonstrates how analysing the relative contribution of aid in a systematic way can help to nuance the bold claim that millions of lives were saved because of the quick response of the international community.

3.6 Chapter summary

This chapter has identified a series of constraints to the analysis of impact. Clearly, impact cannot be simply or straightforwardly measured, particularly in the dynamic and chaotic environments of complex emergencies. There is a lack of consensus as to what properly constitutes a humanitarian outcome, and serious constraints are imposed by the lack of accurate, useable demographic and other data. Techniques that are standard in the social science community, such as the use of control groups, are not widely used, and humanitarian practitioners tend to lack the skills needed to gather and interpret information. Finally, there are fundamental problems around the attribution of impact that cannot easily be resolved. This does not, however, imply that progress is impossible; impact can often be analysed, even if it cannot be measured. The next chapter looks in more detail at how humanitarian agencies have approached the problem of measuring impact in the health and nutrition sectors.

Chapter 4

Impact assessment: current practice in the health and nutrition sectors

This chapter examines humanitarian agencies' current practice in assessing impact in the health and nutrition sectors. The chapter relies heavily on the two background papers commissioned for this study (Roberts, 2004; Shoham, 2004). It is based on published and grey literature, and so cannot be a comprehensive review of current practice. The health, nutrition and food security sectors were chosen as illustrative of wider patterns, but are not fully representative.

4.1 Impact assessment and the health sector

This section is based on a background review of 15 final reports of health-related programmes funded by the US Bureau of Population, Refugees and Migration (BPRM), submitted in 2003 (see Roberts, 2004). These reports were evaluated against the following five criteria:

- Was there a health-related objective?
- Was the baseline rate measured or a comparison group identified?
- Was the health-related outcome measured and reported?
- Was the societal level of the evaluation appropriate given the intervention?
- Were there any major issues supporting or raising concerns about the reported outcome data?

The societal level of a health project was categorised as being either on the patient level, the household level or the community level. The expectation was that programmes that intervened on a specific level should be evaluated on that level. For instance, a curative health programme might have benefits at the individual level, but it may not be possible to evaluate its impact at a wider level.

Six of the 15 reports did not attempt to measure or report any health-related rates or status. Proposals corresponding to five of these six reports only contained process indicators as the objectives, which meant that the lack of documented health benefits was assured before the projects began. An additional three of the 15 reports contained health data-based objectives but did not present any health-status data, instead reporting process indicators such as the number of clinics supported or consultations given. Only four of the 15 final reports could demonstrate a health benefit, and three others were likely to have produced a population-based benefit although this was not documented.

The results of this analysis confirm the general conclusion reached at a July 2002 SMART Monitoring and Evaluation Workshop: that while NGOs and agencies often want to monitor health outcomes, they usually monitor process indicators (USAID, 2002a). Problems with process indicators seen in the BPRM review include:

- The cited activity may be related to the health outcome, but the significance of this effort depends on the activities being done well and in sufficient numbers.
- The health-related objective is only distantly related to the health outcome.
- In some cases, the links between the process indicator and the outcome were simply implausible.

According to OFDA, only two of the 20 NGOs it or ECHO funded in the eastern Democratic Republic of Congo (DRC) in 2000 and 2001 could show health benefits associated with their programmes (pers. comm., 29 January 2002). According to a CDC review of project reports in Somalia in 1991–93 (Boss, Toole and Yip, 1994), the range of methodologies employed in surveillance and surveys and outcomes measured were so variable, and of such poor quality, that widespread comparisons were impossible and much of the data was not credible. Spiegel et al. (2002) reviewed 125 nutritional surveys in Ethiopia in 1999 and 2000. The surveys were carried out by 14 organisations, with a wide range of survey expertise. Only 67 of the 125 attempted to identify a sample that represented the population served. Only six of the surveys possessed the minimum number of clusters (30) and children (900) suggested by most nutritional manuals. Most survey reports did not describe what sampling methods were employed, and few presented confidence intervals around the results. Sixteen reports were 'rapid assessments', with no attempt to take a representative sample. These unstructured surveys measured an average global malnutrition of 32% and severe malnutrition of 5%. This contrasted with the 67 surveys that attempted to be population-based, which found 12% global and 1% severe malnutrition. Spiegel (2002) concluded that most of the surveys were of such poor quality as to be unhelpful in making relief policy decisions.

The measurement of nutrition is relatively standardised compared to other health outcomes, such as mortality and mental health status. For some project objectives, such as the prevention of HIV transmission, there is not even an agreed outcome to be measured. The difficulty of assessing

outcomes such as mortality is a principal reason for the use of process indicators in place of health outcomes. Without improved staff skills and capacity, it is likely that humanitarian agencies will continue to rely heavily on process indicators, and will not be expected to prove that programmes have influenced the health of the targeted beneficiaries. Donor requirements can have an effect here. For example, CIDA, within its Programme Against Hunger, Malnutrition and Disease, has started to pilot reporting requirements for its non-food aid emergency programmes. These require agencies to attempt to measure the cost-effectiveness of interventions in terms of a dollar amount per death averted. CIDA believes that this is beginning to stimulate improvements in how agencies report against and document health impacts (pers. comm., 2004).

Two examples of projects that have attempted to measure impact and produce data on their cost-effectiveness are given below. This suggests that, even in difficult and unstable situations, it can be possible to demonstrate impact given sufficient commitment and skills in programme design, implementation and monitoring.

Starting in December 2000, the International Rescue Committee (IRC) began a general health programme to support government services in Katana Health Zone in the DRC. The agency conducted population-based mortality surveys in the area, which had 345,000 mostly rural residents. The programme consisted of the provision of drugs, supplies, training and medical oversight in clinics, water provision and hygiene education in villages, measles immunisation and vitamin A provision and support to local health committees, including the donation of vouchers for the most indigent members of the community. IRC claims to have reduced the excess CMR by 60% (from 4.9 to 2.8 deaths per 1,000/month, where the baseline is assumed to be 1.5) between six and 12 months after implementation, and by 70% (from 2.8 to 1.9 deaths per 1,000 per month) between 12 and 24 months after implementation. IRC reports that the total cost per death averted (including overhead and management salaries) was \$227 the first year, and \$132 over the two years of funding.

In support of its view that these results were a consequence of the health programme, IRC reported that:

- Attendance at the clinic rose by 147% between 1999 (around 7,400 visits per month) and 2001 (around 18,300 visits per month).
- Seventy per cent of treatments were for malaria and diarrhoea, the main reported causes of death, and deaths from these diseases decreased in 2001 and 2002.
- CMR in the five eastern provinces of DRC was estimated by IRC to have increased slightly in 2001 compared to 2000.

- A survey in November 2001 found that 60% of residents that had experienced fever in the preceding two weeks had sought treatment at a clinic.

