Evaluating household energy and health interventions

A catalogue of methods

World Health Organization

Evaluating household energy and health interventions

A catalogue of methods

World Health Organization

WHO Library Cataloguing-in-Publication Data

Evaluating household energy and health interventions : a catalogue of methods.

1.Air pollution, Indoor - prevention and control. 2.Wood fuels. 3.Energy policy.
I.World Health Organization.

ISBN 978 92 4 159691 6 (NLM classification: WA 754)

Cover photos by Nigel Bruce (top left), Nigel Bruce/Practical Action (bottom left) and Alan Gignoux/World Bank (right).
Designed by minimum graphics
Printed in France

Contents

List of tables

List of figures

List of boxes

Acknowledgements

This catalogue of methods was written by Jonathan Rouse, Independent Consultant, based on an initial outline prepared by Lisa Büttner, Winrock International and under the guidance of Eva Rehfuess, World Health Organization.

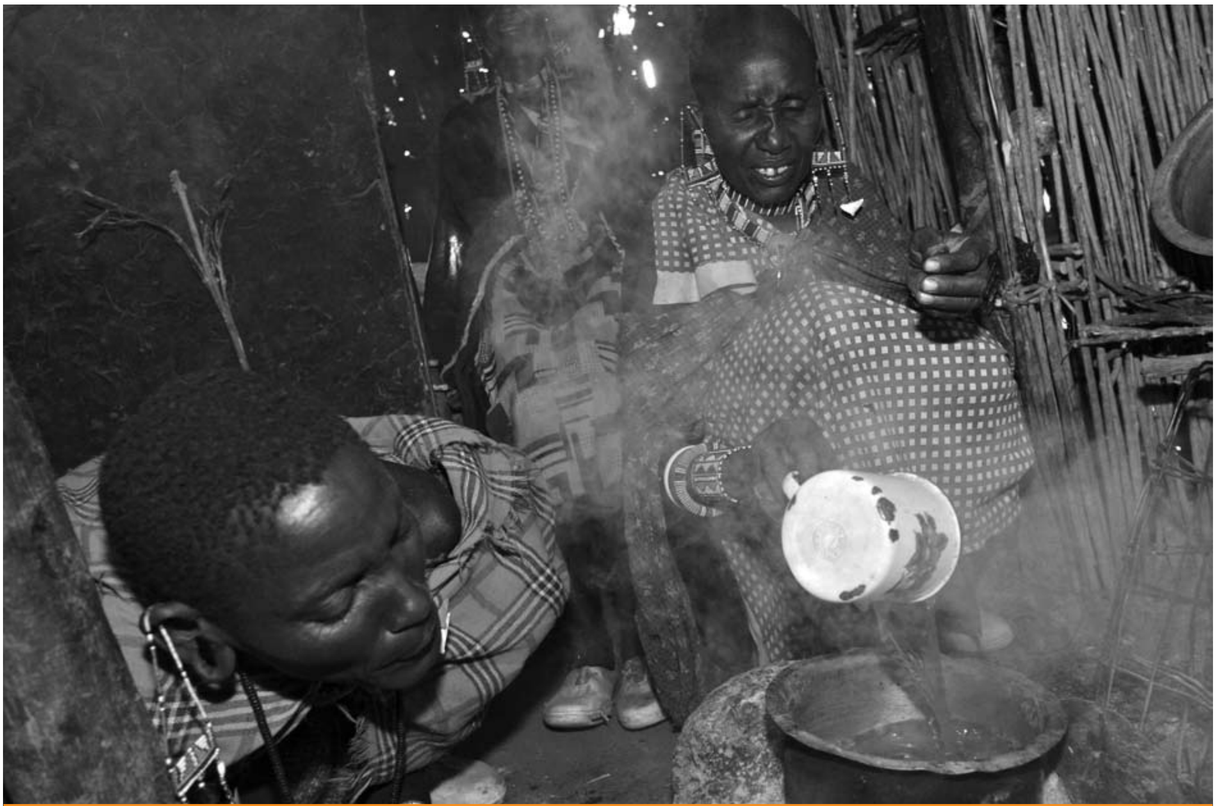
This publication is a contribution to the Partnership for Clean Indoor Air. The United States Agency for International Development (USAID) and the United States Environmental Protection Agency (USEPA) provided essential financial support.

The following individuals made important contributions of methods and suggestions and/or reviewed different drafts of the catalogue:

- Lutfiyah Ahmed, Winrock International
- Fred Arnold, ORCMacro, United States
- Jonathan Avis, Eco Securities Ltd, Oxford, England
- Grant Ballard-Tremeer, Eco Ltd, London, England
- Robert Bailis, University of California at Berkeley, United States
- Liz Bates, Practical Action, England
- Tami Bond, University of Illinois, United States
- Helen Bromley, University of Liverpool, England
- Nigel Bruce, University of Liverpool, England
- Elizabeth Cecelski, ENERGIA
- Dana Charron, CEIHD, University of California at Berkeley, United States
- Susmita Dasgupta, World Bank
- Brenda Doroski, USEPA, United States
- Karabi Dutta, Breathe Easy Network, India
- Majid Ezzati, Harvard University, United States
- Patrick Flynn, Trees, Water and People, United States
- Attila Hancioglu, UNICEF, United States
- Mike Hatfield, Aprovecho Research Center, United States
- Darby Jack, Columbia University, New York, United States
- Kirstie Jagoe, University of Liverpool, England
- Sharna Jarvis, Shell Foundation, England
- Dan Kammen, University of California at Berkeley, United States
- Marlis Kees, German Technical Cooperation, Eschborn, Germany
- Patrick Kinney, Columbia University, New York, United States
- Marcelo Korc, Pan American Health Organization, Columbia
- Jacob Moss, USEPA, United States
- John Mitchell, USEPA, United States
- Kyra Naumoff, University of California at Berkeley, United States
- David Pennise, CEIHD, University of California at Berkeley, United States
- Annette Prüss-Üstün, World Health Organization
- Roger Samson, Resource Efficient Agricultural Production, Canada
- Sumeet Saxena, The East-West Center, Honolulu, United States
- Peter Scott, Approvecho Research Centre, United States
- Kirk Smith, University of California at Berkeley, United States
- Dean Still, Aprovecho Research Center, United States
- Eleanne van Vliet, Columbia University, New York, United States
- Karen Westley, Shell Foundation, England

Acronyms

ARACHNE	Ambulatory real-time analyser for climate and health-relevant nasty emissions
ARECOP	Asia Regional Cookstove Programme
ARI	Acute respiratory infections
ALRI	Acute lower respiratory infections, e.g. pneumonia
ARTI	Appropriate Rural Technology Institute, India
AURI	Acute upper respiratory infections, e.g. common cold
CCT	Controlled cooking test
CDM	Clean Development Mechanism
CEIHD	Center for Entrepreneurship in International Health and Development
CH_4	Methane
CO_2	Carbon dioxide
CO	Carbon monoxide
COPD	Chronic obstructive pulmonary disease, e.g. chronic bronchitis
DALYs	Disability-adjusted life years, a measure of the burden of disease due to death and disability
DHS	Demographic and Health Surveys (USAID/ORCMacro)
FGD	Focus group discussion
GHG	Greenhouse gas
GTZ	Gesellschaft für Technische Zusammenarbeit (German Technical Cooperation)
GTZ/ProBEC	GTZ Programme for Biomass Energy Conservation
GTZ/HERA	GTZ Household Energy Programme
GWP	Global warming potential
HEH	Household energy and health
IAP	Indoor air pollution
IRB	Institutional review board
KPT	Kitchen performance test
LSMS	Living Standards Measurement Study (World Bank)
LPG	Liquefied petroleum gas
MPA	Methodology for participatory assessment
NMHC	Non-methane hydrocarbons
PCIA	Partnership for Clean Indoor Air
PM	Particulate matter
PM_{10}	Particulate matter of a diameter of less than 10 micrometres
$PM_{2.5}$	Particulate matter of a diameter of less than 2.5 micrometres
RCT	Randomized controlled trial
TERI	The Energy Research Institute India
TSP	Total suspended particulates
UCB	University of California at Berkeley
USEPA	United States Environmental Protection Agency
VITA	Volunteers in Technical Assistance
WBT	Water boiling test
WHO	World Health Organization
WHS	World Health Survey (WHO)



KAREN ROBINSON/PRACTICAL ACTION

1. Background

More than half the world's population relies on solid fuels, including biomass fuels (wood, dung, agricultural residues) and coal, to meet their basic energy needs. Cooking and heating with solid fuels on open fires or traditional stoves results in high levels of indoor air pollution (IAP). Globally, IAP is responsible for approximately 1.5 million deaths annually.[1] Estimates of the contribution of IAP to disease and death are also available on a country-by-country basis.[2]

A wide variety of interventions are available to reduce IAP levels, exposure and associated health effects. However, few studies have been undertaken to establish the effectiveness of these interventions at reducing pollution, improving health or resulting in social, economic or environmental benefits.

[1] World Health Organization. *Fuel for life: household energy and health*. Geneva, WHO, 2006.

[2] World Health Organization. *Indoor air pollution: national burden of disease estimates*. Geneva, WHO, 2007.

Evaluation is critical for generating the evidence needed to convince policy-makers and donors that household energy interventions can be successful in tackling one of the major threats to public health and in overcoming a major barrier to socio-economic development. Evaluation will also document experiences vital for deciding on the best intervention strategy in a given setting and for making sound policy recommendations.

A way forward in Rome

In March 2004, the Partnership for Clean Indoor Air (PCIA) in collaboration with the Italian Ministry for the Environment and Territory convened a *Harmonized Health and Exposure Assessment Protocols Workshop* in Rome. Thirty participants from around the world shared and discussed existing evaluation methods with a view to developing a consolidated evaluation resource. While participants agreed that it was not feasible to develop a harmonized protocol to suit the needs of all projects and settings, it was felt that a catalogue of methods could provide a range of evaluation options, while ensuring some comparability of the methods employed and results obtained. Based on the recommendations of the Rome workshop, the World Health Organization (WHO) prepared this catalogue of methods as a step towards:

- evaluating which interventions are effective in achieving intended impacts and how these can be implemented in a sustainable, acceptable way;
- sharing lessons learnt with implementers and other stakeholders to facilitate effective scaling up; and
- making the case with policy-makers and donors for large-scale investments in IAP reduction.

A work in progress

Household energy projects and programmes around the world use a variety of methods to evaluate the quality of their interventions, and their impacts on IAP levels, health and well-being, family livelihoods and the environment. This catalogue of methods attempts to collect methods employed and evaluation experiences gathered during the recent past, yet it is by no means comprehensive. Moreover, existing methods are constantly being refined and new tools being developed. Therefore, this catalogue of methods should be seen as a work in progress – to be updated and improved upon as new knowledge and methods become available.

Structure of the catalogue

Chapter 2 *About this catalogue* clarifies the purpose of the catalogue, and seeks to answer some of the questions readers are likely to have.

This catalogue considers seven thematic areas of evaluation, and Chapter 3 *Evaluation areas* provides an overview of each. It describes some of the challenges associated with each aspect of evaluation, and provides examples of questions to address and methods to use.

Chapter 4 *Choosing evaluation methods* is concerned with helping organizations choose appropriate evaluation methods according to their objectives, resources and the type of intervention being evaluated. Whilst promoting broad evaluation, it also highlights the resource implications of different aspects of evaluation. This chapter ends by presenting five example evaluation plans to indicate how different methods can be used together to provide a coherent focus for evaluation.

Chapter 5 *Planning and undertaking evaluation* provides some practical guidance including adapting and pilot-testing methods, study design, sample selection, fieldwork issues and data analysis.

Chapter 6 concludes by re-emphasizing the importance of evaluation and Chapter 7 lists sources of further information.

This catalogue is accompanied by a CD-ROM which contains method summaries, complete evaluation methods and further reading.

CURT CARNEMARK/WORLD BANK

2. About this catalogue

What is the purpose of this catalogue, and who is it for?

This catalogue of methods is intended to help governmental agencies, non-governmental organizations and universities involved with household energy interventions develop an evaluation strategy appropriate to their needs. It provides information on a diverse range of evaluation options ranging from simple questionnaires to complex monitoring techniques. The catalogue provides basic guidance on choosing between different evaluation options according to feasibility, organizational objectives, type of intervention and so on. It also outlines practical issues related to study design, ethical considerations, analysis and reporting. Ultimately, it is intended to save organizations time and effort in identifying evaluation methods and developing an evaluation strategy. By describing advantages and drawbacks of different approaches it aims to help organizations identify methods appropriate to their goals and organizational capacities.

This catalogue is appropriate for a wide range of organizations. For example, the methods described in this document could be employed by:

- organizations with limited resources looking to undertake simple evaluations to improve their interventions and report back to beneficiaries and donors; or
- organizations/partnerships with significant resources planning a scientific evaluation to contribute credible evidence to the international knowledge base.

Does this catalogue contain all the information required to conduct successful evaluation?

This catalogue provides resources and ideas to help structure evaluation planning. Although the catalogue of methods does not attempt to

Table 1 Health impacts of indoor air pollution[1]

Health outcome	Evidence[a]	Population	Relative risk[b]	Relative risk (95% confidence interval)[c]
Acute infections of the lower respiratory tract	Strong	Children aged 0–4 years	2.3	1.9–2.7
Chronic obstructive pulmonary disease	Strong	Women aged ≥ 30 years	3.2	2.3–4.8
	Moderate I	Men aged ≥ 30 years	1.8	1.0–3.2
Lung cancer (coal)	Strong	Women aged ≥ 30 years	1.9	1.1–3.5
	Moderate I	Men aged ≥ 30 years	1.5	1.0–2.5
Lung cancer (biomass)	Moderate II	Women aged ≥ 30 years	1.5	1.0–2.1
Asthma	Moderate II	Children aged 5–14 years	1.6	1.0–2.5
	Moderate II	Adults aged ≥ 15 years	1.2	1.0–1.5
Cataracts	Moderate II	Adults aged ≥ 15 years	1.3	1.0–1.7
Tuberculosis	Moderate II	Adults aged ≥ 15 years	1.5	1.0–2.4

[a] Strong evidence: Many studies of solid fuel use in developing countries, supported by evidence from studies of active and passive smoking, urban air pollution and biochemical or laboratory studies.
Moderate evidence: At least three studies of solid fuel use in developing countries, supported by evidence from studies on active smoking and on animals. Moderate I: strong evidence for specific age/sex groups. Moderate II: limited evidence.

[b] The relative risk indicates how many times more likely the disease is to occur in people exposed to indoor air pollution than in unexposed people.

[c] The confidence interval represents an uncertainty range. Wide intervals indicate lower precision; narrow intervals indicate greater precision.

be comprehensive, it includes a broad range of evaluation methods. Other methods developed specifically for household energy and health interventions are likely to exist, as well as generic evaluation tools. Organizations wishing to use evaluation methods not described in this catalogue may still find some of the general guidance useful.

This catalogue cannot be seen as a substitute for specialist training, or as a replacement for expert knowledge and experience. Organizations looking to contribute to the international evidence base will probably need to seek expert assistance.