Employing Hill's criteria of causation (described in Chapter 3), there is a baseline; the project goals, most of the implementation and the evaluation were on the population level; the benefit occurred after implementation; the findings are biologically plausible (although one visit per resident per year seems low); and alternative explanations for the reductions cannot be ruled out given the variance over time and the dramatic changes in violent conflict. Although IRC reports that the violence did not dramatically subside until 2002, the magnitude of the reduction and the fact that IRC's two other areas of health programmes had similar reductions (but somewhat less dramatic) implies significance and repeatability – that the impact of the health programme could be replicated elsewhere. Finally, the fact that the CMR was measured by an apparently valid survey method implies that IRC probably did reduce mortality in Katana.

In the second example, this time in Kisangani in DRC, IRC supported laboratory activities in hospitals, specifically the testing of blood supplies for HIV, which had not been done for the preceding three years. At the start of the programme, 7% of blood donors were HIV-positive, and 200 transfusions were being done per month, almost exclusively among children experiencing extreme anaemia induced by malaria. At the end of the programme two years later, 17% of blood donors were HIV-positive and 120 transfusions were being done monthly. IRC assumed that all transfusions of HIV-positive blood could infect an individual, that all blood would be used (as it was always solicited for a specific patient) and that 93% of recipients at the start of the project and 83% at the end of the project were HIV-negative, and capable of being infected. IRC reported a total programme cost of \$476 per case of HIV averted at the start, and \$327 at the end.

4.2 Impact assessment and the food and nutrition sector

This section is largely based on a review of current practice in impact measurement and analysis in the food and nutrition sector (Shoham, 2004). Food and nutrition interventions usually aim to save lives, reduce levels of malnutrition (wasting and sometimes micro-nutrient deficiency), prevent increases malnutrition and, increasingly, protect livelihoods (WFP, 2003a). Sometimes these objectives are stated explicitly; for example, nutrition programmes may aim to reduce levels of wasting to pre-emergency levels (where information for this exists), or to below levels that would trigger an emergency response, such as less than 10%. In the case of micronutrient deficiency, the objective is more likely to be expressed in terms of eradicating the incidence of disease. Objectives with regard

to mortality will be stated in terms of reducing or maintaining the crude mortality rate at levels seen in a normal population. Food security objectives may be stated in terms of ensuring access to sufficient food to meet daily requirements. Food security programmes may also aim to prevent activities which are detrimental to households or the community, such as the sale of assets or distress migration.

Food security and nutrition information has rarely been used to assess intervention impacts at the population level. However, for some types of intervention, such as therapeutic feeding programmes, impact is routinely and rigorously assessed at the project and individual beneficiary level.

4.2.1 General ration programmes

Assessing the impact of general rations through monitoring of nutrition indicators is problematic. In most cases, where general rations are triggered by a high prevalence of wasting (i.e. > 20% or 10–19% with exacerbating factors), selective feeding programmes will also have been implemented. Thus, a reduction in levels of wasting will in part be due to these programmes, as well as the general ration. It will therefore be difficult to separate out the impact of general rations from selective feeding on the prevalence of wasting.

The most feasible means of assessing the impact of general rations may be to combine assessment of process indicators and nutritional impact monitoring. Key process indicators are food basket monitoring (for quality and quantity) and targeting (inclusion and exclusion error). Information on speed of change in population-level nutritional status in relation to the onset of a general ration programme will also strengthen confidence in the assessment of impact. Equally, if disruption to the general ration leads to a rapid deterioration in population-level nutritional status, impact can be confidently inferred (providing there is no evidence of other simultaneous marked changes in food security or health circumstances).

4.2.2 Selective feeding programmes

All selective feeding programmes (therapeutic and supplementary) collect data on key programme-related outcomes/impacts, including weight for height, average daily weight gain, mortality, default and average length of stay in programmes. For health centre-based therapeutic feeding programmes, where beneficiaries are entirely dependent on the programme inputs, these indicators clearly show programme impact at the individual and project-beneficiary level. For supplementary feeding programmes (SFPs) the link between programme and impact at individual or project-beneficiary level is less clear-cut, as other food security factors can have a significant impact on programme performance. Targeted SFPs are predicated on the basis of adequate household

food security (either through general rations or some other means). Frequently, however, SFPs are implemented in the absence of adequate food security/general rations, so impact would be expected to be compromised.

All selective feeding programmes produce monthly, mid-term and end-of-programme data summaries. Although these describe key outcomes in terms of nutritional impact and mortality at programme level, they cannot be used to infer impact at population level unless there is good data on programme coverage, default and readmissions. Obtaining good data on programme coverage is not straightforward, and there is currently some debate as to the most appropriate methodology for doing this (Emergency Nutrition Network, 2003a).

There are established minimum standards for the expected impact of selective feeding programmes in emergencies (MSF, 1995; Sphere, 2004). However, although each agency evaluates the impact of these programmes individually in terms of targets (e.g. average daily weight gain, mortality), there is currently no overview of impact across a large number of selective feeding programmes. This is of concern because many factors can undermine impact (especially of supplementary feeding programmes) in emergency situations, such as breaks in the food aid pipeline and overcrowding at feeding centres.

A review by Beaton and Ghassemmi (1982) concluded that supplementary feeding programmes in stable situations had little impact in terms of growth performance. These findings arguably had a significant influence on the perception of these programmes, which in turn contributed to a reduction in their scale. A similar review of emergency supplementary feeding programmes is long overdue, as it cannot be assumed that these types of emergency programme automatically achieve their objectives, or that if they do, they do so in a cost-effective manner.

The lack of a comprehensive overview of the impact or effectiveness of supplementary feeding at project level can partly be understood in terms of the various pressures which encourage their implementation, such as the inability of smaller agencies to resource and implement general ration programmes, or delays in general ration implementation. It can also be explained in terms of the absence of any body/institution with a mandate to coordinate and evaluate emergency nutrition activities. Agencies such as ICRC have policies whereby SFPs are rarely implemented, so that general ration provision is planned on the basis of catering for the additional nutritional needs of the mild and moderately malnourished. It has been argued that these expanded general ration programmes have been as effective in meeting the needs of the mild and moderately malnourished as general ration programmes in conjunction with SFPs (Curdy, 1994).

Another constraint on assessing the impact of SFPs is the lack of clarity or explicitness of objectives. The impact and effectiveness of these programmes can only be measured in relation to stated objectives and delivery goals (process indicators). The objectives for supplementary feeding programmes are often narrowly defined as preventing mortality amongst mild and moderately malnourished individuals (targeted supplementary feeding) and preventing increasing levels of malnutrition at population level (blanket supplementary feeding). However, other objectives for these programmes may also be invoked, such as ensuring food access in situations of conflict, where general ration distributions may be targeted by combatants; or enhancing the household food security of refugee-impacted households (Borrel, 1997). These types of objective rarely appear in agencies' emergency nutrition guidelines, but may be stated in programme proposals submitted to donors. Such objectives are often country- and population-specific. However, they may not be stated as explicit objectives or may be stated in such a way (e.g. qualitatively) that it is difficult to measure whether the objective has been achieved. Clearly, in the examples given above some form of food security monitoring would need to be introduced to test whether programmes have achieved the desired impact.