What do we already know, and what are the knowledge gaps?

We know that indoor smoke contributes to childhood pneumonia as well as chronic obstructive pulmonary disease (COPD) and lung cancer (from coal) in adults, making it responsible for 1.5 million deaths per year. We also suspect that inhaling smoke may be linked to a range of other health outcomes, such as tuberculosis, low birth weight and cataracts, based on a limited number of studies in developing countries and complementary evidence on exposure to tobacco smoke and outdoor air pollution. Table 1 indicates the strength of evidence for the link between IAP exposure and health outcomes for different population groups.

We do not know the exposure-response relationship between IAP and different health outcomes, i.e. what levels of IAP cause different health outcomes. Consequently, we also do not know by how much it is necessary to reduce IAP levels in order to see benefits to health.

Several interventions can effectively reduce IAP levels (Table 2). Switching from wood, dung or charcoal to more efficient modern fuels, such as kerosene, liquefied petroleum gas (LPG) and biogas, brings about the largest reductions. A study in rural Tamil Nadu, India, compared the levels of respirable particles between homes where cooking was done using gas or kerosene and homes using wood or animal dung. Average pollution levels of 76 µg/m^3 and 101 µg/m^3 in kitchens using kerosene and gas, respectively, contrasted with levels of 1500 to 2000 µg/m^3 in kitchens where biomass fuels were used.[2]

[1] Smith KR, Mehta S, Feuz M. Indoor air pollution from household use of solid fuels. In: Ezzati M et al., eds. *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*. Geneva, WHO, 2004.

[2] Parikh J, Balakrishnan K, Laxmi V, Haimanti B. 2001. Exposures from cooking with biofuels: pollution monitoring and analysis for rural Tamil Nadu, India. *Energy* 26: 949–62.

Improved stoves – provided they are adequately designed, installed and maintained – can also cut back IAP levels considerably. Cheap wood-burning stoves in East Africa lower pollution by 50 per cent;[1] *plancha* stoves in Latin America reduce indoor smoke levels by as much as 90 per cent. Eaves spaces and extraction through smoke hoods can also curb levels of carbon monoxide and particulate matter. For example, a smoke hood installed into Maasai homes in Kenya reduced the concentration of respirable particles by up to 80 per cent, from more than 4300 $\mu g/m^3$ to about 1000 $\mu g/m^3$.[2] Changes to cooking location and ventilation characteristics, such as placement of doors and windows, were shown to have a significant impact on pollution levels in Bangladesh.[3]

Changing cooking behaviours, such as drying fuel wood before use or keeping children away from the fire, also plays a role. Such changes are unlikely to bring about reductions as great as those from switching to a cleaner fuel or the installation of a chimney stove, but they are important supporting measures for all interventions.

Yet, so far, there is little evidence that demonstrates the success of these interventions in reducing the health burden in women and children. Consequently, the key question regarding the health impacts of interventions remains:

- Which interventions reduce IAP and respiratory health outcomes, and by how much?

The first ever randomized controlled trial attempting to answer this question has recently been completed in Guatemala, evaluating the impact of reduced indoor smoke levels on childhood pneumonia and women's respiratory health. In this case the intervention, a *plancha* stove, was implemented exclusively for the sake of research. These scientific studies are complex, time-consuming and costly. They make an essential contribution to knowledge but it is not feasible to undertake such randomized controlled trials for many different interventions in many different settings.

There is thus an urgent need for the more thorough evaluation of ongoing and planned intervention projects and programmes to complement the evidence from scientific studies. Moreover, the evaluation of projects and programmes can answer important questions regarding the successful implementation of interventions in a sustainable way. In particular:

- How can interventions which meet the needs of users in the long-term and reduce IAP in real-life conditions be scaled up?
- Which interventions result in the greatest benefits at the lowest cost, providing the best value for money for limited financial resources?
- Which interventions will bring most widespread benefits to health, welfare and the environment?

Why evaluate?

Ultimately we evaluate to determine whether a given intervention has been well-received, adopted and retained by the users, and to examine whether it has been effective in achieving various objectives related to pollution levels and the health and socio-economic conditions of the target group, especially women and children.

Thorough evaluation of an intervention project/programme can:

- *Ensure that interventions address communities' needs and concerns*. Often the poor are most affected by IAP and the success or failure of interventions to address it. The indicators of success from the perspective of the users may be very different from those set by funders or implementers of the same project.
- *Help implementers and donors focus their efforts on the most effective strategies* to improve household energy and health. Implementing organizations have a duty to ensure that their interventions are safe and effective, and that they do not waste resources. Evaluation is a way for organizations to review their work and ensure that they are having a lasting

[1] Ezzati M, Mbinda MB, Kammen DM. 2000. Comparison of emissions and residential exposure from traditional and improved cookstoves in Kenya. *Environmental Science and Technology* 34(4): 578–83.

[2] Practical Action. *Reducing indoor air pollution in rural households in Kenya: working with communities to find solutions. The ITDG Smoke and Health Project, 1998–2001*. Available at: http://www.itdg.org/docs/advocacy/smoke-project-report-kenya.pdf

[3] Dasgupta S, Huq M, Khaliquzzaman M, et al. 2006. Indoor air quality for poor families: new evidence from Bangladesh. *Indoor Air* 16(6):426–44.

[4] http://www.who.int/indoorair/interventions/guatemala/

positive impact. Evaluation during the implementation of a project or programme may reveal that an intervention is not achieving its intended impacts, pointing to necessary 'midcourse corrections' and enabling improvements in the future. Moreover, funding organizations want to make sure that their money is spent well, and will base strategic decisions on evaluation outcomes.

- *Generate the evidence needed by local, national and international policy-makers and donors* that interventions to reduce IAP make a difference to people's lives and are a worthwhile investment of scarce resources. Although some aspects of household energy and health interventions are well understood and accepted, many knowledge gaps remain. Evaluation results can help to fill these and thus contribute evidence to the international knowledge base. There is particularly a need to complement evidence based on scientific research, with evidence based on interventions implemented in real-life situations.
- *Help implementers make the case for the value of their work* and attract more funding for ongoing and future activities. Careful evaluation enables organizations to provide evidence to donors when seeking further funding to continue or upscale activities. It can also help inform governments about how to allocate their limited resources.
- *Contribute to economic evaluation*. Such analyses demonstrate the economic returns on investment in the intervention and enable the comparison of cost-effectiveness or costs and benefits of different interventions. Ultimately, economic evaluation helps inform policymakers on how to allocate budgets and answers the question *Which interventions offer the greatest benefits at the lowest cost?*. For example, should a government spend a given sum of money on a small-scale improved stoves programme, or on a large-scale behaviour change campaign?

What kind of interventions and projects or programmes should we evaluate?

In principle, many interventions can reduce exposure to IAP and related health outcomes, but for the majority we have little information on how they actually impact IAP levels and people's health and livelihoods. The methods presented in this catalogue can be used to evaluate any of the interventions listed in Table 2.

This catalogue is appropriate for a wide range of projects or programmes. For example, the methods described could be applied to:

- a small-scale project promoting improved cooking stoves in a village to determine level of adoption, performance of the technology and effectiveness at reducing IAP levels;
- a medium-scale project disseminating behaviour change messages throughout a district to determine the level of adoption and perceived impact on health and welfare; or
- a large-scale programme encouraging fuel-switching across a region to determine the level of switching and the impact on health, family livelihoods and the environment.

Many NGOs across the globe are implementing small-scale household energy projects. Individually and cumulatively it is important that their effectiveness is understood, particularly prior to scaling up. The type of evaluation conducted must, however, be appropriate to the size of intervention. For example, it is not worth spending US$ 50 000.- on evaluation for a US$ 100 000 project, yet, for a US$ 1 million project this investment is certainly worthwhile.

What aspects of projects or programmes should we evaluate?

The type of intervention and the intended use and audience for the evaluation results will determine what can or should be measured. Implementers, researchers, donors and different sectors (e.g. health, energy and environment) will be interested in monitoring different aspects given their respective objectives, expertise and resources. At the same time, it is assumed that there is benefit in identifying a core set of indicators that are useful for:

- identifying which interventions are most effective; and
- making the case with policy-makers and donors about the need to reduce IAP and related health outcomes through household energy interventions.

Table 2 Interventions for reducing exposure to indoor air pollution[1]

Changing the source of pollution	Improving the living environment	Modifying user behaviour
Improved cooking devices • Improved biomass stoves without flues • Improved stoves with flues **Alternative fuel-cooker combinations** • Briquettes and pellets • Kerosene • Liquefied petroleum gas • Natural gas • Biogas, Producer gas • Solar cookers • Modern biofuels (e.g. ethanol, methanol, plant oils) • Electricity **Reduced need for the fire** • Retained heat cooker (haybox) • Efficient housing design and construction • Solar water heating • Pressure cooker	**Improved ventilation** • Smoke hoods • Eaves spaces • Windows **Kitchen design and placement of the stove** • Kitchen separate from house reduces exposure of family (less so for cook) • Stove at waist height reduces direct exposure of cook leaning over fire	**Reduced exposure by changing cooking practices** • Fuel drying • Pot lids to conserve heat • Food preparation to reduce cooking time (e.g. soaking beans) • Good maintenance of stoves and chimneys and other appliances **Reductions by avoiding smoke** • Keeping children away from smoke, e.g. in another room (if available and safe to do so)

This catalogue of methods aims to provide each audience with tools to meet a range of objectives while remaining comparable as much as possible. Although it is not possible to harmonize evaluation of different interventions around the world, it is possible for many evaluations to incorporate some key indicators. Chapter 3 considers seven thematic evaluation areas and, for each of these, describes key questions to generate comparable data.

When should we evaluate?

It is possible to include monitoring and evaluation elements in a project from the very outset, or to evaluate projects retrospectively. Currently, the most commonly used approach to monitor impacts involves a baseline survey prior to introducing the intervention, and follow-up surveys 6 and 12 months after the intervention has been implemented (see also Box 1). Chapter 5 describes different available study design options.

[1] World Health Organization. *Fuel for life: household energy and health*. Geneva, WHO, 2006.

> **BOX 1 Avoiding snapshot evaluation**
>
> Conducting evaluation at only one point in time can result in unrepresentative results. For example, if a fuel-use questionnaire is administered during dry summer months, it may not account for space heating during winter or the use of damp fuel during the rainy season.
>
> Snapshot evaluation can be avoided either by conducting monitoring at different times and seasons of the year, or by investigating and discussing the impact of seasons on the intervention. Several participatory methods, such as seasonal charting, are designed for this task.

Isn't evaluation very expensive and time-consuming?

Depending on the approach taken, evaluation certainly can be expensive and time-consuming. This catalogue describes a range of methods, including simple approaches particularly suited to smaller organizations and small-scale projects. These include aspects of routine project planning and development, such as testing technologies and undertaking initial surveys to understand needs and demands, as well as monitoring the uptake of an intervention by different population groups. Other methods, however, require specialized equipment and training. Where this is the case, it is clearly indicated. Many of the more sophisticated methods also require specialist expertise that often can only be obtained in collaboration with universities. Chapter 4 presents ideas on capacity building and accessing funds and support for evaluation.

What about evaluating unsuccessful projects or programmes?

Understanding barriers and constraints to intervention success is of critical importance. Honest reporting on interventions which have not achieved the intended impacts can enable lessons to be drawn and applied by the implementing organization as well as others facing similar situations around the world.

DAVID WHITFIELD/CEDESOL

3. Evaluation areas

This catalogue considers seven thematic areas of evaluation as particularly relevant to household energy and health interventions:

A. Adoption;
B. Market development;
C. Performance;
D. Pollution levels and personal exposure;
E. Health and safety;
F. Time and socio-economic impacts; and
G. Environmental impacts.

Evaluation also tends to distinguish between process and outcome evaluation:

- *Process evaluation* assesses what interventions have been implemented, in how many homes, with whom, when and how. It measures the extent to which an intervention has been adopted, and the factors influencing the development of a viable market that will in turn influence the sustainability, replicability and scalability of a given intervention. Process-related evaluation areas include *Adoption* and *Market development*.
- *Outcome evaluation* measures the extent to which an intervention achieves the specific outcomes desired by the beneficiaries, implementers or donors. Outcome-related evaluation areas include *Performance, Pollution levels and personal exposure, Health and safety, Time and socio-economic impacts* and *Environmental impacts*.

Presenting the methods

In this chapter, more than twenty methods are compiled in seven sections corresponding to the thematic areas of evaluation described above. For each thematic area, the purpose of evaluation is described under the heading *What does this type of evaluation tell us?*. This is followed by a set of key questions intended to provide the reader with a feel for the issues to be explored, but not designed to be used directly in surveys. Challenges in undertaking the assessment are summarized under the heading *What are the challenges?*. Each section concludes with a table that lists available methods and rates their difficulty. As described in more detail below, a summary of each method, the method itself and any additional relevant information is provided on the CD-ROM accompanying this catalogue.

Some methods consist of a number of sections which cut across several thematic areas.

- Where feasible, these have been divided into individual components relating to different thematic areas of evaluation. For example, Winrock International's 'Questions on cooking practices' (Section B) and 'Fuel use' (Section D) represent parts of the same overall evaluation tool but are listed separately. It is possible to adapt and use certain sections of methods in isolation. It is, however, important to be aware of the whole tool as linkages exist between different sections. For example, IAP monitoring relies on a post-monitoring time-activity questionnaire for context and validation.
- Where methods cannot be divided into individual components they are duplicated under all sections to which they are relevant. For example, the three stove performance tests developed by the household energy and health team at the University of California at Berkeley (C3, C4, C5) are listed under both methods to evaluate performance and methods to evaluate environmental impacts.

Identifier codes

Each of the methods included in this catalogue has been assigned an identifier code (e.g. D2). A, B, C, D, E and F refer to the seven thematic areas of evaluation described above. Generic evaluation methods that cut across several areas are marked Y. In addition, each of the different methods classified under the thematic areas of evaluation was assigned a number.

Linking to the CD-ROM

The identifier code of each method corresponds to a folder in the accompanying CD-ROM. These folders contain further details including:

- Method summary: an overview of the method, contact information, comments on use in the field and an indication of resource requirements.
- Complete method: original survey, protocol or test method.
- Any further information relating to the method.

Not all of this information is available for each of the methods. For example, where no written protocol exists, only a method summary is provided. Where possible, generic evaluation methods and other materials are included in the CD-ROM in relevant folders.

Recommended and additional methods

Where appropriate, this catalogue distinguishes between recommended methods and additional methods. Recommended methods have often been developed with a specific purpose in mind, have been employed by different organizations and are considered to be relatively well tested and standardized. In contrast, additional methods are often more general in nature.

Rating methods

Methods in this catalogue are rated as follows:

✎	Feasible for most organizations, including those with minimal evaluation experience and resources.
✎✎	Feasible for organizations with some qualitative, quantitative and analytical research experience and resources. Specific training required.
✎✎✎	Feasible for organizations with IAP monitoring devices and/or the expertise to analyse results. Specific training required.
✎✎✎✎	Feasible for organizations with the above plus access to laboratory facilities.
✎✎✎✎✎	Evaluation methods feasible only in partnership with technical specialists.