Although there are many examples where improvements in a population's nutritional status as evidenced by repeated nutritional surveys is attributed in part to emergency selective feeding programmes, there are very few examples of studies which explicitly set out to analyse the impact of these programmes at the population level. Where this has been done, the findings have been inconclusive. For example, a study of the Zimbabwe Community Supplementary Feeding Programme (Munro, 2002) focused mainly on coverage and targeting (whether the most malnourished were enrolled). These proxy indicators for population-level impact showed only moderate coverage and a high level of exclusion and inclusion errors.

In the case of therapeutic feeding programmes, attempts at impact assessment are problematic due to the generally low coverage of health centre-based programmes. Indeed, this is a main reason for the advent of community-based therapeutic feeding programmes (discussed in Chapter 2), which promise far greater coverage through out-patient and community-based care. It may be possible to attribute the impact of therapeutic feeding programmes on severe malnutrition at a population level where there is good coverage (for example over 80%, though this is extremely unusual), and where the programme data shows that targets are being met. This raises again the point that it is crucial at what level impact is measured. Therapeutic feeding programmes, where impact is measured at an individual or a project level, may seem to be very effective, but if

coverage is low their impact when analysed at the level of the wider population can be relatively small.

4.2.3 Interpreting nutritional status data

Using malnutrition rates as an indicator of impact can be dangerous, and there is a need for contextual information in order to be able to interpret and understand the factors behind malnutrition levels. Reliance on child anthropometry alone will not provide an understanding of factors which are determining current nutritional status, or which are likely to influence short-term nutritional trends. Nutritional survey results can mask imminent famine unless combined with food security/livelihood analysis; low rates of malnutrition may exist alongside (and therefore mask) a severe erosion of livelihoods and exhaustion of coping strategies. Furthermore, there is often a need to respond before nutritional deterioration can be measured. Relying on malnutrition and mortality rates as indicators of impact may be problematic given that the objective of some humanitarian aid programmes may be to prevent these rates from climbing.

In some situations, micronutrient deficiency disease outbreaks may occur before widespread protein energy malnutrition, so that nutrition indicator monitoring systems need to expand indicators to include micronutrient status. Analysis of the nutritional status of adults (BMI) in conjunction with the nutritional status of under-fives in the same household can help determine the degree to which nutritional problems are related to disease/caring practices, rather than food security constraints (James et al., 1999). Furthermore, in contexts where child nutritional status is protected at the expense of adult food consumption, measuring the nutritional status of adults can lead to earlier detection of nutritional stress caused by food insecurity. There is a need for research into how best to integrate nutritional indicator and food security assessments; this will pose institutional as well as technical challenges.

4.2.4 Micronutrient status

Micronutrient status is rarely assessed at the onset of an emergency situation. There are several reasons for this:

- the lack of a field-friendly method of assessment (Seal, 1998);
- the relative rarity of clinical micronutrient problems in emergencies, especially since the introduction of fortified blended foods (CSB) into general rations in the early 1990s;
- low levels of wasting can mask poor micronutrient status (Assefa, 2001);
- some forms of deficiency disease such as pellagra and scurvy appear to affect older age groups more predominantly, and would therefore be missed in a standard nutritional survey which measures and weighs children under five years of age (Duce et al., 2003); and

- no practical tests are available which meet all the necessary criteria for biochemical assessments in the difficult circumstances encountered in the field (Seal, 1998).

It is imperative that the impact of interventions to address micronutrient outbreaks is assessed through a combination of monitoring the micronutrient adequacy of food rations (accounting for tablet distribution if appropriate) and clinical information on the incidence of diseases following an intervention. Food rations can be monitored through food basket monitoring at distribution points or within households. The speed at which incidence of the disease diminishes following an intervention is a key indicator of impact. Most micro-nutrient deficiency diseases respond rapidly to improved diet (Duce et al., 2003). Further discussion of questions relating to measuring the impact of micro-nutrient interventions are in the Shoham (2004) background paper.

4.2.5 Livelihoods approaches

Humanitarian interventions are increasingly aiming to protect livelihoods, as well as preventing mortality and protecting nutritional status. Operational agencies which specify livelihoods protection as an intervention objective rarely do so in a quantitative manner. Thus, objectives will be framed in terms such as preventing the sale of key assets, preventing distress migration or preventing indebtedness.

A WFP technical consultation on emergency needs assessment has identified the information needed to determine whether livelihoods were being protected by an intervention. Relevant information included:

- estimates of primary asset liquidation;
- changes in productive capacity (human capital);
- population movements (political or economic); and
- changes in market conditions induced (commercial and wages).

In order to assess humanitarian impact at household level in terms of livelihoods protection, it would therefore be necessary to have baseline information on primary assets, human capital, normal population movements and market conditions. Without baseline quantification of these variables, any attempt at meaningful impact analysis would be problematic.

There appear to be very few examples of livelihoods-based tools of analysis being used to assess the impact of humanitarian interventions, although interest appears to be growing. Livelihoods-based assessment tools have mostly been used to conduct emergency needs assessments. A range of assessment tools have been developed, including household economy approaches used by Save the Children, livelihoods approaches used by Oxfam and CARE and WFP's

vulnerability assessment and mapping methods. These approaches are all based to varying degrees on entitlements theory and concepts of vulnerability and coping strategies (Jaspars and Shoham, 2002).

Many factors militate against impact assessment using livelihoods frameworks. In addition to the broader methodological constraints already discussed, there may also be other factors more specific to livelihoods-type information. Data on coping strategies may be difficult to elicit where these strategies are perceived by the community as unlawful, immoral or damaging (Jaspars and Shoham, 2002). This is likely to pose particular problems in situations of chronic conflict and instability. The lack of easily quantifiable intervention objectives also militates against the assessment of impact on livelihoods. Unlike nutrition objectives, which are usually expressed in terms of reducing the prevalence of wasting to below a specified level, setting similar quantifiable objectives for livelihoods is less straightforward. What level of asset protection is desirable or recommended? What degree of loss of social capital is sustainable or allowable? These are difficult questions. Furthermore, measuring these types of variable poses challenges in sampling design, sample size and methods of quantification (Shoham, 1991). Another difficulty is the lack of consensus and standardisation of livelihood assessment methodologies.

There are also potential advantages to livelihoods approaches to assessing impact. One significant benefit is that livelihoods-type assessments invariably rely on key informant interviews or focus group discussions with target groups, beneficiaries or affected populations. In contrast to nutrition surveys, these techniques allow beneficiaries to explain how and whether interventions have impacted their lives/livelihoods and food security.

In 1999, Save the Children UK assessed the impact of food aid on household economies in three areas of Ethiopia (Save the Children UK, 1999). The assessment utilised the HEA approach. The strategy to measure food aid impact included four principal stages:

- Quantifying food aid distributed in the study area.
- Quantifying food aid received at household level and stratifying households by wealth group.
- Tracing utilisation of food aid for consumption, sale, redistribution, exchange, paying back loans and other uses.
- Assessing the impact of each method of utilisation on household economy and food security. This included changes in food consumption (quantity and quality), changes in expenditure patterns and changes in income generation strategies (for example labour migration, the sale of productive assets and whole household migration).

Key findings were that:

- For poorer and some middle-income households that had migrated for work, food aid had encouraged return to home areas.
- Food aid constituted a significant proportion of the diet but was insufficient to prevent hunger. A large percentage of households experienced a deficit in 1999.
- Food aid effectively prevented the sale of animals for grain in the months when food aid was delivered. However, households were forced to increase animal sales in months when distribution did not occur in order to buy grain and non-food items.

CARE's Coping Strategy Index (CSI) is a new tool to establish baseline data specifically for monitoring trends and judging the impact of humanitarian programmes (CARE and Save the Children US, 2002; CARE, WFP and ERREC, 2003). The CSI has been developed by WFP and CARE after initial piloting in Kenya (Maxwell, 2001). It enumerates both the frequency and the severity of coping strategies of households faced with short-term insufficiencies of food. It goes beyond commonly used calorie indicators to incorporate elements of future vulnerability. The CSI enumerates all consumption-related coping strategies commonly used by a population. The approach is designed as a rapid means of assessing livelihoods status through a set of proxy indicators related to consumption. Four general categories of coping are measured, with individual strategies defined specifically according to location and culture. These are:

1. Dietary change (such as eating less preferred but less expensive foods).

2. Increasing short-term food access (borrowing, gifts, wild foods, consuming seed stock).
3. Decreasing the number of people to feed (such as through short-term migration).
4. Rationing strategies (mothers prioritising children/men, limiting portion size, skipping meals, not eating for whole days). Work on the CSI in Ghana concluded that the index offers detailed information about people's decision-making and behaviour, and is much less time-consuming and less expensive in terms of data collection and analysis than using benchmark indicators like consumption, poverty or nutrition (Maxwell, 1999). In Eritrea and Malawi, CSI has been used to assess population vulnerability and to provide a baseline from which to measure future vulnerability in the context of humanitarian interventions – i.e, to measure impact.

4.3 Chapter summary

The two background papers carried out for this study largely confirm that, as a whole, impact assessment within the humanitarian community is poor. But they also point to promising developments, such as the use by IRC of mortality data to demonstrate health impact, or the development of CARE's coping strategies index as a way of assessing the impact of aid on livelihoods. The problem therefore seems to be less an absence of tools than a lack of skill and capacity to utilise them fully. Humanitarian field workers often lack the skills they need to carry out the sort of qualitative or quantitative assessments that would allow impact to be effectively analysed. Improving the analysis of impact will therefore require investment in improving skills and capacity.

Chapter 5

Conclusions and recommendations

This study has reviewed current knowledge about, and practice in, impact assessment in humanitarian assistance. To do so, it has drawn on experiences in other areas, such as international development, and from trends in the wider public policy environment in the West. Despite tremendous improvements in the technical and programmatic aspects of humanitarian aid, assessment of the impact of humanitarian assistance is still poor in comparison to the level of analysis common in development aid. The methodological and practical difficulties seem so great that it is tempting to conclude that it is unrealistic to expect meaningful analysis of impact in the humanitarian sphere. Yet the fact that the humanitarian system has not been particularly good at analysing impact does not imply that improvement is impossible. This study recognises the constraints to assessing the impact of humanitarian assistance, such as the volatile environments in which interventions generally take place, the high turn-over of staff, the lack of access to crisis situations and the short lifespan of many projects. However, these should not serve as justification for not considering more seriously the question of impact.

- The humanitarian system has been poor at analysing impact. Reviews of evaluations have consistently found that questions of impact are not adequately addressed.
- Promising approaches to the analysis of impact are starting to be developed. Examples include greater investments in gathering mortality data and the development of tools such as the Coping Strategy Index to analyse the impact of aid on livelihoods.
- Reviews have suggested that humanitarian field workers often lack the necessary skills to carry out the sort of improved qualitative or quantitative assessments that would allow impact to be effectively analysed. Improvements in the analysis of impact will therefore require investment in improving skills and capacity.

It is important that the measurement of impact is not reduced to a narrow set of technical questions at the expense of the wider context in which aid is delivered. The conceptual framework presented in Chapter 2 shows the wider dimension of impact, where a set of complex inter-related factors all have a degree of influence. The principles of humanitarian aid must not be sidelined, and an analysis of the full effects of humanitarian aid must not be restricted through a focus on measurable results. For example, humanitarian agencies in some contexts may have a positive impact on reducing threats to lives through

their presence, as much as through any measurable impact of activity.

5.1 Beyond the project level

Impact assessment has tended to focus on projects and programmes. There is a role for the analysis of impact at all of the various levels of assistance: the project, the programme, the sector, the country and the organisation. For example, sector-wide or system-wide assessment of impact would shed light on a number of important questions, such as the overall impact of the humanitarian enterprise, the coverage of humanitarian aid in a given context or the role of humanitarian aid in relation to other factors. The Rwanda evaluation remains the only system-wide evaluation of humanitarian aid. Such evaluations may be particularly important in enabling the large questions to be asked, and responsibilities for meeting humanitarian outcomes to be properly assigned.

- Concern for the impact of humanitarian aid should not be narrowly restricted to the project level. There is also a need for greater investment in approaches that aim to build up the evidence base of what works through more detailed research and for system-wide evaluations that can ask difficult and important questions about the responsibility for humanitarian outcomes and the broader political dimensions within which the humanitarian system operates.
- Project-based approaches that focus on determining the impact of a particular intervention through a causal pathway from inputs to impact should be complemented by approaches that start with changes in people's lives and situate change within the broader external environment.
- Questions of impact should not be considered only as part of the evaluation process. In the humanitarian sphere, a concern with significant change in the short term implies a need for impact to be considered in ongoing monitoring processes and through techniques such as real-time evaluation.