These ratings should be seen as a rough guide. Organizations will need to use the information provided in this catalogue at their own discretion to decide which methods are appropriate for their use.

A. Adoption

PETER SCOTT/APROVECHO RESEARCH CENTER

What does this type of evaluation tell us?

Evaluating adoption attempts to answer some of the most basic questions about the implementation of an intervention. It aims to determine the number of households or persons reached by a given project or programme and whether the intervention has been adopted by users as expected. It could also consider who (e.g. which socio-economic group) has adopted the intervention, and whether people are likely to continue using it.

Most interventions, including those based exclusively on the promotion of an improved cooking stove, require users to change the way they cook or use fuel. This is particularly true for interventions that attempt to change behaviour, such as drying of wood, pre-soaking lentils and beans or keeping children away from the kitchen during cooking. Adoption evaluation can be applied to all types of intervention.

One vital element of intervention evaluation is considering the extent to which the priorities of beneficiaries have been met. Adoption evaluation implicitly considers this, as it is beneficiaries that decide whether or not to adopt, and continue to use or apply, technologies and behaviours.

Key questions

- How many households have adopted the intervention? Why/why not? How many still rely on traditional practices?
- To what extent have households adopted the intervention (e.g. percentage of time/occasions when improved versus traditional practices are used, including technology, maintenance and other behaviours)?
- For how long have interventions been maintained? What barriers exist to their continued use? (This should capture maintenance issues, such as availability of spare parts, expertise, time availability and cost.)
- To what extent did the intervention reach its target audience (e.g. percentage of different socio-economic groups which adopted the intervention)? What barriers exist to a more widespread adoption?
- Have appropriate behaviours (e.g. moving children outside during cooking, using pot lids) been adopted to support new technologies/housing designs?
- What additional unplanned impacts have resulted from the intervention?

What are the challenges?

Determining whether an intervention has been fully or only partly adopted can be challenging. An example would be examining the extent to which cooks replaced polluting cooking practices (e.g. open fire, traditional stove) with improved stoves and/or complementary cooking devices (e.g. hayboxes, pressure cookers). Is it a total replacement, or are both used simultaneously, do the seasons determine which device is used, and what are the implications for the impact of the intervention?

Available methods

This aspect of evaluation is relatively easy to conduct and represents standard practice in reporting the success of a project or programme, for example when reporting back to a funding agency. Available methods to evaluate adoption are listed in Table 3.

Table 3 Evaluating adoption

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
A1	Demographic and Health Surveys	USAID/ORCMacro	Questions on fuel type, stove type and cooking location	✎
A2	World Health Survey	WHO	Questions on fuel type, stove type and cooking location, and heating practices	✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Questions on cooking practices (Section B), technology (Section C) and fuel use (Section D), enumerator observation form (Section F)	✎✎
Additional methods				
D4	Indoor air quality post-monitoring questionnaire	TERI/HEED	All	✎
Y4	Measuring successes and setbacks	GTZ/HERA	Monitoring and evaluation with users (Section 4.3)	✎
Y5	Methodology for participatory assessment	ARECOP	Section C	✎✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

B. Market development

JULIO ETCHART/WORLD BANK

What does this type of evaluation tell us?

Evaluating market development gets to the core of what makes interventions sustainable: the creation of a viable, self-sustaining market and the balance between supply and demand. Market development evaluation considers the extent of market penetration, the profitability of enterprises involved, dependencies on subsidies and the nature of the market and supply chain. Market development can be applied to all those interventions that promote the purchase of a product (e.g. improved stove, smoke hood, haybox) or fuel (e.g. LPG, ethanol).

Key questions

Customer and demand

- Who are the decision makers and customers? How can the market be characterized and segmented?
- What percentage of the market has been penetrated by the intervention?
  - What is the current size of the market for the products/services in the intervention? (i.e. all those that have adopted = y)
  - What is the potential market for the products/services in the intervention? (i.e. all those who would adopt if they were willing and able = x).
  - Percentage of market penetration = (y/x)*100
- Which credit/financing mechanisms are available for the purchase of the intervention?
- How do people's willingness and ability to pay affect the market? How has this been addressed (e.g. by ensuring satisfaction, provision of credit)?
- Which factors determining demand have been addressed by the project or programme (e.g. number of new energy kiosks where people can buy LPG)?
- What is the growth rate of adoption?
- How effective have promotional activities been?
- How has the project itself affected adoption and behaviour patterns, and what are the implications of the withdrawal of the implementing organization? For example, if spare parts are provided by the implementing organization, what happens when the organization withdraws? What structures have been put into place to ensure continuity after the project ends?

Manufacture and supply

- Who are the manufacturers, distributors, installers and other entrepreneurs in the supply chain for improved stoves/technologies (including components supply) and cleaner fuels?
- What is the profitability or rate of return for enterprises in the supply chain?
- How many manufacturers of improved stoves/technologies and/or suppliers of cleaner fuels have been set up and/or supported during the project?

- How many improved stoves/technologies have been produced, sold and disseminated?
- Are the enterprises profitable? Is this dependent on subsidies or support from the implementing organization (e.g. access to credit or markets)? How is this being addressed?
- What are the limiting factors for ensuring the supply of cleaner fuels or improved stoves/technologies, and how have they been addressed?
- Which credit/financing mechanisms are available for the production of improved stoves/technologies?
- What actions have been taken by local or national governments to facilitate growth in supply as a result of lobbying (e.g. reductions in tariffs for import of gas stove components)?

Sustainability and scale-up

- How are demand and supply balanced?
- Is the present market growth reliant on the implementing organization, and what will be the implications of withdrawal?

What are the challenges?

The tools for evaluating market development are complex. Not all organizations have skilled enterprise or market development staff, and it may be necessary to forge partnerships or source expertise for this aspect of evaluation.

Available methods

Evaluating demand overlaps with a number of other evaluation areas, namely *Adoption*, *Performance*, and *Time and socio-economic impacts*. These aspects of evaluation are all strongly linked to people's response to the interventions, i.e. whether the interventions meet their needs and expectations. Available methods to evaluate market development are listed in Table 4.

Table 4 Evaluating market development

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
B1	Breathing Space commercialization toolkit	Shell Foundation	All	✎✎
Additional methods				
C4	Kitchen performance test	Household energy and health team, UCB	Questions on user satisfaction	✎✎
Y4	Measuring successes and setbacks	GTZ/HERA	Monitoring and evaluation with producers and distributors (Section 4.2)	✎
Y5	Methodology for participatory assessment	ARECOP	Section C	✎✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

C. Performance

JONATHAN ROUSE

What does this type of evaluation tell us?

This type of evaluation is concerned with measuring impacts of interventions on combustion efficiency, wood consumption, cooking time and general performance.

It is critical that improved stoves and other devices are tested for efficiency and emissions during the project planning and development phases to ensure that they are safe, and represent an improvement over traditional practices. Further tests may be applied during the evaluation of a project to determine user satisfaction and operational performance.

Laboratory or field-based testing?

Laboratory testing is usually a critical first step in establishing the performance of a new device. Laboratory testing controls variables such as wood type, and can be based on standardized testing methods. This has obvious advantages in terms of comparing stoves with one another as part of the same project or programme as well as between different projects or programmes around the world.

Yet, the use of a given cooking technology in the field presents a range of challenges and variables that are very different from those encountered in a laboratory: efficiency, emissions and specific fuel consumption are highly dependent on stove installation, maintenance and operation. Therefore, it is also important to test stoves in ways that reflect actual usage as closely as possible, for example in people's homes.

Absolute or comparative values?

Although knowing absolute values for efficiency is of some use, comparative figures tend to be more useful and relevant for assessing the impact of an intervention and from the perspective of beneficiaries. For example, an improved stove could be correctly described as '17% efficient', or more usefully as 'twice as efficient and fast as a traditional stove' (see also Box 2).

What are important performance parameters?

There are no established performance criteria for cooking stoves, but there is relatively broad agreement on which parameters should be measured and which performance tests should be used (see also Box 3).

Efficiency

Efficiency is a measure of how much of the energy in wood is transferred into the pot and is often seen as the most effective way of determining stove performance. Efficiency is a product of combustion efficiency and heat transfer efficiency (i.e. how well the energy released from the wood is transferred to the pot):

Efficiency = combustion efficiency * heat transfer efficiency

Consequently, high stove efficiency does not necessarily mean a clean stove, as an increase in heat transfer efficiency may be achieved at the expense of combustion efficiency, and vice versa. Efficiency tests also reward the stove for the production of steam which is considered wastage during cooking.

Specific fuel consumption

Specific fuel consumption is defined as fuel used per unit of meal cooked, for example 'kilograms wood per pot of beans cooked'.

BOX 2 A simple evaluation: comparative cooking test[1]

A household energy project in India had very few resources but wanted to evaluate improved stoves in a way that was relevant and appropriate to the users. In the village of Chibau Khera the improved *Mina* stove was developed for use by mostly female domestic cooks. These were therefore chosen to be the testers, and the basis of the test was cooking a typical family meal in the village.

Public tests simulated the cooking of a typical family meal sufficient to feed six people. Two women cooked identical meals side by side, one on an improved *Mina* stove, the other on a traditional u-shaped stove. The same type of wood was provided to both women and the amount used weighed.

The purpose of the tests was described to the women. They were asked to use the stoves as they would in their own homes, and to try to use the fuel efficiently by keeping small fires for simmering and keeping burning wood well inside the firebox. No other instruction was given and the women were left to cook without any interference from the fieldworkers.

The tests revealed that the improved *Mina* stove saved 30 minutes (35%) in cooking time and used 0.5 kg (25%) less wood than the traditional u-shaped stove. In addition, the women using the *Mina* stove commented that the stove emitted considerably less smoke and that having two pot holes was more convenient.

Many women, men and children observed the tests which were followed by an announcement of the results and a meal. These public tests did much to raise the profile and popularity of the stove.

BOX 3 What about emissions testing?

Emissions testing can be used as a measure of combustion efficiency and represents a useful way of comparing different stove types or stoves in different settings. Emissions testing is, however, mostly used as a way of determining the impact of an intervention on the environment, notably through the release of greenhouse gases. Therefore, the emissions tests included in this catalogue are listed in Section G *Environmental impacts*.

The evaluation methods recommended in this catalogue are based on specific fuel use for cooking or other specified tasks, and reflect real-life conditions more accurately. Unlike efficiency testing, specific fuel consumption accounts for steam production as wastage.

Turn down ratio

This is also known as control efficiency, determined by noting the difference in fuel consumption per minute between high power (bringing water to a boil or frying) and low power (simmering). Stoves with a higher turn down ratio are likely to use less fuel during a real cooking task, which involves bringing food to a boil and then cooking it at a simmer for an extended period of time.

Key questions

- By what percentage does the intervention reduce specific fuel consumption in the laboratory, and in users' homes? How much fuel is saved on average?
- Is the improved stove or cooking device more or less efficient, convenient, time-consuming and user-friendly for specific cooking needs compared with traditional practices or technologies?
- Do improved stoves or cooking technologies continue to perform after an initial trial period (e.g. 3 months)? After 1 year? After 5 years? What are the reasons for discontinued use or performance (e.g. lack of maintenance)?
- What behavioural factors influence performance (e.g. maintenance of stove, dryness of fuel, use of pot lids, cultural beliefs)?
- How easily can cooks adjust the temperature for specific dishes? How does the turn-down capability of the improved stove or cooking technology rate?

[1] Institute of People's Action and Development Systems, Lucknow, India.

- How effective have complementary cooking devices, such as retained heat cookers, been?

What are the challenges?

Most methods recommended for testing the performance of cooking devices require some training and basic equipment. The comparative cooking test is an exception, as it does not necessitate specific resources and could be undertaken by any organization.

Available methods

The methods presented in Table 5 primarily focus on improved stoves and other cooking technologies but they can also be used to measure the effectiveness of certain behaviour changes related to cooking stoves (e.g. changes in the way fuel is used) as well as interventions to supplement traditional cooking practices, such as hayboxes.

Table 5 Evaluating performance

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
C2	VITA stove performance tests	Enterprise Works/VITA	All	✎✎
C3	Water boiling test	Household energy and health team, UCB	All	✎✎✎
C4	Kitchen performance test		All	✎✎✎
C5	Controlled cooking text		All	✎✎✎
Additional methods				
C1	Comparative cooking test	—	—	✎

A – Adoption; B – Market developm ent; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

D. Pollution levels and personal exposure

NIGEL BRUCE/PRACTICAL ACTION

What does this type of evaluation tell us?

From a health point of view, reducing exposure to IAP levels is the primary objective of household energy interventions. This section addresses understanding the impact of interventions on pollution levels and personal exposure to IAP. This is particularly important given the difficulty in assessing health outcomes directly (see Section E). Instead, reductions in pollution levels and personal exposure can be used as a proxy for likely reductions in adverse health outcomes.

The chain of events that links household energy practices to adverse health outcomes via IAP concentrations and exposures is referred to as the environmental health pathway. Figure 1 illustrates the variety of strategies to measure health impacts as a result of exposure to IAP. Moving along this pathway, assessing health impacts becomes more accurate but also more costly and difficult. Ultimately, the choice of strategy depends on an organization's technical and financial constraints as well as their objectives.

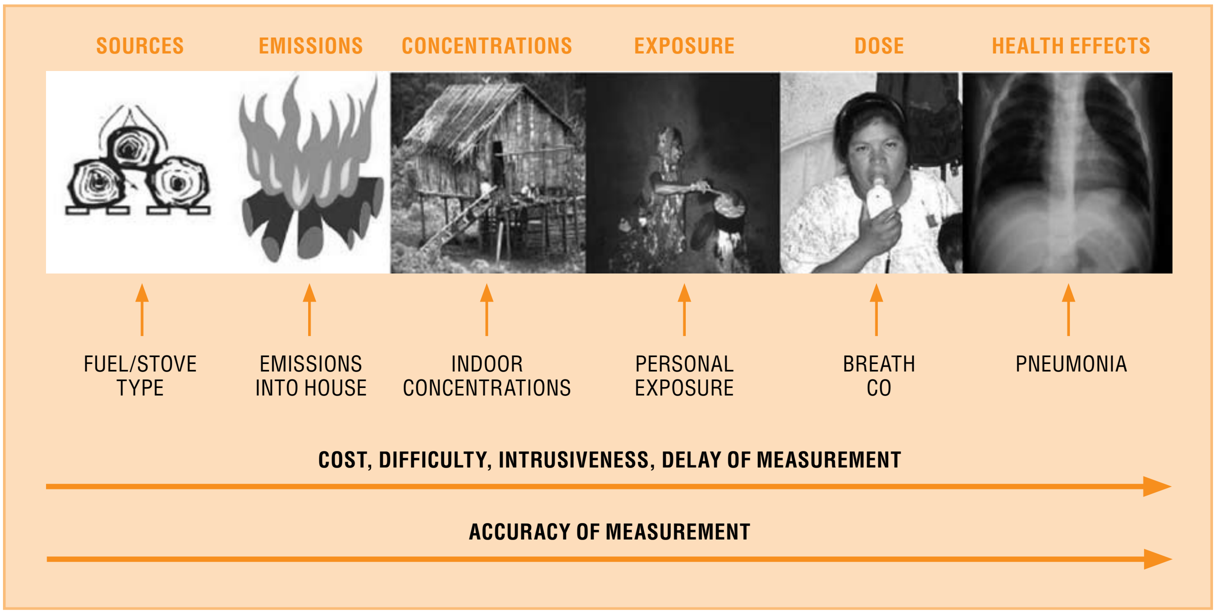
Given the difficulties of assessing health impacts directly, this section focuses on more achievable methods of estimating personal exposure to IAP:

1. The most simple and economical way to estimate exposure to IAP is through surveys.
2. In a next step, questionnaire data could be supplemented with area measurements.
3. The most sophisticated and expensive strategy is to conduct personal exposure monitoring.

Area monitoring. Measuring room pollution levels (i.e. area measurements) is a commonly used proxy for personal exposure. A monitor is placed in a standard location in a room and the concentration of a given pollutant is measured for a specific period of time. Such measurements reflect, for example, the exposure of individuals with limited mobility, such as infants, elderly or sick household members, who spend most of their time in the area being monitored. Most people, however, tend to move from high-pollution environments to low-pollution environments. In these cases, IAP monitoring can be combined with time-activity data that record participants' activities on an hour-by-hour basis on a time chart. Further accuracy might be gained from combining time activity data with pollution measurements taken on a minute-by-minute basis, to provide a better indication of levels of exposure at a given time.

Personal monitoring. The most accurate way of determining exposure is personal exposure monitoring, where participants are required to wear IAP monitors for a 24- or 48-hour period. The data account for their location and behaviour, including changes instigated by the intervention (e.g. spending more time in a less polluted kitchen), because the monitors move with the individual. Personal exposure monitoring also allows researchers to investigate exposures for specific vulnerable groups, such as women and children. Assessing the carbon monoxide level in a person's breath (CO breath) is a measure of recent exposure to IAP (within the last 5 to 6 hours) but interpreting the estimate is difficult as CO breath measures do not

Figure 1 The environmental health pathway[1]



SOURCE: KIRK SMITH, UCB

change linearly with dose. Moreover, CO breath is not suitable for assessing IAP exposure among small children as they have difficulty blowing into CO monitors.

What are we interested in measuring?

Cookstove emissions contain a wide range of harmful pollutants. With respect to measuring exposure to IAP, researchers agree that particulate matter (PM) and carbon monoxide (CO) should be monitored as they are the pollutants considered most damaging to health. In areas where a variety of fuels (notably coal but also garbage or industrial waste) is used, additional emissions may need to be monitored such as carcinogens or locally specific pollutants from coal (e.g. fluoride or arsenic in certain parts of China).

The debate on which size of particles should be measured, given currently available evidence, continues. Smaller particles with a diameter of less than 2.5 microns ($PM_{2.5}$ and PM_1) are likely to be most harmful as they penetrate deep into the human lung. Larger particles (above PM_{10}) are more likely to get filtered by the upper respiratory tract, although it is argued that insufficient research has been done to rule out their importance. It is thought that measures of $PM_{3.5}$ match the respirable fraction more closely than $PM_{2.5}$. However, the emissions of $PM_{2.5}$ and $PM_{3.5}$ from wood combustion are approximately equal, therefore measurements of $PM_{2.5}$ can be considered a close approximation of $PM_{3.5}$. In summary, international air pollution standards are based on PM_{10} and $PM_{2.5}$, and these two particle sizes continue to be the basis for much of the outdoor and IAP monitoring undertaken around the world. Considering available technologies and the relative cost and technical difficulty of monitoring, it is recommended that organizations focus on measuring levels of $PM_{2.5}$.

What is considered to be a good reduction?

The level of IAP experienced in many homes is one or two orders of magnitude greater than existing internationally accepted guideline values. For example, the WHO global air quality guidelines, which apply to both outdoor and indoor environments, recommend that the limit for annual mean PM_{10} concentrations be set at 20 µg/m³.[2] By comparison, the Chinese standard for indoor PM_{10} is set considerably higher, at 150 µg/m³. China has been one of the first countries to establish such an indoor air qual-

[1] World Health Organization. *Indoor air pollution and household energy monitoring: workshop resources*. Geneva, WHO, 2005.

[2] World Health Organization. *WHO air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide*. Global update 2005. Geneva, WHO, 2006.

ity standard (GB/T 1883-2002) although compliance for biomass stoves is not yet compulsory.

A dose-response relationship for IAP exposure and adverse health impacts is not known. Consequently, we cannot be sure by how much exposure to IAP needs to be reduced to realize benefits to health. Given the extreme starting levels of IAP, it is likely that a fairly gross reduction is required.

Key questions

- How large is the reduction in room pollution levels?
- How large is the reduction in personal exposure for mothers, children, the elderly or other specific target populations?
- What other factors, such as ventilation or personal behaviour may affect the results, and how have they been accounted for?

What are the challenges?

Assessing CO and PM levels and, even more so, personal exposure assessment is time- and labour-intensive and requires specialized equipment. In-depth training and institutional capacity is required to ensure effective monitoring and reliable results.

Choosing a time period over which to monitor will affect results. For example, monitoring undertaken during a single mealtime, over a period of 8 hours, 24 hours or a week may produce very different results and insights. It is important to appreciate that a single cooking activity may not be representative of general IAP levels, or indeed that cooking practices and fuel choices observed during one season may not reflect the situation during other seasons.

It is relatively simple to measure CO breath. It is, however, important to record when the measurement was taken in relation to smoking, cooking or other activities leading to high exposure as CO breath has a relatively short half-life of approximately five hours. Moreover, it is advisable to repeat CO breath measurements, for example, when visiting the house to place and collect equipment for room pollution monitoring.

Wearing personal monitoring equipment for an extended period of time and during all daily activities (such as cooking, sleeping, going to the market) is often considered intrusive by household members. Interacting with households prior to the study, for example through focus group discussions, is an important means of ensuring that study participants are willing to wear the monitors at all times.

Any research on human subjects, such as personal exposure monitoring, requires approval by an ethical review panel or an institutional review board to ensure that ethical issues are being considered (see Chapter 5).

Available methods

Boxes 4, 5, 6 and 7 provide a brief overview of different IAP-monitoring technologies. Available methods to evaluate pollution levels and personal exposure are listed in Table 6.

BOX 4 Carbon monoxide – colour-change diffusion tubes

Colour-change diffusion tubes, also known as stain tubes or dosimeter tubes, are small glass tubes containing a compound that changes colour on exposure to CO. The tubes give a time-weighted average for exposure to CO and are read by matching the colour change of the indicator compound with a colour key. They are small and discreet and can be easily worn on clothing although the breakable glass tube may make them unsuitable for use on children. As they do not give a precise numerical reading, they are subject to considerable inaccuracy.

Colour-change diffusion tubes are relatively inexpensive and do not require a power supply. These tubes can be purchased for around US$6 each and are not reusable. Re-usable plastic clips to allow the tubes to be worn for personal exposure monitoring cost around US$10 each.

Summary of requirements

Power supply:	✘
Calibration device:	✔ Ruler
Computer & software:	✘
Laboratory facilities:	✘

BOX 5 Carbon monoxide – real-time electro-chemical gas monitors

Real-time gas monitors can give a digital output of CO levels at one minute (or even one second) intervals. Data can be downloaded to a personal computer. This allows peaks as well as mean values to be recorded that can in turn be related to particular events or activities. Some gas monitors are small, robust, battery-operated and can be worn by adults.

Several devices are available within the price range US$250 to US$600.

Summary of requirements

Power supply:	✔	Battery (cell)
Calibration device:	✘	
Computer & software:	✔	
Laboratory facilities:	✘	

BOX 6 Particulates – gravimetric pump and filter

This method of measuring airborne particulates involves a filtration device attached to a pump. The pump draws air through filter papers over a number of hours. This method measures the mass of total suspended particulates (TSP) to indicate average pollution levels. Measuring specific particle sizes, such as $PM_{2.5}$, necessitates the use of a cyclone, which requires specific flow rates. The amount of particles collected on a filter depends on the particle concentration in the environment, on the sampling time and on the flow rate. The flow rate for measuring TSP should be standardized at a low rate to avoid overloading of the filter.

Pumps can be used for measuring adult personal exposure by placing the devices in backpacks with an intake attachment in the participant's breathing zone.

The pump and filter method is considered the standard and has been widely used. It is, however, more expensive than many of the light-scattering monitors, because there are significant overheads involved in preparing and analysing filter papers (around US$40 per data point, undertaken in a climate-controlled laboratory with a 5–6 place analytical laboratory balance).

Commonly used pump and filter devices cost approximately US$1000.

Summary of requirements

Power supply:	✔	240V/ 120V/ 12V
Calibration device:	✔	
Computer & software:	✘	
Laboratory facilities:	✔	

BOX 7 Particulates – real-time light scattering monitors

These monitors measure the scatter of light resulting from suspended particles in the air. They are able to measure changes in PM concentrations from minute to minute rather than relying on daily mean values. Laboratory facilities are not needed and the monitors themselves are quieter and easier to use. Some devices are small enough to be worn by adults. Some training and skills are required to use the monitors effectively.

A number of light scattering monitors are available on the market, priced between US$500 and $6500 per unit.

Summary of requirements

Power supply:	✔	Battery (cell)
Calibration device:	✘	
Computer & software:	✔	
Laboratory facilities:	✘	

Table 6 Evaluating pollution levels and personal exposure

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
D1	UCB light-scattering particle monitoring protocol[a]	IAP team, UCB	All	✎✎✎
D2	CO dosimeter tube protocol[a]		All	✎✎
D3	HOBO CO logger and calibration check protocols[a]		All	✎✎✎
D1, D2, D3	IAP post-monitoring questions[a]		All	✎
D4	Indoor air quality post-monitoring questionnaire	TERI/HEED	All	✎
D5	House, household and monitoring[b]	Practical Action/ University of Liverpool	PM pump and filter area monitoring (Section B)	✎✎✎✎
D5	House, household and monitoring[b]	Practical Action/ University of Liverpool	Digital CO (area) and CO (personal exposure) monitoring (Section B)	✎✎✎
D5	House, household and monitoring[b]	Practical Action/ University of Liverpool	Post-monitoring questionnaire (Section C)	✎
D6	Measuring breath CO[b]	Practical Action/ University of Liverpool	All	✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for PM pump and filter area monitoring (Section G)	✎✎✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for digital CO (area) and CO (personal exposure) monitoring (Section G)	✎✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Post-monitoring questions (Section H)	✎✎✎
Additional methods				
D8	Protocols for assessing daily integrated exposure	TERI/East-West Centre	All	✎✎✎✎✎
Y1	Indoor air pollution survey	World Bank Bangladesh	Questions on ventilation	✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

[a] These methods are accompanied by a document that describes the installation of IAP monitoring instruments in homes as well as data forms.
[b] These methods are accompanied by an interviewers' and supervisors' manual as well as data collection forms.

E. Health and safety


NIGEL BRUCE

What does this type of evaluation tell us?

Some household energy interventions are designed and implemented primarily to reduce IAP (and hence to have a major impact on respiratory disease), while for others the improvement of health is only one of many objectives. In general, the evaluation of the impacts of household energy interventions in developing countries on illness and death is considered to be of high importance.

In the assessment of any given intervention, the emphasis placed on the evaluation of various health outcomes will, among others, depend on the following critical factors:

- The objective measurement of certain health outcomes, such as childhood pneumonia, is technically complex and resource-intensive. It can therefore only be undertaken by appropriately qualified organizations and individuals.
- Several diseases associated with IAP, such as chronic obstructive pulmonary disease, cataract and tuberculosis, take many years to develop. Therefore, studying the impact of interventions on these health outcomes prospectively requires long-term investigations. A cross-sectional study design can overcome this problem but is associated with caveats (see Chapter 5).
- Intervention impacts on health depend on the nature of the intervention. For example, changes that improve ventilation are unlikely to reduce the risk of burns.
- The importance and nature of information collected about health impacts will depend on how that information is to be used. Will any insights gained be used to demonstrate health improvements at community level or are they intended for arguing the case for the intervention being a cost-effective means of reducing serious illness at national or international level?

All of these issues, and their implications for the approach taken to evaluation, are discussed further below, structured around different health and safety impacts. The following section is intended to help decide how best to focus efforts and resources depending on the goals of a given intervention, the objective for evaluation and organizational capacity.

Direct health impacts of reduced indoor air pollution exposure

As described in Chapter 2, exposure to IAP is linked with varying strengths of evidence to a range of health outcomes (Table 1). Given their major contribution to the burden of disease in developing countries, childhood pneumonia and COPD are of primary health concern and a major stimulus for the development of interventions to reduce IAP. As more evidence becomes available on other important health issues including low birth weight, tuberculosis and cataract, these are also likely to become the focus of attention for studies evaluating the impacts of reduced exposure to IAP.

It is expected that, over time, a reduction in exposure to IAP will result in a reduced risk for many, if not all, of these health outcomes.

It is therefore desirable that – where resources and technical capacity can be made available – evaluation should be employed to demonstrate such impacts. However, not all risk reductions will become apparent in a short time period and therefore, studying the effects of reduced IAP on some health outcomes is not feasible in the short-term. For example, it is expected that a reduction in IAP levels will result in a fairly rapid reduction in incidence of conditions such as childhood pneumonia or low birth weight. On the other hand, it will probably take several years to demonstrate changes in the progression of COPD, and even longer to see reductions in the risk of lung cancer, cataract, tuberculosis and other chronic conditions.

The inherent complexity involved in defining and assessing most of these disease conditions demands that a range of practical, technical and study design issues need to be carefully examined. Assessment methods involve various combinations of questionnaires, measurements by trained field staff, clinical examination by physicians and investigations, such as chest X-ray and lung function tests. In practice, many of these assessment techniques present considerable challenges, as discussed further below.

Key questions related to direct health impacts of reduced indoor air pollution exposure

- Has the intervention led to a decrease in the incidence of pneumonia among children under five years of age?
- Has the intervention led to a reduction in other health outcomes, such as low birth weight or perinatal mortality?
- Has the intervention, over a period of several years, led to a reduction in chronic respiratory symptoms and an improvement in lung function among women?

Impacts on safety during the collection and use of fuel

In many developing countries women are responsible for fuel collection and, depending on the geographical setting, accessibility of fuel and the political situation, may be at risk of injuries and violence. Children are also often involved with fuel collection, either accompanying adults or on their own. To date there has been little systematic study of the risks associated with fuel collection. Consequently, this is an important topic for evaluation work.

A number of published studies, together with anecdotal evidence, links the use of open fires or unsafe stoves to burns and scalds, in particular among young children and cooks. Less well studied, but also commonly reported, is the danger of fires (including major house fires) caused by knocking over candles or kerosene lamps.