5.2 Measuring, analysing or demonstrating impact?

The particular constraints imposed by humanitarian emergencies may mean that a measurement of impact, with its quantitative, scientific connotations, is sometimes impossible. The best being the enemy of the good, it is important to acknowledge other dimensions of impact assessment: impact can be analysed or demonstrated

without being necessarily measured. The choice of method will depend on the objectives for which impact is being analysed, and the degree of validity that is expected. A useful distinction is between the measurement of impact in a scientific sense, and a more deductive approach that seeks to analyse it. Many evaluations tend to make statements about impact based on observation. A proper measurement of impact requires the collection of data, the existence of a baseline and adequate evidence of a correlation between the observed phenomenon and the impact. These different forms of impact analysis serve different purposes, require different methods and will imply considerable differences in the amount of time and technical expertise required.

- Impact in any context is difficult to measure and attribute; this is particularly so in the dynamic and chaotic environments of complex emergencies. This does not mean, however, that it is impossible. Greater efforts could be made.
- The humanitarian system often lacks the skills and capacity to successfully measure or analyse impact. Greater investment needs to be made in human resources and research and evaluation capacity if a greater focus on results is to be realised.
- The humanitarian system has remained consistently poor at ensuring the participation of affected populations. This is as true in impact analysis as in other aspects of the humanitarian response. Much could be learnt from innovations in participatory approaches in the development sphere, and possibly from customer-focused approaches in the private sphere. The humanitarian system is largely ignorant of the views of affected people as to the assistance being provided.
- There is a place for both the art and the science of impact measurement, and scientific, analytical and participatory approaches can often be complementary.

5.3 The evidence-base of humanitarian aid

Humanitarian aid tends to rely on limited evidence regarding both the nature and the impact of its actions. The fact that humanitarian aid often aims to prevent something from happening exacerbates the fragility of the evidence-base. Impact is often about showing that something that would have happened did not do so. This is relatively straightforward in some circumstances; in the case of a cholera outbreak, for example, the impact of an intervention to mitigate the risk of people contracting cholera may be relatively easy to demonstrate. The impact of a supplementary feeding programme, or a reproductive health care programme, or a programme to support primary health care, may be more difficult to assess. In these programmes, humanitarian agencies tend to use output indicators as proxies for impact, without sufficient evidence that there is a link between intervention and impact.

- Analysis of impact could be improved through greater clarity about the objectives of humanitarian assistance and more consistent assessment of needs.
- Process indicators can sometimes be used as proxies for impact when there is strong evidence between the action being monitored and an expected impact. An example would be measles vaccinations, which are known to reduce mortality. There is a need for greater investment in strengthening the evidence about how activities such as supplementary feeding or support to health clinics relate to humanitarian outcomes such as reductions in mortality or malnutrition.

5.4 The new public management agenda

Results-based management has placed high expectations on humanitarian agencies to demonstrate that they achieve positive impacts. While this study generally sees important benefits in shifting from a focus on outputs/activities to a focus on outcomes/impact, the rhetoric of results-based management is not always matched by the practice: some donors continue to focus on outputs or activities in their contractual arrangements and reporting. In that sense, there is insufficient incentive for humanitarian agencies to pay greater attention to the outcomes/impact of their action. Furthermore, perverse incentives may emerge in humanitarian aid, just as they have in the public domain.

- Results-based management systems are being introduced in a number of humanitarian organisations, but it is too early to say whether they will significantly improve the measurement and analysis of impact. Experience from elsewhere suggests that there will be a need for caution due to possible perverse effects and the possibility that measurement will remain largely focused on outputs and not impact.
- An increased focus on results also brings with it a risk that harder-to-measure aspects of humanitarian action, such as protection and the principles that underpin the humanitarian endeavour, could be neglected.
- There may be room for humanitarian actors to explore further the potential for learning from experience in the private sector.

5.5 The way forward: approaches to impact assessment

This study has described the different forms and functions an assessment of the impact of humanitarian action can take. The choice of the appropriate approach for assessing impact may vary according to the context, the level of analysis and the degree of accuracy sought, as well as the overall purpose of the exercise. There are significant differences between routine monitoring and one-off assessments through surveys; between quantitative and qualitative/participatory approaches; or between statements about impact in project evaluations, or lengthy

impact assessments through research and case studies. All these different forms and roles of impact assessment have their own advantages in specific contexts. The choice of method will depend on the objectives for which impact is being analysed, and the degree of validity that is expected.

This study suggests that sufficient and appropriate tools and methods exist that can provide reliable analysis of the impact of humanitarian aid. The general lack of knowledge about the impact of humanitarian programmes stems more from the under-use or inappropriate use of different methods. The various constraints characteristic of humanitarian programming should not serve as an excuse for not considering the question of impact. These constraints do, however, imply that the approach has to be adapted to the context and circumstances. It would not be helpful to expect over-burdened programme managers to rigorously analyse impact without equipping them with the capacity and resources to do so. Improving impact assessment is closely linked to the drive to improve downwards accountability and the need to make good on commitments to greater participation. However, people are rarely asked what impact they feel aid has had on their lives. Practice is beginning to emerge, for example through

DEC surveys of beneficiaries, but much more could be done to develop qualitative and participatory approaches to the analysis of impact.

Taken as a whole, the humanitarian system has been poor at measuring or analysing impact, and the introduction of results-based management systems in headquarters has yet to feed through into improved analysis of impact in the field. Yet the tools exist: the problem therefore seems to be that the system currently does not have the skills and the capacity to use them fully. This suggests that, if donors and agencies alike want to be able to demonstrate impact more robustly, there is a need for greater investment in the skills and capacities needed to do this. Given the large (and rising) expenditures on humanitarian assistance, it is arguable that there has been significant under-investment in evaluation and impact analysis. Many of the changes identified in this study would have wider benefits beyond simply the practice of impact assessment: greater emphasis on the participation of the affected population, the need for clearer objectives for humanitarian aid, more robust assessments of risk and need and more research into what works and what does not would be to the advantage of the system as a whole.