Impacts on these aspects of safety depend on many factors, in particular the nature of the intervention. For example, it can be expected that a reduced need for fuel resulting from a more fuel-efficient stove might lead to a reduction in injury or physical assault during fuel collection.[1] A switch to a more enclosed stove can be expected to reduce the risk of burns and scalds, while a move from kerosene to LPG would reduce the risk of fires and eliminate the serious problem resulting from children drinking kerosene, which is often kept in soft drink bottles.

All of these safety impacts can be assessed relatively easily and reliably through questionnaires. Qualitative methods, such as focus group discussions, can provide more detailed information about the context of these risks. Thereby qualitative methods can help to understand how best to reduce risks, or why an intervention may not have had the desired benefit.

Key questions related to impacts on safety

- Has the intervention improved the safe use of the fire or stove (e.g. reduced risk of burns and scalds among children and women)?
- Has the intervention had an impact on health problems associated with fuel collection by women and children (e.g. backache, snake bites, physical assault)?

[1] It should be noted that this assumption does not necessarily hold true. For example, in the refugee camps in Darfur, Sudan, the introduction of more fuel-efficient stoves did not always protect women from assault, as they continued to collect wood for sale as an income-generating opportunity. This situation emphasizes the need for careful and in-depth study.

User perceptions of the health effects of indoor air pollution and intervention impacts

Users, based on their daily experience, may emphasize effects of IAP and traditional household energy practices on health that are very different from those considered to be important by health researchers and development organizations. For example, women cooks often report symptoms such as cough, chest discomfort, headache, stinging eyes and backache (in particular when switching from cooking on the ground to cooking on a raised stove) as being associated with smoke and traditional cooking practices, and describe that these have improved subsequent to the installation of an intervention. Local perceptions of health effects of IAP, as well as the perceived impacts of interventions on the health of various family members, are important both in terms of generating interest and demand, and in promoting acceptance and longer-term sustainability.

These commonly reported symptoms undoubtedly affect wellbeing and quality of life, particularly of women and young children, and hence are important in their own right. There is also some anecdotal evidence that the desire to avoid discomfort by the husband and other family members affects family relationships and interaction in the kitchen. An important issue for further study is how these commonly reported symptoms relate to the risk of serious disease outcomes.

When examining user perceptions, clinical examination of specific diseases is not required, unless in the context of a health impact study that intends to compare commonly reported symptoms with specific measures of these disease outcomes. Structured interview questionnaires and qualitative methods will both be useful in describing symptoms, and documenting and understanding what changes people ascribe to the intervention. It is very important to design questions and run focus groups in a balanced way, and to avoid creating the expectation that respondents should give answers that favour the intervention. This may happen all too easily, particularly where the implementing organization has a favourable long-standing relationship with the community, where respondents have received interventions free or subsidized, or where the perception exists that positive responses will increase the likelihood of further aid.

Key questions related to user perceptions

- What health risks do women and other household members associate with IAP and traditional cooking practices? Has the intervention brought about any change in these, for better or worse?
- Has there been any improvement, as perceived by the mother, in children's respiratory health (e.g. cough, rapid breathing, difficulty in breathing) or any other aspect of children's health and well-being (e.g. stinging eyes, headaches, burns, scalds, injuries during fuel collection)?
- Has there been any perceived improvement in women's respiratory health (e.g. wheezing, whistling, coughing) or any other aspect of women's health and well-being (e.g. eye irritation, headache, backache, physical assault during fuel collection)?

Indirect impacts on health

Household energy and health interventions can potentially bring about a range of changes to household livelihoods, to the way women, men and children use their time, and to the status of the dwelling itself. Many of these are dependent on the type of intervention as well as the local setting and culture. These impacts, as well as their assessment, are covered more fully in Section F *Time, socio-economic and other impacts*. They can, however, also have important indirect effects on the health and well-being of different family members, and this health perspective should be kept in mind. While the relationship between such changes and specific health benefits is complex and not easy to demonstrate, there are some important and well-accepted principles around how social, economic and gender-based development can impact health and wellbeing.

What are the challenges in the assessment of key health outcomes?

Respiratory health of children

One of the biggest challenges in assessing childhood acute respiratory infections (ARI) is the distinction between harmless acute upper respiratory infections (AURI, such as the common cough or cold) and potentially life-threatening acute lower respiratory infections (ALRI, such as pneumonia and bronchiolitis).

AURI are very common (5–8 disease episodes per child per year during the first year of life), while pneumonia occurs at about one-tenth of this rate (0.3–1 disease episodes per child per year during the first year of life in developing countries).

An additional complication is that a child's vulnerability to respiratory infections decreases markedly during the first year of life and, more gradually, during the first five years of life. Consequently, a child's resistance to and experience of ALRI will be very different at the beginning and at the end of a 12-month intervention study and thus it will not be possible to attribute any changes in incidence to an intervention without the use of a control group. The study design and data analysis methods will need to take these features into account.

Current experience indicates that pneumonia and other ALRI in children can only be reliably diagnosed by a physician's examination of the chest with a stethoscope. Although field staff can be trained to recognize key symptoms and signs of pneumonia (cough or difficulty breathing, rapid breathing and chest indrawing[1]), such assessments are rather non-specific and only about 30% of cases identified in this way actually have pneumonia. Most accurately, a physician's diagnosis of pneumonia is confirmed by an X-ray of the lungs. Therefore, an epidemiological study to assess the impact of an intervention on childhood pneumonia in a scientifically valid and statistically significant way requires frequent measurements, long-term follow-up and well-trained medical staff.[2]

On the other hand, if a reduction in exposure to IAP results in a reduction in childhood respiratory symptoms, such as cough, such changes can be assessed through a questionnaire. In this case mothers are asked to report on the respiratory health (cough and rapid breathing) of their children during the last two weeks or month, and the responses can be used as an indication of general respiratory health – at least as perceived by the mother. As noted above, open-ended questions and discussions may be an even more informative way to reveal a mother's perception of the impact of the intervention on child health.

Changes in health visits or hospitalization rates may offer a means of assessing child ALRI incidence, but only in areas where most of the population has regular access to health services as well as the financial means to pay for treatment. Unfortunately, this is not common in countries with high levels of IAP exposure and high incidence of ALRI. Furthermore, this link is not well understood and it could be that greater awareness of respiratory disease problems resulting from the implementation of an intervention account for the increased demand for health care and hide a potential decrease in childhood respiratory disease.

In summary, there is a great need for the research community to examine the links between household energy interventions, reduced IAP levels and changed exposure patterns, morbidity and mortality from childhood pneumonia as well as health-care seeking and hospitalization rates. Scientific studies should be conducted for major types of interventions and in different settings. Accurately evaluating a reduction in childhood pneumonia is both too complex and too costly to be undertaken as part of the evaluation of small-scale intervention projects. A more realistic and valuable approach for such projects may be to examine perceived changes in the general respiratory health of children. Consequently, the methods referred to in this section aim to document perceived changes in children's health. Such an approach is particularly relevant where health messages form part of a promotional activity: knowing more about how families perceive the health impacts of smoke and different health benefits of the intervention will be helpful in designing effective health messages.

Chronic respiratory disease in adults

COPD is a group of lung diseases that are characterized by cough, phlegm and shortness of breath (due to limited airflow and airway inflam-

[1] Lower chest wall indrawing is defined as the inward movement of the bony structure of the chest wall with inspiration. It can be considered a useful indicator of severe pneumonia if it is consistently present in a calm child (American Thoracic Society).

[2] Lanata CF, Rudan I, Boschi-Pinto C, Tomaskovic L, Cherian T, Weber M, Campbell H. 2004. Methodological and quality issues in epidemiological studies of acute lower respiratory infections in children in developing countries. *International Journal of Epidemiology* 33:1362–1372.

mation) and progressive lung tissue destruction (Box 8). It is not clear how reduced exposure to IAP will impact on COPD: with mild disease (in particular among younger women), IAP reduction may lead to gradual symptom reversal; established disease is at least partly irreversible but a reduction in IAP may prevent or slow disease progression in the same way that is seen after quitting smoking.

BOX 8 Defining chronic obstructive pulmonary disease

While several definitions exist for COPD,[1] the classification by the Global Initiative for Chronic Obstructive Lung Disease continues to be most widely used:

'COPD is a disease state characterized by airflow limitation that is not fully reversible. The airflow limitation is usually both progressive and associated with an abnormal inflammatory response of the lungs to noxious particles or gases.'[2]

As with pneumonia, COPD can only be accurately diagnosed by a physician through a combination of lung function measurement (i.e. spirometry) and a reported history of key chronic symptoms (i.e. persistent cough and production of sputum). Questionnaires alone can be used to investigate levels of, and changes (over sufficient time) in these key respiratory symptoms. Great care is required to ensure that the expressions used to describe patterns of symptoms are meaningful in a given language and relevant to local conditions such as seasons. While results from questionnaires can be strongly indicative and create persuasive evidence, they cannot be used as conclusive diagnostic tools. In addition, cultural factors may affect responses, for example in areas where tuberculosis (and associated coughing) is stigmatized, respondents may be unwilling to report the presence of a persistent cough.

Some organizations are currently using spirometry to assess women's respiratory health and to document any changes associated with the intervention. A spirometer is a device which can be used to measure lung volume and respiratory air flow. Analysis of these measures, in combination with symptoms, can be used to diagnose and measure the severity of COPD and other lung diseases. However, collecting accurate lung function data is challenging as it requires careful quality control and well-trained field staff. Field staff must be able to instruct the study participant effectively on how to carry out the breathing manoeuvre required (i.e. fast and complete exhalation). An element of learning usually occurs, which means that spirometry repeated over time will tend to improve in quality. Lung function tests also require specialist assistance in analysing and interpreting the data.

Available methods

Table 7 lists available methods to evaluate health and safety.

[1] Mannino DM. 2002. COPD: epidemiology, prevalence, morbidity and mortality, and disease heterogeneity. *Chest* 121:121S–126S.

[2] Global Initiative for Chronic Obstructive Lung Disease. 2005. *Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. Executive summary.* Available at http://www.goldcopd.org/

Table 7 Evaluating health and safety

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
A1	Demographic and Health Surveys	USAID/ORCMacro	Questions on health	✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Women's and children's health and well-being questionnaire (qualitative) (Section A.7)	✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Questions on health impact perception (Section E)	✎✎
E1	Spirometry	Practical Action/ University of Liverpool	All	✎✎✎✎✎
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

F. Time, socio-economic and other impacts

NIGEL BRUCE

What does this type of evaluation tell us?

This aspect of evaluation is concerned with understanding the short-term, visible impacts of interventions on people's lives. Many women identify time and money savings resulting from fuel efficiency as the most important outcomes of interventions. Other impacts on their status, the cleanliness of their homes, and further indirect effects are also often reported. Unlike CO emissions or stove efficiency, these impacts are directly observable. This makes them key to people's willingness to adopt, pay for and maintain improved cooking practices.

Evaluation of socio-economic impacts can also contribute to economic analysis.[1] For example, cost–benefit analysis considers all costs and benefits of interventions including impacts on health, household livelihoods and the environment. Many household energy interventions lead to important paybacks in terms of prevented illness and death, reduced expenditure on health care and fuel (where applicable) and time savings.

Socio-economic impacts include:

- **Time use.** An improved cooking stove which consumes less fuel will usually result in less time spent on collection for those who gather wood or other biomass fuels. Equally, shorter cooking times (for example when moving from a one-pot to a two-pot stove) also free up time. Saved time can have further effects: perhaps it will be spent in school or engaged in an income-generating activity.
- **Changes in expenditure.** In situations where fuel is purchased, fuel savings will result in lower expenditure on fuel. Occasionally, such as when users shift from using gathered wood to LPG, expenditure on fuel may increase – but hopefully be offset by time savings and other benefits. Purchase and maintenance costs must also be considered. Like savings in time, financial savings may have additional indirect effects, such as increased expenditure on food and therefore better nutrition.
- **Prestige and status.** A cleaner house due to less smoke or the prestige of owning a modern stove can result in a perceived rise in status of users. For example, using a modern stove could improve a woman's self-confidence, and a cleaner house may encourage visitors. These can both contribute to an individual's or family's empowerment, i.e. political, social and/or economic strength as well as confidence. Perceived prestige and status of interventions can also be used in promotional activities.
- **Other impacts – including problems.** Users will often identify benefits not foreseen by the implementing organization. Examples include improved portability of the cooking device, the ease of keeping it alight, or less blackened pots to wash. Users may also identify drawbacks of using the intervention, and it is important to understand these. The removal of smoke from homes is sometimes

[1] Details of WHO's work on economic analysis relating to household energy can be found at www.who.int/indoorair/interventions/cost_benefit/.

perceived to have negative impacts, including fear of snakes living in the smoke-free thatched roofs, reduced protection against malaria-carrying mosquitoes[1] or termites, no longer deterred by smoke, attacking the wooden structure of the house.

Key questions

- What impact has the intervention had on time allocation for women, children, men and the elderly (e.g. fuel collection, cooking, childcare)?
- What impacts has the intervention had on women's status, self esteem, decision-making power or other aspects of empowerment? What have these been due to (e.g. kitchen cleanliness, decision-making power, visitors)?
- What impacts has the intervention had on the household economy? Have women engaged in income-generating activities as a result of extra free time?
- What other impacts (both positive and negative) have been reported by women and men as a result of the intervention (e.g. effect of smoke reduction on insects, kitchen cleanliness, number and types of meals prepared)? How important are these perceived to be?

What are the challenges?

Assessing changes in time allocation and household expenditure lends itself to quantitative evaluation, and can be assessed using questionnaires or time-charts. Many of the other broader impacts are best assessed through qualitative methods, such as focus group discussions and open-ended questions. One of the most effective methods of gathering information on women's work is observation. Ideally this is undertaken in circumstances which reflect normality as much as possible, and by a known and trusted observer who has explained the purpose of the assessment. Gathering time-activity data on children can be particularly challenging but is most valuable, in particular with regard to their location. Gathering accurate information on socio-economic impacts relies on skilled researchers. This is discussed further in Chapter 5.

Available methods

Available methods to evaluate time, socio-economic and other impacts are listed in Table 8. There are also a number of basic questionnaires for understanding socio-economic impacts of interventions, including:

- IAP survey questions, World Bank, Bangladesh;
- Impact assessment questionnaire, GTZ/ProBEC, Africa; and
- Evaluation questionnaire, Trees, Water and People, United States.

The University of Liverpool is developing a standardized process for designing health and socio-economic studies to contribute to comparability between studies. Further information and example questionnaires are available at https://liv.ac.uk/hehevaluation/

[1] Biran A, Smith L, Lines J, Ensink J, Cameron M. 2007. Smoke and malaria: are interventions to reduce exposure to indoor air pollution likely to increase exposure to mosquitoes? *Transactions of the Royal Society of Tropical Medicine and Hygiene* 101:1065–1071.

Table 8 Evaluating time, socio-economic and other impacts

ID	Method	Organization	Relevant section of method	Rating
Recommended methods				
D1, D2, D3	IAP post-monitoring questions	IAP team, UCB	All	✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Post-monitoring questionnaire (Section C)	✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Post-monitoring questions (Section H)	✎
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎
Additional methods				
D7	Focus group discussion guide	Winrock International	All	✎✎
F1	Energy policies and multitopic household surveys	World Bank	Questions on fuel consumption, collection and purchase for different uses	✎
Y5	Methodology for participatory assessment	ARECOP	All	✎✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

G. Environmental impacts

NIGEL BRUCE

What does this type of evaluation tell us?

Household energy interventions can impact on the environment in two main ways:

- through emissions of air pollutants that can contribute to global warming and/or localized pollution; and
- through unsustainable use of wood as fuel, potentially contributing to deforestation and desertification, and through the use of manure as fuel, potentially contributing to loss of soil fertility.

Global environmental impact

Greenhouse gas (GHG) emissions from biomass burning account for over 70% of total emissions in Nepal, Sri Lanka and Vietnam. Although these figures are much higher than in most developing countries, there is growing interest internationally in addressing the contribution to global warming made by more than 2 billion people worldwide that rely on biomass for cooking and heating in their homes. Some desk studies have been undertaken which indicate that household energy interventions could have a significant impact on GHG emissions, and that improved cooking stoves could have their part to play.

The Clean Development Mechanism (CDM), part of the Kyoto Protocol, provides a framework for the international trading of carbon credits. This potentially opens up the possibility of funding for larger-scale household energy interventions which demonstrably result in reductions in GHG emissions.

Many different gases are thought to contribute to global warming. The methods listed in this section measure some or all of the following:

- carbon monoxide (CO);
- carbon dioxide (CO_2);
- methane (CH_4);
- other hydrocarbons; and
- suspended particulates of varying sizes (also know as aerosols).

Of these, only CO_2 and CH_4 are included within the Kyoto Protocol, but others such as CO and non-methane hydrocarbons (NMHC) are also believed to contribute to global warming. Gases can be assigned a global warming potential (GWP) value, relative to that of CO_2 (CO_2 GWP = 1). CO has a much higher GWP than CO_2, and CH_4 considerably higher still, making them more potent greenhouse gases. Aerosols affect the environment in more complex ways. Some, such as 'black carbon', have a global warming effect while others, such as organic carbon, are understood to have a cooling effect.

Emissions testing is not only carried out for assessing impacts on the global environment: it is also a useful way of comparing the combustion efficiency of different stoves, and the performance of stoves in different settings (e.g. laboratory versus users' homes). Emissions testing is not usually an effective way of assessing IAP or exposure, although it can give an indication of potential IAP levels.

Emissions can be expressed in terms of an emissions factor, namely grams emitted per unit activity. Examples include:

- g/kg wood burnt;
- g/MJ energy-delivered-to-the-pot ;
- g/meal; and
- g/cooking task.

What are the challenges in the assessment of global environmental impact?

Combustion is a complex process. Even though the methods for measuring emissions require sophisticated equipment, they can be used by non-specialized trained staff and generate data which are reasonably easy to analyse.

The source of biomass needs to be understood and considered, as this has an impact on the GHG contribution made by a stove. If wood is not harvested renewably (i.e. as much is grown as is burned) CO_2 emissions should be counted as GHG emissions. Even if biomass is harvested renewably, because combustion is never 100% efficient, it may still be necessary to consider certain products of incomplete combustion with global warming potential (e.g. organic compounds).

Local environmental impact

In settings, where ample fuel resources are readily available and can be harvested in a sustainable way, assessing local environmental impact may not be relevant. Yet, where fuel wood is scarce, demand for wood can lead to deforestation. In the context of an already unstable local environment this may add to land degradation and desertification. Many improved stoves programmes conducted in the past, e.g. the Chinese National Improved Stoves Programme, were set up to counteract such environmental pressures. A shortage of wood fuel can also force people to use manure as fuel, thus depriving fields of their natural fertilizer.

It is generally the poor who are most vulnerable to local environmental degradation such as desertification, loss of soil fertility and landslides caused by erosion. This makes understanding the impact on the local environment important for all organizations with poverty-reduction and human welfare objectives.

Local environmental impact can be assessed in two main ways:

- using questionnaires to determine fuel use, fuel sources and perceived trends and impacts on the local environment; and
- analysing data from stove performance tests in the context of local environmental conditions.

What are the challenges in the assessment of local environmental impact?

Measuring local environmental impact is generally considered to be difficult. The questionnaire methods are prone to subjectivity. The stove performance tests may provide accurate data on efficiency and wood use, but it is not easy to relate this to environmental impact. For example, a highly inefficient stove may have little impact on a robust environment, while even a highly efficient stove may have a grave impact on forest resources if used in an area of high population density with scarce forest resources.

Key questions

- By how much does the intervention lower emissions (e.g. CO_2, CO, methane, particulates) compared with traditional practices?
  - In the laboratory?
  - In users' homes?
- To what extent has the intervention led to reductions in deforestation, desertification or erosion pressures? Does the intervention involve a reforestation component and, if so, to what extent is it being implemented?
- Has the intervention impacted on soil fertility through reduced use of dung and crop waste as fuel?

Available methods

Currently, no ready-made tools are available for assessing the contribution that interventions make to improving local and global environments. Yet a number of existing methods can be adapted to inform this area of evaluation, most notably emissions measurement and fuel-use measurement and surveys.

The methods for measuring emissions can be divided into two categories: those which use a hood (mostly not portable and therefore laboratory-based) and those which use probes to measure emissions either in the room (the chamber method) or in the smoke plume. It is also possible to estimate emissions drawing on prior scientific studies and secondary data. It is recognized that the objectives of many interventions and organizations do not include environmental protection, and that this aspect of evaluation may not be widely undertaken.

Table 9 Evaluating environmental impacts

ID	Method	Organization	Relevant section of method	Rating
C2, C3, C4, C5	All technology performance tests	Enterprise Works/VITA Household energy and health team, UCB	Can be used for estimating quantities of fuel used per house/village/programme	✎✎ ✎✎✎
G1	ARACHNE emissions monitoring (without hood)	Department of Civil and Environmental Engineering, University of Illinois	All	✎✎✎✎
G2	Emissions collection hood	Aprovecho Research Center	All	✎✎✎✎
Y2	Biomass energy conservation questionnaire	GTZ/ProBEC	All	✎

A – Adoption; B – Market development; C – Performance; D – Pollution levels and personal exposure; E – Health and Safety; F – Time and socio-economic impacts; G – Environmental impacts; Y – Generic methods.

STEPHEN BATES/PRACTICAL ACTION

4. Choosing evaluation methods

This chapter is designed to help organizations choose methods appropriate to their evaluation objectives, type of intervention promoted, available resources and levels of knowledge and expertise.

The choice of evaluation methods depends on a number of factors, including:

- *Which evaluation areas (A-G) are most important in the light of the objectives for evaluation? What kind of information is required?*
  - For example, what is the objective of the intervention project or programme? What is the aim of the evaluation, and who is the target audience for results (e.g. donors, beneficiaries, implementing organization)? Are the data intended to produce statistically significant results?
- *What type of intervention project/programme is to be evaluated, and at what stage?*
  - For example, what type of intervention is the evaluation concerned with (e.g. fuel change, improved stove, behavioural change)? Is it a large-, medium- or small-scale project/programme? Has the project/programme already commenced (i.e. requiring retrospective evaluation) or is it possible to conduct a baseline and one or more follow-up assessments?
- *What capacity and resources are available to plan and undertake evaluation and analyse, interpret and record findings?*
  - For example, are staff members with evaluation experience available, and do they have enough spare capacity? Is equipment for IAP monitoring available, and is there access to laboratory facilities? Is there sufficient funding to conduct evaluation?

Factors such as evaluation objective and type of intervention are not further discussed in this catalogue. Other issues such as study design are discussed in more detail in Chapter 5. This chapter focuses on choosing methods according to the capacity and resources of a given organization.

Resource considerations

Evaluation is often time-consuming, expensive and requires skill in planning, data collection, analysis and reporting. Many organizations may find it necessary to develop their human resources further through training, to partner with organizations with more evaluation experience or to involve specialists, such as health-care workers or university researchers.

It is paramount in evaluation planning to ensure that:

- the right amount of data is collected, avoiding wasting resouces by collecting neither too much nor too little data (see Chapter 5); and
- the right type of data is collected, avoiding omitting essential data or linkages, and avoiding gathering irrelevant data.

Key questions to consider include:

- What specialist evaluation skills do staff have? Who will plan the evaluation, collect data, and analyse and interpret results?
- How much time can staff devote to evaluation? Is this sufficient for the planned scale of evaluation (i.e. sample size)?
- What capacity building, training or further resources can be secured?
- Do the necessary Information Technology resources, equipment and laboratory facilities exist, or can access to these facilities be ensured?
- Are there sufficient financial resources for the planned evaluation, including data gathering, data analysis and reporting costs?

If the necessary skills and resources of funds required to undertake evaluation are not yet available, there are several avenues for mobilizing evaluation funds and resources (see also Box 9). These include:

- securing evaluation funding through the project/programme donor and ensuring that evaluation is included in any new proposals;
- forging partnerships with local or national governmental organizations that may be interested in the results (e.g. Ministry of Health);
- collaborating with universities to gain access to research funding, equipment (e.g. IAP monitoring equipment, laboratories) and to explore the possibility of recruiting students to undertake evaluation work as part of their studies (e.g. postgraduates working on Master of Science projects);
- convincing local manufacturers and industry to invest in evaluation (e.g. market research); and
- generating an independent evaluation budget (e.g. by putting aside a small percentage of sales from stoves).

BOX 9 Support for evaluation

At present no internationally accessible service exists that is able to provide assistance and advice on evaluation. However, regional expertise and resources are being developed, such as through regional evaluation training workshops (www.who.int/indoorair/interventions/training). Several organizations facilitate cooperation on evaluation, such as the World Health Organization (www.who.int/indoorair), the Center for Entrepreneurship in International Health and Development (http://ceihd.berkeley.edu), the University of Liverpool (www.liv.ac.uk/hehevaluation) and the Partnership for Clean Indoor Air (www.pciaonline.org/).

Further considerations

Qualitative versus quantitative methods

Most evaluation methods produce either quantitative or qualitative data. Quantitative methods track changes in 'quantifiables', while qualitative methods reveal perspectives, perceptions or behaviours.

Broadly, quantitative approaches are more suited to medium- or large-scale evaluations, where information can be standardized and where respondents are likely to give accurate answers to questionnaires. They are particularly suited to more technology-focused interventions and include performance testing, IAP monitoring and questionnaires based on closed questions. Qualitative approaches include in-depth open-ended interviews with groups or individuals, observations and the use of participatory methods, such as focus group discussions. These are useful for understanding what kind of issues to consider in any evaluation. Qualitative approaches can be used alone, as well as play an important role in developing quantitative approaches, and in providing contextual data to explain the results of quantitative analysis.

It should be noted that quantitative data analysis and qualitative data analysis are very different and require distinct skills. For example, an expert in analysing emissions or exposure data may not be skilled to analyse interview transcripts on women's welfare or empowerment.

Combining evaluation areas and methods

In order to fully understand the impact of an intervention, an evaluation will consider each of the seven evaluation areas A-G (see Chapter 3). In reality, organizations have their own evaluation objectives and resource constraints. Therefore, evaluating each area in depth is usually not possible or appropriate. It is, however, important to choose and combine evaluation areas and methods carefully to deepen the knowledge of impact and to answer some of the key outstanding questions in relation to household energy and health interventions (see Box 10). For example:

- Data on IAP levels in different rooms combined with information on where different family members spend their time indicates who is exposed to different pollution levels at different times of day and for how long.
- Data on adoption rates coupled with socio-economic impacts provides insights as to why people are or are not adopting an intervention.
- Data on stove performance analysed alongside information about fuel expenditure can reveal unexpected outcomes, such as the use of traditional stoves alongside improved stoves.

BOX 10 A range of methods applied across a range of areas

In Kenya, Practical Action undertook IAP monitoring combined with surveys of health, socio-economic status and time use. The qualitative information gathered provided a useful context to the data on IAP levels. For example, informal discussions revealed that a number of households brewed beer inside their homes, resulting in particularly high IAP levels during prolonged periods of time. Data for these homes were analysed separately, making the results for all homes more accurate.

Effects, such as the brewing of beer in homes, which distort or mask a true change or association between two variables are known as confounders. Other confounding factors might be related to socio-economic status (e.g. income, wealth, education), housing type or nutritional status. Qualitative methods are invaluable in identifying and understanding these factors.

Considering how different areas relate to one other and how methods feasible to a given organization's resources and experience could complement one other, should inform the choice of evaluation methods.

Example evaluation plans

To simplify the process of choosing evaluation methods, several example evaluation plans have been prepared to account for different organizational requirements and objectives, types of intervention, budgets and so on. These evaluation plans combine different evaluation areas and identify a range of appropriate methods leaving the reader to select specific methods. They are designed to be a useful starting point for developing an evaluation strategy.

Organizations wishing to use these example evaluation plans should choose one which reasonably fits their overall objectives. The example evaluation plans will then need to be customized according to organizational requirements, evaluation objectives, type of intervention and setting by selecting, adding and adapting appropriate methods.

The following example evaluation plans are general evaluations that assess a wide range of impacts:

- Example evaluation plan 1 *Basic overview of intervention impacts* and
- Example evaluation plan 5 *Advanced assessment of intervention impacts*.

The following example evaluation plans have a thematic focus:

- Example evaluation plan 2 *Market development, time and socio-economic impacts*,
- Example evaluation plan 3 *Technology performance and indoor air pollution monitoring* and
- Example evaluation plan 4 *Indoor air pollution and personal exposure monitoring and health impacts*.

Table 10 Example evaluation plan 1

Basic overview of intervention impacts

An evaluation strategy suitable for organizations with limited evaluation experience and small-scale projects or programmes. This menu covers all evaluation areas using simple methods.

Organizational requirements

- Limited evaluation experience, time and resources.
- Basic skills in surveys, qualitative field research and data analysis.

Intervention requirements

Appropriate for smaller-scale projects and programmes and interventions of all types. If the intervention is focused around behaviour change or housing design, certain tools may need to be adapted.

Resource requirements

- Access to very simple technologies (e.g. scales for weighing wood).
- Some IT resources and analytical skills needed for data analysis and reporting.

Contribution to international evidence base

Depending on sample sizes and overall study design, some data could contribute to the international knowledge base. However, these evaluation tools will produce mainly anecdotal evidence, most useful for project planning and feedback to donors and beneficiaries.

ID	Method	Organization	Section of method	Rating
Evaluation area-specific methods				
A1	Demographic and Health Surveys	USAID/ORCMacro	Questions on fuel type, stove type and cooking location	✎
A2	World Health Survey	WHO	Questions on fuel type, stove type and cooking location, and heating practices	✎
C1	Comparative cooking test	—	—	✎
F1	Energy policies and multitopic household surveys	World Bank	Questions on fuel consumption, collection and purchase for different uses	✎
Generic methods				
Y1	Indoor air pollution survey	World Bank Bangladesh	All	✎
Y2	Biomass energy conservation questionnaire	GTZ/ProBEC	All	✎
Y3	General impact questionnaire	Trees, Water and People	All	✎
Y4	Measuring successes and setbacks	GTZ/HERA	All	✎
Additional methods and resources				
Y6	Example informed consent forms	WHO	All	✎

Table 11 Example evaluation plan 2

Market development, time and socio-economic impacts
A more complex evaluation strategy suitable for organizations with some evaluation experience and projects or programmes of all sizes.