References

- ActionAid (2001) *Monitoring and Evaluating Advocacy: A Scoping Study*.
- ALNAP (2002) *Humanitarian Action: Improving Performance through Improved Learning*. London: Overseas Development Institute.
- ALNAP (2003a) *ALNAP Global Study on Beneficiary Consultation and Participation Monograph Series: Sri Lanka, Angola, Colombia, Afghanistan*. London: Overseas Development Institute.
- ALNAP (2003b) *Participation by Crisis Affected Populations in Humanitarian Action: A Handbook for Practitioners*. London: Overseas Development Institute.
- ALNAP (2003c) *Annual Review 2003. Humanitarian Action: Improving Monitoring To Enhance Accountability and Learning*. London: Overseas Development Institute.
- ALNAP (2004) *Humanitarian Protection*. London: Overseas Development Institute.
- Anderson, M. (1996) *Do No Harm: Supporting Local Capacities for Peace through Aid*. Cambridge, MA: The Collaborative for Development Action, Inc.
- Assefa, F. (2001) 'Scurvy Outbreak and Erosion of Livelihoods Masked by Low Wasting Levels in Drought-Affected Northern Afghanistan', *Field Exchange*, vol. 12, pp. 14–16.
- Bacquet, S. (2002) 'Scurvy Outbreak in Afghanistan: An Investigation by ACF and WHO', *Field Exchange*, no. 17, pp. 28–30.
- Beaton, G. and H. Ghassemmi (1982) 'Supplementary Feeding Programmes for Young Children in Developing Countries', *American Journal of Clinical Nutrition*, vol. 35, pp. 864–913.
- Borrel, A. (1997) 'Selective Feeding Programmes for Detainees in Rwanda', *Field Exchange*, vol. 2, pp. 4–6.
- Borrel, A. and P. Salama (1999) 'Public Nutrition from an Approach to a Discipline: Concern's Nutrition Case Studies in Complex Emergencies', *Disasters*, vol. 23, no. 4, pp. 326–43.
- Borton, J. (1995) *Assessing the Impact and Effectiveness of Relief Assistance*. Paper prepared for the 1995 World Disaster Report.
- Borton, J., A. Wood and R. Apthorpe (2001) *Evaluating International Humanitarian Action: Reflections from Practitioners*. London: Zed Books.
- Boss, L. P., M. Toole and R. Yip (1994) 'Assessments of Mortality, Morbidity and Nutritional Status in Somalia During the 1991–92 Famine. Recommendations for Standardization of Methods', *Journal of the American Medical Association*, vol. 272, no. 5, pp. 371–76.
- British Agencies Aid Group (2002) *Breaking New Ground? BOAG Agencies' Experience of Developing Systems for Assessing Organisational Performance*.
- CARE and Save the Children US (2002) *Rural Malawi Nutrition and Food Security Assessment*.
- CARE, WFP and ERREC (2003) *Eritrea Rural Livelihoods Security Assessment*.
- Chambers, R. (1997) *Whose Reality Counts? Putting the Last First*. London: Intermediate Technology Publications.
- Chapman, J. (2003) *System Failure: Why Governments Must Learn to Think Differently*. London: DEMOS.
- Chelimsky, E. (1995) 'The Political Environment of Evaluation and What it Means for the Development of the Field', *Evaluation Practice*, vol. 16, no. 3, pp. 215–25.
- Cliffe, L. and R. Luckham (2000) 'What Happens to the State in Conflict?: Political Analysis as a Tool for Planning Humanitarian Assistance', *Disasters*, vol. 24, no. 4.
- Collins, S. (forthcoming, 2004) *Community Based Therapeutic Care for Acute Malnutrition*, HPN Network Paper. London: Overseas Development Institute.
- Collins, S. (2001) 'Changing the Way We Address Severe Malnutrition during Famine', *The Lancet*, vol. 358.
- Collins, S. and K. Sadler (2002) 'Outpatient Care for Severely Malnourished Children in Emergency Relief Programmes: A Retrospective Cohort Study', *The Lancet*, vol. 360.
- Collinson, S. (2003) *Politically Informed Humanitarian Programming: Using a Political Economy Approach*, Network Paper 41. London: ODI.
- Commission of the European Communities (2001) *Implementing Activity Based Management in the Commission*.
- Cornwall, A. and G. Pratt (eds) (2003) *Pathways to Participation: Reflections on PRA*. London: ITDG Publications.
- Curdy, A. (1994) *The Relevance of Supplementary Feeding Programmes for Refugees, Displaced or Otherwise Affected Populations*. Paper prepared for a workshop on the improvement of the nutrition of refugees and displaced people in Africa, Machakos, Kenya, 5–7 December 1994.
- Danida (1999) *Danish Humanitarian Assistance*.
- Darcy, J. and C.-A. Hofmann (2003) *According to Need? Needs Assessment and Decision-making in the Humanitarian Sector*, HPG Report 15. London: Overseas Development Institute.
- Dart, J. and R. Davies (forthcoming) *A Dialogical Story Based Evaluation Tool: The Most Significant Change Technique*, *American Journal of Evaluation*.
- Davies, R. (2001) *A Review of NGO Approaches to the Evaluation of Advocacy Work*.
- De Waal, A. (1997) *Famine Crimes: Politics and the Disaster Relief Industry in Africa*. Oxford: James Currey.
- DEC (2000) *Independent Evaluation of Expenditure of DEC Kosovo Appeal Funds*. London: Overseas Development Institute in association with Valid International.
- DEC (2001) *Independent Evaluation of Expenditure of DEC India Earthquake Appeal Funds*. London and Ahmedabad: Humanitarian Initiatives/Disaster Mitigation Institute.
- DFID (2002) *How Effective Is DFID?*. London: DFID.
- Donnelly, A., B. Dalal-Clayton and R. Hughes (1998) *A Directory of Impact Assessment Guidelines*.
- Duce, M. et al. (2003) 'Suspected Thiamine Deficiency in Angola', *Field Exchange*, vol. 20, pp. 26–28.
- Duffield, M. (1994) 'The Political Economy of Internal War', in J. Macrae and A. Zwi (eds) *War and Hunger: Rethinking International Responses to Complex Emergencies*. London: Zed Books.
- Earle, L. (2003) *Lost in the Matrix: The Logframe and the Local Picture*. Paper presented at the Fifth INTRAC Evaluation Conference on Measurement, Management and Accountability.