Organizational requirements
- Some evaluation experience, and medium time and resources.
- Resources and capacity for in-depth surveys, qualitative field research and data analysis.

Intervention requirements
Appropriate for projects or programmes of all sizes and all types of interventions. If the intervention is focused around behaviour change or housing design, certain tools may need to be adapted.

Resource requirements
- Access to very simple technologies (e.g. scales for weighing wood).
- Some IT resources and analytical skills needed for data analysis and reporting.

Contribution to international evidence base
Depending on sample sizes and overall study design, some data could contribute to the international knowledge base. However, these evaluation tools will produce a large amount of qualitative data which could be difficult to interpret conclusively.

ID	Method	Organization	Section of method	Rating
Evaluation area-specific methods				
A1	Demographic and Health Surveys	USAID/ORCMacro	Questions on fuel type, stove type and cooking location	✎
A2	World Health Survey	WHO	Questions on fuel type, stove type and cooking location, and heating practices	✎
B1	Breathing Space commercialization toolkit	Shell Foundation	All	✎✎
C1	Comparative cooking test	—	—	✎
D1, D2, D3	IAP post-monitoring questions	IAP team, UCB	All	✎
D4	Indoor air quality post-monitoring questionnaire	TERI/HEED	All	✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Post-monitoring questionnaire (Section C)	✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Questions on cooking practices (Section B), technology (Section C), and fuel use (Section D) and enumerator observation form (Section F)	✎✎
D7	Focus group discussion guide	Winrock International	All	✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Post-monitoring questions (Section H)	✎
F1	Energy policies and multitopic household surveys	World Bank	Questions on fuel consumption, collection and purchase for different uses	✎
Generic methods				
Y1	Indoor air pollution survey	World Bank Bangladesh	All	✎
Y2	Biomass energy conservation questionnaire	GTZ/ProBEC	All	✎
Y3	General impact questionnaire	Trees, Water and People	All	✎
Y4	Measuring successes and setbacks	GTZ/HERA	All	✎

Table 11 Continued

ID	Method	Organization	Section of method	Rating
Additional methods and resources				
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎
Y5	Methodology for participatory assessment	ARECOP	All	✎✎
Y6	Example informed consent forms	WHO	All	✎

Table 12 Example evaluation plan 3

Technology performance and indoor air pollution monitoring

A challenging evaluation strategy suitable for organizations with considerable skills and evaluation resources and medium- to large-scale projects or programmes.

Organizational requirements

- Considerable resources, skills, capacity and time for evaluation, in particular:
  - Capacity and skills to develop an evaluation strategy and study design.
  - Skilled personnel for undertaking the evaluation, including specialists in IAP monitoring and technology performance testing.

Intervention requirements

Not appropriate for small-scale projects or programmes due to the need for sufficiently large sample sizes and substantial financial and technical resources.

Resource requirements

- Access to technology performance testing devices, including digital thermometers, scales and stopwatches.
- IT resources and skills for analysing significant volumes of data.
- Pump and filter-type devices for measuring particulate concentrations, and access to laboratories for processing filter papers.
- CO colour change diffusion tubes.

Contribution to international evidence base

If well-planned and carefully undertaken, data generated from these evaluations will contribute to the international knowledge base.

ID	Method	Organization	Section of method	Rating
Evaluation area-specific methods				
C3	Water boiling test	Household energy and health team, UCB	All	✎✎✎
C4	Kitchen performance test	Household energy and health team, UCB	All	✎✎✎
C5	Controlled cooking test	Household energy and health team, UCB	All	✎✎✎
D1, D2, D3	IAP post-monitoring questions	IAP team, UCB	All	✎
D2	CO dosimeter tube protocol	IAP team, UCB	All	✎✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	PM pump and filter area monitoring (Section B)	✎✎✎✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Post-monitoring questionnaire (Section C)	✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for PM pump and filter area monitoring (Section G)	✎✎✎✎

Table 12 Continued

ID	Method	Organization	Section of method	Rating
Evaluation area-specific methods				
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for digital CO (area) and CO (personal exposure) monitoring (Section G)	✎✎✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Post-monitoring questions (Section H)	✎
Generic methods				
Y1	Indoor air pollution survey	World Bank Bangladesh	All	✎
Y4	Measuring successes and setbacks	GTZ/HERA	All	✎
Y5	Methodology for participatory assessment	ARECOP	All	✎✎
Additional methods and resources				
C2	VITA stove performance tests	Enterprise Works/VITA	All	✎✎
D4	Indoor air quality post-monitoring questionnaire	TERI/HEED	All	✎
D8	Protocols for assessing daily integrated exposure	TERI/East-West Centre	All	✎✎✎✎✎
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎
	'How to' guide to household smoke monitoring	Practical Action	See CD ROM: D5	—
	Monitoring indoor air pollution	World Bank Bangladesh	See CD ROM: Y1	—
Y6	Example informed consent forms	WHO	All	—

Table 13 Example evaluation plan 4

Indoor air pollution and personal exposure monitoring and health impacts

A challenging evaluation strategy suitable for organizations with considerable skills and evaluation resources and medium- to large-scale projects or programmes.

Organizational requirements

- Considerable resources, skills, capacity and time for evaluation, in particular:
  - Capacity and skills to develop an evaluation strategy and study design.
  - Skilled personnel for planning and undertaking the evaluation, including specialists in IAP and health evaluation.
- Resources and capacity for in-depth surveys and qualitative field research.

Intervention requirements

Not appropriate for small-scale projects or programmes due to the need for sufficiently large sample sizes and substantial financial and technical resources.

Resource requirements

- IT resources and skills for analysing significant volumes of data.
- Light-scattering technology particulate monitors.
- CO colour change diffusion tubes.
- Real-time electrochemical CO monitors.
- CO breath monitors.

Contribution to international evidence base

If well-planned and carefully undertaken, data generated from these evaluations will contribute to the international knowledge base.

ID	Method	Organization	Section of method	Rating
Evaluation area-specific methods				
D1, D2, D3	IAP post-monitoring questions	IAP team, UCB	All	✎
D1	UCB light-scattering particle monitoring protocol	IAP team, UCB	All	✎✎✎✎
D2	CO dosimeter tube protocol	IAP team, UCB	All	✎✎
D3	HOBO CO logger and calibration check protocols	IAP team, UCB	All	✎✎✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Digital CO (area) and CO (exposure) monitoring (Section B)	✎✎✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Post-monitoring questionnaire (Section C)	✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Women's and children's health and well-being questionnaire (qualitative) (Section A.7)	✎✎
D6	Measuring breath CO	Practical Action/ University of Liverpool	All	✎✎
Generic methods				
Y4	Measuring successes and setbacks	GTZ/HERA	All	✎
Y5	Methodology for participatory assessment	ARECOP	All	✎✎
Additional methods and resources				
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Questions on health impact perception (Section E)	✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for PM pump and filter area monitoring (Section G)	✎✎✎✎

Table 13 Continued

ID	Method	Organization	Section of method	Rating
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	IAP monitoring datasheet and accompanying enumerator's manual for digital CO (area) and CO (personal exposure) monitoring (Section G)	✎✎✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Post-monitoring questions (Section H)	✎
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎
Y6	Example informed consent forms	WHO	All	—
	'How to' guide to household smoke monitoring	Practical Action	See CD ROM: D5	—
	Monitoring indoor air pollution	World Bank Bangladesh	See CD ROM: Y1	—

Table 14 Example evaluation plan 5

Advanced assessment of intervention impacts

A very advanced evaluation strategy suitable for organizations with considerable skills and evaluation resources. This menu is applicable to medium- or large-scale projects or programmes with access to health specialist knowledge.

Organizational requirements

- All of the skills required for example evaluation plans 2, 3 and 4 plus:
  - Health professional to conduct spirometry and interpret results.
- Resources and capacity for in-depth surveys and qualitative field research.

Intervention requirements

Not appropriate for small-scale projects or programmes due to the need for sufficiently large sample sizes and substantial financial and technical resources.

Resource requirements

- Access to technology performance testing devices, including digital thermometers.
- IT resources and skills for downloading and analysing significant volumes of data.
- CO colour change diffusion tubes.
- Real-time electrochemical CO monitors.
- Light-scattering technology particulate monitors.
- Spirometers.
- Emission-monitoring device.
- Access to complex and expensive (approximately US$20 000) devices required for emissions testing.

Contribution to international evidence base

If well-planned and carefully undertaken, data generated from these evaluations will contribute to the international knowledge base.

Evaluation area-specific methods

ID	Method	Organization	Section of method	Rating
B1	Breathing Space commercialization toolkit	Shell Foundation	All	✎✎
C3	Water boiling test	Household energy and health team, UCB	All	✎✎✎
C4	Kitchen performance test	Household energy and health team, UCB	All	✎✎✎
C5	Controlled cooking test	Household energy and health team, UCB	All	✎✎✎
D1	UCB light-scattering particle monitoring protocol	IAP team, UCB	All	✎✎✎

Table 14 Continued

ID	Method	Organization	Section of method	Rating
D1, D2, D3	IAP post-monitoring questions	IAP team, UCB	All	✎
D2	CO dosimeter tube protocol	IAP team, UCB	All	✎✎
D3	HOBO CO logger and calibration check protocols	IAP team, UCB	All	✎✎✎
D5	House, household and monitoring	Practical Action/ University of Liverpool	Women's and children's health and well-being questionnaire (qualitative) (Section A.7)	✎✎
D7	Household energy practices, indoor air pollution and health perceptions survey	Winrock International	Questions on cooking practices (Section B), technology (Section C) and fuel use (Section D) and enumerator observation form (Section F)	✎✎
E1	Spirometry	Practical Action/ University of Liverpool	All	✎✎✎✎✎
F1	Energy policies and multitopic household surveys	World Bank	Questions on fuel consumption, collection and purchase for different uses	✎
G1	ARACHNE emissions monitoring (without hood)	Department of Civil and Environmental Engineering, University of Illinois	All	✎✎✎✎
Generic methods				
Y4	Measuring successes and setbacks	GTZ/HERA (Section 4.3)	Monitoring and evaluation with users	✎
Additional methods and resources				
D8	Protocols for assessing daily integrated exposure	TERI/East-West Centre	All	✎✎✎✎✎
E2	Guidelines for evaluating the impacts of household energy programmes	University of Liverpool	All	✎✎
G2	Emissions collection hood	Aprovecho Research Center	All	✎✎✎✎
Y6	Example informed consent forms	WHO	All	✎

JONATHAN ROUSE

5. Planning and undertaking evaluation

Planning and undertaking evaluation involves several stages, including:

- **Choice of methods:** choosing evaluation areas and methods according to information needs, the type of intervention and organizational capacity and resources.
- **Capacity-building:** identifying and training evaluators.
- **Study design:** selecting the evaluation framework.
- **Sample selection:** deciding who and how many to evaluate.
- **Ethics:** securing consent from participants and ethical approval.
- **Adaptation of methods:** adapting tools to local language and culture, piloting and revising as appropriate.
- **Data collection:** undertaking the evaluation and collecting data.
- **Data analysis:** entering and analysing data and reporting results.

The first stage, choosing evaluation methods, has been discussed in detail in Chapter 4; this chapter is concerned with the remainder. The intention is to raise awareness of the issues associated with each of the stages rather than to cover all aspects in detail. Please note that these stages do not necessarily need to be followed sequentially. For example, the study design affects the choice of evaluation methods and vice versa.

Study design

Some of the methods described in this catalogue were developed as research tools and need to be placed into an evaluation framework (e.g.

baseline survey and follow-up after 6 and 12 months). Other methods have a built-in retrospective evaluation element: i.e. they ask how things have *changed* as a result of an intervention. Methods will need to be chosen and adapted according to the stage of an intervention project or programme and the study design.

Selecting a particular study design depends on a number of factors, including:

- outcomes of interest;
- local conditions;
- available expertise;
- financial and staff resources; and
- stage of project or programme (i.e. planning, underway, completed).

Choosing the most appropriate study design is critical and seeking specialist knowledge may be required. Three key study design options are described in this catalogue; Box 16 describes an additional study design whose complexity goes beyond the scope of evaluation.

For interventions not yet underway:

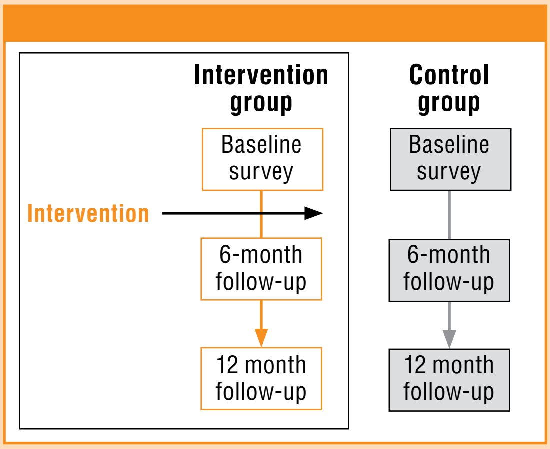
- Before-and-after design with control group (Boxes 11 and 12).
- Before-and-after design (no control group) (Box 13).

For interventions already underway or completed:

- Cross-sectional design (Boxes 14 and 15).

BOX 11 Before-and-after design with control group

This study design involves identifying a control group (ideally similar to the intervention group in every way except that this group does not receive the intervention), and undertaking baseline and follow-up surveys in both groups.



This study design is considered to be very robust, i.e. due to the presence of a control group it is less vulnerable to changes in seasons, economic or political instability and, very importantly, age-related changes in the vulnerability of the study population. It is, in fact, the only option for those wishing to accurately assess the impact of interventions on children's health.

Because of the control group, this study design requires a large sample size resulting in a lot of data to collect and process. The need to identify control homes of comparable composition and with similar socio-economic characteristics and household energy habits can be challenging. It is important not to interfere with the natural adoption process of the intervention, which may lead to some of the control group adopting the intervention during the study. Finally, there are ethical implications of using a control group which does not benefit from the impacts of an intervention. Depending on the context of the project/programme and choice of control group, this problem can be overcome by providing the control group with access to the intervention at the end of the evaluation study.

BOX 12 Case study – before-and-after design with control group

The Appropriate Rural Technology Institute undertook this evaluation study at two sites in Maharashtra, India to assess the multiple impacts of one- and two-pot improved cooking stoves.

300 homes were studied in total: 150 intervention homes cooking on an improved stove, and 150 control homes from an adjacent community cooking on a traditional open fire. Homes were chosen according to the presence of a woman aged between 15 and 45 years, and of at least one child aged less than 5 years.