- ECHO (2004) *ECHO Aid Strategy 2004*.
- Emergency Nutrition Network (2003a) 'Community Based Approaches to Managing Severe Malnutrition'. Proceedings of an inter-agency workshop, Dublin, 8–10 October 2003.
- Emergency Nutrition Network (2003b) 'Community Based Therapeutic Care: Summary of Published Research', *Field Exchange*, vol. 19.
- Estrella, M. and J. Gaventa (1998) *Who Counts Reality? Participatory Monitoring and Evaluation: A Literature Review*, IDS Working Paper 70.
- Frerks, G. and D. Hillhorst (2002) *Evaluation of Humanitarian Assistance in Emergency Situations*. Geneva: UNHCR Working Paper No 56.
- Garfield, R. (2001) 'Studies of Young Child Malnutrition in Iraq: Problems and Insights, 1990–9', *Nutrition Reviews*, vol. 58, no. 9, pp. 269–77.
- Garner, P. (1997) 'Do Objectives of Health-aid Programmes Impair Their Effectiveness?', *The Lancet*, vol. 349, pp. 722–23.
- Golden, M. (2000) 'Comments on Including Infants in Nutrition Surveys', *Field Exchange*, no. 9, pp. 16–17.
- Griekspoor, A., E. Sondorp and T. Vos (1999) 'Cost-effectiveness Analysis of Humanitarian Relief Interventions: Visceral Leishmaniasis Treatment in the Sudan', *Health Policy and Planning*, vol. 14.
- Guba, E. and Y. Lincoln (1989) *Fourth Generation Evaluation*. London: Sage Publications.
- Hailey, J. and M. Sorgenfrei (2003) *Measuring Success? Issues in Performance Management*. Paper presented at the Fourth INTRAC International Evaluation Conference.
- Hallam, A. (1998) *Evaluating Humanitarian Assistance Programmes in Complex Emergencies*. Good Practice Review 7. London: Overseas Development Institute.
- Hansch, S., S. Lillibridge, G. Egeland, C. Teller and M. Toole (1994) *Lives Lost, Lives Saved, Excess Mortality and the Impact of Health Interventions in the Somalia Emergency*.
- Hill, B. A. (1965) 'The Environment and Disease: Association or Causation?', *Proceedings of the Royal Society of Medicine*, 58.
- Hulme, D. (1997) *Impact Assessment Methodologies for Microfinance: A Review*.
- IFRC (2003) *Taking the Paper Out of Relief Work Can Save Lives*. Geneva: IFRC.
- Inter Agency Health Programme Review Initiative (2003) *Establishing Inter-agency Health Programme Evaluations in Humanitarian Crises*.
- INTRAC (1999) 'Evaluating Impact: The Search for Appropriate Methods and Instruments', *Ontrac*, no. 12.
- INTRAC (2001) *NGOs and Impact Assessment*. NGO Policy Briefing Paper No 3. Oxford: INTRAC.
- James, P. et al. (1999) 'The Potential Use of Maternal Size in Priority Setting When Combating Childhood Malnutrition', *European Journal of Clinical Nutrition*, vol. 53, pp. 112–19.
- Jaspars, S. and J. Shoham (2002) *A Critical Review of Approaches To Assessing and Monitoring Livelihoods in Situations of Chronic Conflict and Political Instability*, Working Paper 191. London: ODI.
- Joint Evaluation of Emergency Assistance to Rwanda (JEEAR) (1996) *The International Response to Conflict and Genocide: Lessons from the Rwanda Experience*.
- Kaiser, T. (2002) *Participatory and Beneficiary Based Approaches to the Evaluation of Humanitarian Programmes*, New Issues in Refugee Research, Working Paper No. 51, Geneva: Evaluation and Policy Analysis Unit, UNHCR.
- Keen, D. (1994) *The Benefits of Famine: A Political Economy of Famine and Relief in Southwestern Sudan, 1983–89*. Princeton, NJ: Princeton University Press.
- Koponen, J. and P. Mattila-Wiro (1996) *Effects or Impacts? Synthesis Study on Evaluations and Reviews 1988 to Mid 1995*. Ministry for Foreign Affairs, Department for International Development.
- Kruse, S.-E. (2003) *Searching for Impact and Methods: Meta-Evaluations of NGO Experience*.
- Lautze, S. (1997) *Saving Lives and Livelihoods: The Fundamentals of a Livelihood Strategy*. Medford, MA: Tufts University.
- Leaning, J. and S. Arie (2001) *Human Security: A Framework for Assessment in Conflict and Transition*. Working Paper Series, vol. 11, no. 8. Cambridge, MA: Harvard Center for Population and Development Studies, Harvard School of Public Health.
- Le Billon, P. (2000) *The Political Economy of War: What Relief Agencies Need to Know*. Network Paper 33. London: Overseas Development Institute.
- Lloyd Laney, M. (2003) *Advocacy Impact Assessment Guidelines*.
- Macrae, J. et al. (2002) *Uncertain Power: The Changing Role of Official Donors in Humanitarian Action*. HPG Report 15. London: Overseas Development Institute.
- Macrae, J. and N. Leader (2000) *Shifting Sands: The Search for Coherence Between Political and Humanitarian Action*. HPG Report 8. London: Overseas Development Institute.
- Maxwell, D. (1999) 'Alternative Food Security Indicators: Revisiting the Frequency and Severity of Coping Strategies', *Food Policy*, vol. 24.
- Maxwell, D. (2001) 'The Coping Strategies Index: Monitoring Food Security Status in Emergencies', *Field Exchange*, no. 13.
- Maxwell, S. (2003) *Lost in Translation? Implementation Constraints to Results-Based Management*.
- Maxwell, S. and T. Conway (2003) *Evaluating Country Programmes*. DAC Working Party on Aid and Evaluation.
- Mayne, J. (1999) *Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly*. Ottawa: Office of the Auditor General of Canada.
- Mayoux, L. (2003) *Advocacy for Poverty Eradication and Empowerment: Ways Forward for Advocacy Impact Assessment*.
- Medair (2002) *The Humanitarian Approach*.
- MSF (1995) *Nutrition Guidelines*. Paris: MSF.
- MSF (2001) 'Chronicle of Ajiep, Bahr el Ghazal, South Sudan, 1998', in *Reflections on Humanitarian Action: Principles, Ethics and Contradictions*. London: Pluto Press.
- MSF Holland (1996) *Mini-Symposium on Evaluations and Impact Studies of Humanitarian Relief Interventions*.
- Munro, L. (2002) 'Zimbabwe's Child Supplementary Feeding Programme: A Re-assessment Using Household Survey Data', *Disasters*, vol. 26, no. 3, pp. 242–61.
- National Audit Office (2002) *Department for International Development. Performance Management – Helping To Reduce World Poverty*. London: Report by the Comptroller and Auditor-General.
- Natsios, A. (2003) 'Remarks by Andrew S. Natsios,

- Administrator, USAID', InterAction Forum, Closing Plenary Session.
- NCVO (2003) *Measuring Impact. Case-studies of Impact Assessment in Small and Medium-sized Voluntary Organisations*. The UK Voluntary Sector Research Group.
- Odhiambo, O., S. Holden and C. Ackello-Ogutu (1998) *OXFAM Wajir Pastoral Development Project. Economic Impact Assessment*.
- ODI (1996) *The Impact of NGO Development Projects. Briefing Paper 2*. London: Overseas Development Institute.
- OECD/DAC (1997) *Searching for Impact and Methods: NGO Evaluation Synthesis Study*. Report prepared for the OECD/DAC Expert Group on Evaluation, Ministry for Foreign Affairs, Finland.
- OECD/DAC (1999) *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*. Paris: OECD
- OECD/DAC (2000) *Result Based Management in the Development Cooperation Agencies: A Review of Experience*. Paris: OECD.
- OECD/DAC (2002) *Glossary of Terms Used in Evaluations*. Paris: OECD
- O'Neill, O. (2002a) *Called to Account*. BBC Reith Lectures.
- O'Neill, O. (2002b) *Trust and Transparency*. BBC Reith Lectures.
- Oxfam (2003) *What Are the Lessons To Be Learnt from the Food Security Crisis in Malawi, Zambia and Mozambique and the Implications for Oxfam's Future Food Security work in Southern Africa*.
- Pretty, J. (1993) *The Trustworthiness of Findings from Participatory Methods*. London: IIED.
- Ravaillon, M. (2002) 'The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation', *World Bank Economic Review*, 15 (115).
- Refugee Studies Centre, Fritz Institute, and Norwegian Refugee Council (2003) 'Delivering the Goods: Rethinking Humanitarian Logistics', *Forced Migration Review*, vol. 18.
- Roberts, L. (2004) *Measuring the Impact of Humanitarian Health Programmes*. Background Paper. London: Overseas Development Institute.
- Roberts, L. et al. (2001) 'Keeping Clean Water Clean in a Malawi Refugee Camp: A Randomized Intervention Trial', *Bulletin of the World Health Organisation*, vol. 79, pp. 280–87.
- Roche, C. (1999) *Impact Assessment for Development Agencies: Learning to Value Change*. Oxford: Oxfam, Novib.
- Rychetnik, L., M. Frommer, P. Hawe and A. Shiell (2002) 'Criteria for Evaluating Evidence on Public Health Interventions', *Journal of Epidemial Community Health*, vol. 56.
- Salama, P. et al. (2001) 'Malnutrition, Measles, Mortality and the Humanitarian Response During a Famine in Ethiopia', *Journal of the American Medical Association*, vol. 286, no. 5, pp. 563–71.
- Save the Children UK (1999) *SCF-UK Food Aid Impact Assessment – North Wollo, South Wollo and East Hararghe, Ethiopia, November–December 1999*.
- Save the Children UK (2003) *Toolkits: A Practical Guide to Planning, Monitoring, Evaluation and Impact Assessment*. London: Save the Children.
- Save the Children UK (2004) *Global Impact Monitoring Guidelines 2004*.
- Seal, A. (1998) 'The Measurement of Micronutrient Deficiencies – A Long Way To Go', *Field Exchange*, vol. 5, no. 13.
- Shoham, J. (2004) *Review of the Role of Nutrition and Food Security Information in Assessing the Impact of Humanitarian Interventions*. Background Paper. London: Overseas Development Institute.
- Shoham, J. et al. (2001) 'DFID-funded Technical Support Facility to FAO's FIVIMS. Theme 2: The Use of Nutritional Indicators in Surveillance Systems'.
- Shoham, J. (1991) 'Attempts at Incorporating Socio-economic Data into Early Warning and Food Information Systems'. IGADD regional workshop, Nairobi, 7–9 October 1991.
- Slim, H. (1995) 'The Continuing Metamorphosis of the Humanitarian Practitioner: Some New Colours for an Endangered Chameleon', *Disasters*, vol. 19, no. 2.
- SMART (2002a) *Draft Proposal for Discussion: Monitoring Progress of Relief Assistance Nutritional Status and Crude Mortality Rate (CMR)*.
- SMART (2002b) *Improving the Monitoring, Reporting and Evaluation of Humanitarian Assistance*.
- Sphere Project (2004) *Humanitarian Charter and Minimum Standards in Disaster Response*, 2nd edition.
- Spiegel, P. et al. (2002) *Methodology Case-Study: Ethiopia*. Presentation at SMART Monitoring and Evaluation Workshop, Washington DC, 24 July.
- Spiegel, P. et al. (2001a) 'Developing Public Health Indicators in Complex Emergency Response', *Prehospital and Disaster Medicine*, vol. 16, no. 4.
- Spiegel, P. et al. (2001b) 'The Accuracy of Mortality Reporting in Displaced Persons Camps during the Post-Emergency Phase', *Disasters*, vol. 25, no. 2.
- Starling, S. (2003) *Balancing Measurement, Management and Accountability: Lessons Learned from SC UK's Impact assessment Framework*.
- Tomashek, K. M. et al. (2001) 'Randomized Intervention Study Comparing Several Regimens for the Treatment of Moderate Anemia among Refugee Children in Kigoma Region, Tanzania', *American Journal of Tropical Medicine and Hygiene*, vol. 64, nos 3–4, pp. 164–71.
- Toole, M. (1999) 'The Public-Health Consequences of Inaction', in K. Cahill (ed.) *A Framework for Survival*. London: Routledge.
- UNHCR (2001) *Real-time evaluation of UNHCR's Response to the Afghanistan Emergency*.
- UNHCR (2003) *The WHALE: Wisdom we Have Acquired from the Liberia Experience*. Report of a regional lessons-learned workshop, Monrovia, Liberia, 26–27 April 2001.
- United Nations 2003, *Consolidated Appeal for the Sudan Assistance Programme*.
- UNOCHA (2003) *Southern Africa Regional Consolidated Appeal, July 2003–June 2004*. New York and Geneva: United Nations.
- USAID (2000) *Complex Humanitarian Emergencies and USAID's Humanitarian Response*. Washington DC: Center for Development Information and Evaluation.
- Van Dyke, M. and R. Waldman (2004) *The Sphere Project Evaluation Report*. New York: Mailman School of Public Health, Columbia University.
- Wallace, T. and J. Chapman (2003) *Some Realities Behind the Rhetoric of Downward Accountability*. Working Paper presented at the Fifth INTRAC Evaluation Conference.
- Watkins, B. (2003) *Market Analysis and Emergency Needs Assessment*:

- A Review of the Issues*. Background paper for WFP technical group meeting on ENA issues, Rome 28–30 October.
- WFP (2003a) *Report of the Technical Meeting on Key Issues in Emergency Needs Assessment*. Rome, 28–30 October.
- WFP (2003b) *WFP's Evaluation Policy*.
- White, H. (2002) Combining Quantitative and Qualitative Techniques in Poverty Analysis, *World Development*, vol. 30, no. 3.
- White, H. (2001) *Combining Quantitative and Qualitative Approaches in Poverty Analysis*. Brighton: Institute of Development Studies.
- Woodruff, B. A. (2002) *Review of Survey Methodology Presented at SMART Workshop*, 23 July 23 2002.
- World Vision Sudan (1999) *Nutritional Surveys as a Management Tool*.
- Young, H. and S. Jaspars, S. (1995) *Nutrition Matters. People, Food and Famine*. London: Intermediate Technology Publications.



Britain's leading independent
think-tank on international development
and humanitarian issues

Overseas Development Institute
111 Westminster Bridge Road
London SE1 7JD
United Kingdom

Tel. +44 (0) 20 7922 0300
Fax. +44 (0) 20 7922 0399

E-mail: hpg@odi.org.uk
Websites: www.odi.org.uk/hpg
and www.odihpn.org