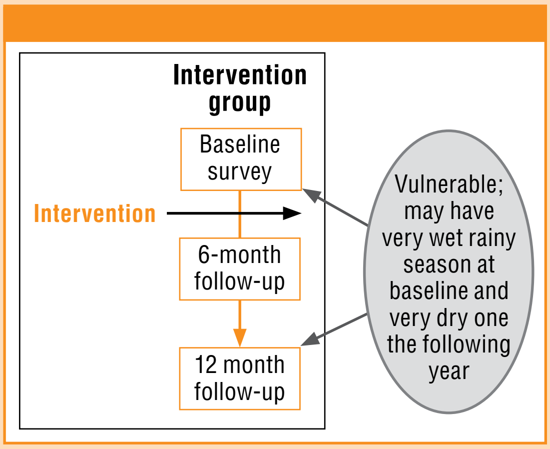
An in-depth survey was conducted to collect data at baseline as well as 6 and 12 months after the intervention was introduced. Brief questionnaires were employed after 3 and 9 months to maintain contact with communities. Moreover, interviews were held with key informants, followed by focus group discussions.

Evaluation areas included:

- pollution levels and personal exposure;
- performance;
- health and safety; and
- time, socio-economic and other impacts.

This case study was kindly provided by ARTI, India.

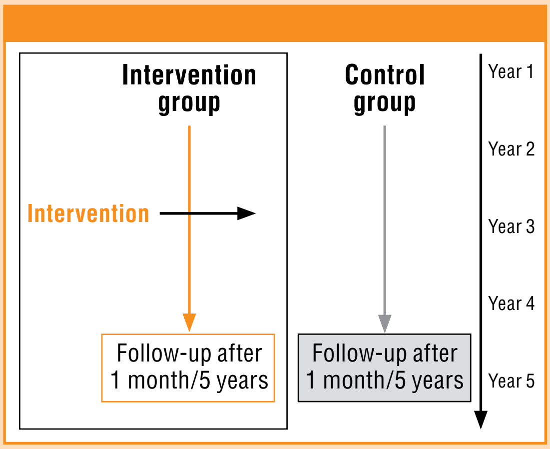
BOX 13 Before-and-after design (no control group)



This study design is similar to the one described in Box 11 with the exception of the control group. Baseline and follow-up surveys are undertaken only with the intervention group.

One of the main advantages of this design is that there is no need to identify and evaluate control homes. The smaller sample size makes the evaluation less resource-intensive and results in less data to collect, process and analyse. This design is suitable where the climatic, economic and political situation is fairly stable. It is vulnerable to any change in factors that affect energy use or health. The absence of a control group means that evaluators will not be able to distinguish changes in variables of interest due to the intervention from changes due to variability in economic, political or climatic conditions. Qualitative enquiry can help determine reasons for observed changes.

BOX 14 Cross-sectional design



This study design involves a one-off study of an intervention group and control group at some point (up to five years) after an intervention project or programme has been implemented.

A main advantage of this design is that it can be undertaken retrospectively. This makes it particularly suitable for evaluating projects which are already underway or completed. Moreover, it does not require a baseline and thus a single visit to each home is sufficient.

Some of the drawbacks of this design include the difficulty of ensuring that the differences between the intervention and control groups are due to the intervention and not other factors. Socio-economic differences, in particular, may impact both the adoption of an intervention and household energy practices. Furthermore, one-off measurements may not be representative as these often do not reflect the situation at other times of the year (see Box 1). In order to achieve statistically significant results large sample sizes are required.

BOX 15 Case study – cross-sectional design

In 2002, a multidisciplinary team of Chinese and American researchers undertook a large-scale retrospective evaluation of the Chinese National Improved Stove Programme using a cross-sectional design. Three provinces were chosen to represent different adoption rates of improved stoves and prevailing fuels:

- Zhejiang: high adoption and widespread electricity and LPG use;
- Hubei: medium adoption and widespread biomass use; and
- Shaanxi: low adoption and widespread coal use.

In China, IAP levels are determined by a complex combination of different fuels and stoves used for a variety of activities. Moreover, due to space-heating, there are substantial summer-winter differences. The researchers carefully categorized different fuels and stoves to generate meaningful data.

Evaluation areas included:

- pollution levels and personal exposure; and
- health and safety.

The study required large sample sizes:

- A household survey was conducted in nearly 3500 households, amounting to more than 20 in each village including an oversample for women and children.
- 24-hour CO and PM monitoring was undertaken in 400 households, and repeated for a sub-sample during the winter.

This case study was kindly provided by Kirk Smith, UCB.

BOX 16 Randomized controlled trial

A randomized controlled trial (RCT) is the 'gold standard' approach to measuring the impact of a given intervention on a given health outcome, such as childhood pneumonia. This research design involves the detailed investigation of a randomized selection of clusters (e.g. villages) or individuals who have received an intervention as well as a control group that has not received the intervention. RCTs, such as the trial undertaken in Guatemala (see Chapter 2), are complex and costly and go much beyond the scope of an evaluation.

Sample size and sample selection

Choosing the right size for an evaluation study is critical. The sample size is the number of homes/individuals/stoves (i.e. sample) being evaluated. If sample sizes are too large, time and money will be wasted gathering and analysing unnecessary data. If sample sizes are too small, findings may not be representative or statistically significant, and may therefore lack credibility.

Organizations intending to contribute evidence to the international knowledge base must ensure that their results are statistically significant. Even basic evaluations whose results are not intended for a wider audience need to consider sample sizes to obtain meaningful results and in the interest of careful use of resources. Small-scale evaluations can produce 'false positives' or 'false negatives', where data gathered for too small a sample size do not allow the evaluator to draw firm conclusions.

Calculating optimal sample sizes is complicated and dependent on many factors including:

- prevalence (e.g. of symptoms);
- typical levels and variation (e.g. pollution);
- expected change attributable to the intervention; and
- study design.

Organizations may choose to seek specialist help or form partnerships with, for example, universities for this aspect of evaluation planning.

Sample selection

Choosing participants, homes or devices can follow many different approaches, including:

- **Random selection.** This approach is suitable for quantitative evaluation and can be undertaken in several ways. The most straightforward approach involves assigning a unique number to households that are eligible to participate in the evaluation, and choosing a sample according to random numbers generated by a computer programme.
- **Theoretical sampling.** This approach is suitable for qualitative evaluation. Discussing issues or asking questions to individuals is repeated until no more (or very few) new responses are recorded. This point, where further study no longer contributes anything new on a particular subject, is called saturation.

Some organizations, such as Practical Action, have adopted a sampling method based on the natural adoption process to explore impacts of interventions during scaling up. This approach monitors only those people who have adopted interventions by choice. It provides evaluators with data based on what is happening in a real-life situation, and is deemed appropriate for market-based projects/programmes and behaviour change interventions.

It is vital that the right people are asked the right questions (Box 17). Practical considerations may also impact sample selection, such as location and accessibility.

BOX 17 Consulting the right people in evaluation

Sometimes those most willing and able to participate in studies are not necessarily those to whom evaluators want to listen. The poor and marginalized are often those least articulate or willing to speak.

For example, an improved stove project is implemented in rural Africa and a survey plans to investigate the socio-economic impacts of the intervention on women. Conducting the survey during daytime means that the majority of women are not at home but working in the field. Therefore, it will be difficult for fieldworkers to locate the appropriate group of women. If the fieldworkers were to ask the women who remained at home (who may not be the poorest, or those who have adopted the intervention) or other family members, they may obtain a false picture of the intervention impact.

Dropout and loss to follow-up

Sometimes participants will refuse to complete studies (dropout) or not be available during follow-up visits because they have moved away or cannot be traced (loss to follow-up). It is not possible to give exact figures on how many people will refuse to cooperate throughout an evaluation study, but many projects or programmes use sample sizes around 20% larger than statistically required so that, accounting for losses, results will still be valid (see Box 18).

BOX 18 Accounting for loss to follow-up in Guatemala

HELPS International evaluated the Onil stove in a small community of 48 families in Guatemala.

A before-and-after design was used to evaluate IAP levels and stove performance in homes using the Onil stove, and in homes cooking on a traditional open fire on a raised wooden box filled with earth.

Calculations based on an 80% reduction of IAP levels result in a required sample size of 30. To allow for 'loss to follow-up' 36 homes were monitored.

This case study was kindly provided by HELPS International, Guatemala.

Adapting evaluation methods

Most evaluation methods described in this catalogue have been developed for a specific geographical and cultural context, and most will need to be adapted to local conditions. All methods will need pilot testing prior to use (Box 19).

BOX 19 Pilot testing tools

Any newly developed or adapted evaluation tool will need to be pilot tested in the field. These are some of the key considerations in pilot testing:

- Does the flow of the questions work?
- Are the words understood? Are they too difficult, too simple, ambiguous (e.g. wheezing, stove names)?
- Do the response categories in quantitative surveys capture all options (e.g. plastic used as fuel in South African slums)?
- Are there any cultural sensitivities in relation to specific questions (e.g. asking about cough symptoms in India)?
- Are the questions interpreted in the same way by different respondents? (This is referred to as reliability.)
- Do they measure what they are supposed to measure? (This is referred to as validity.)
- Are the questions answered in the same way if repeated with the same respondent? (This is referred to as reproducibility.)

Examples of culture- and geography-specific aspects of methods include:

- local language and terminology (e.g. specific words for illnesses/symptoms, use of the Hindi word *chulha* for cooking stove);
- climatic and geographic conditions (e.g. need for space heating);
- cultural taboos (e.g. the reluctance of women to speak about coughing symptoms because of cultural stigmatization related to tuberculosis);
- cultural practices (e.g. the use of a traditional sauna or *temescal* in Guatemala); and
- locally specific cooking devices and practices (e.g. alcohol-brewing in Nepal which significantly contributes to IAP and thus impacts on the choice and effectiveness of a given intervention).

Data collection

Evaluation is undertaken for many reasons but, ultimately, it is undertaken to ensure that people (i.e. beneficiaries) have been well served, and to inform decisions about how to serve people better in the future. Data collection is at the heart of evaluation: both interviewer and interviewee play a key role. This section addresses some of the issues related to choosing the right evaluator and designing and asking questions in the best possible way.

Choice of evaluator

The person that facilitates discussions or administers surveys can have a profound impact on data collection. Choosing the most appropriate evaluator depends on what information is being collected and from whom. For example, in some cultures it would be considered inappropriate for a male researcher to speak with female cooks. In other situations, respondents may feel intimidated or under pressure to respond positively if the project manager of the intervention project/programme is asking the questions.

It is important that evaluators are well-trained and that they and, where applicable, their translators are aware of the aims of the research. It is also vital that they appreciate that refusals to respond or negative responses are valuable out-

BOX 20 People-centred evaluation

Evaluation implicitly involves interaction with people: gaining access to their homes to monitor IAP levels or test the stove, or asking (sometimes personal) questions relating to their health, time use and socio-economic status. Evaluation is by nature intrusive, and it is important for evaluators to be sensitive to the expectations and specific needs of participants.

One of the key principles of participatory approaches is reversing roles. For evaluation this means that the development professionals become the learners and listeners, and the participants become the teachers and informers. Users of stoves are the experts, and evaluators need to recognize their knowledge on how the intervention has worked.

comes, and will not be viewed as failure on their part by the evaluation coordinators.

Designing and asking questions

In designing qualitative and quantitative materials it is important to minimize suggestion in questions, and to phrase questions in an open-ended way. For example:

'Can you tell me about any difference in how much wood the two stoves use?'	✔ *open question*
'Is there any difference in how much wood the two stoves use?'	✘ *closed question*
'The new stove uses more wood, doesn't it?'	✘ *leading question*

People tend to respond more honestly to open questions than to closed or leading questions. It is also more difficult to answer an open-ended question if it has not been properly understood, whereas answering a closed question only requires a simple yes or no.

To aid data analysis open questions can include coded responses, for example:

'Which stove do you use most of the time?'

A. Single pot ❒

B. Double pot ❒

C. Three stone fire ❒

D. Other ....

In the interest of eliciting accurate and meaningful responses, participants should be made aware that:

- the purpose of the study or survey is to enable improvements to be made to the work;
- interviewers are equally interested in whether the situation is worse, the same or better; and
- their answer will not disadvantage them in terms of future assistance.

Be aware of language issues – particularly when using translators – as certain words (e.g. wheeze) may be difficult to translate accurately (see Box 19). If necessary, demonstrate words and concepts to make sure people have understood.

Feeding back to participants

Evaluation information voluntarily provided by participants should be made available to them upon request (see Box 20). Many organizations choose to share evaluation findings, partly as a show of appreciation for participation but also as a promotional tool. For example, results showing that an improved cooking stove adopted by some families in the community has resulted in increased disposable income, lowered IAP levels and reduced coughing in children could convince many more families to purchase one.

Ethical considerations

Wherever research involves human subjects, ethical issues need to be considered carefully (Box 21). Organizations intending to publish evaluation results need to take particular care, and those monitoring IAP levels, personal exposure and health outcomes may need to seek approval from an ethical review panel or institutional review board.

A range of example informed consent forms developed by WHO have been included in the accompanying CD-ROM. Further examples and information can be downloaded from www.who.int/rpc/research_ethics/informed_consent/en/

BOX 21 Three basic principles of ethics in research

Beneficence: 'the duty to do good'

- Research should cause no harm to participants, whether intentionally or by failing to anticipate and avoid harm.
- The research design should maximize benefits and minimize harm.

Respect for persons

Research should uphold the following principles:

- Autonomy or self-determination.
- Voluntariness, including the choice to opt out of activities at a later stage.
- Duty to protect persons with limited autonomy (e.g. children, refugees, women).
- Confidentiality, anonymity and privacy:
  - — Numeric codes (instead of address/name) should be used on all forms/databases.
  - — All records (e.g. health status, socio-economic status) should be stored in locked rooms with only study staff having access.

Justice

- Research should not create injustices, whether in relation to risks and discomforts or in relation to benefits.

Data management and data analysis

Data management

It is important to be systematic about data storage and management. Many of the methods included in this catalogue, for example the methods for testing stove performance and assessing IAP levels, are accompanied by data collection forms. These serve as a template for data entry and storage. Some survey questionnaires, for example those related to impacts on symptoms, time and socio-economic status, are lengthy and require many data entries for each participant. Similarly, qualitative methods, such as focus group discussions, can generate large volumes of data which require careful and skilled management.

Data analysis

Analysing data can be simple or complicated, depending on the methods employed and the type of outcome to be reported.

Producing descriptive statistics is usually relatively straightforward. It requires familiarity with general applications in software packages, such as Microsoft Excel or other spreadsheet programmes, and a basic understanding of mathematics. The outcomes produced tend to be numbers or percentages, such as the proportion of households using different types of stoves or fuels based on an adoption survey, the proportion of women reporting cough symptoms or the weekly amount of different fuels used for cooking.

Computer software is essential for downloading, analysing and storing data from digital IAP monitors, such as the HOBO CO monitor or the UCB particle monitor. Much or all of the required data processing, management and analysis can be performed in Excel or another spreadsheet programme.

Establishing relationships between the intervention and changes in an outcome of interest (e.g. pollution levels, health outcomes) is more difficult. It requires at least a basic understanding of statistics and epidemiology and experience with software packages, such as Microsoft Excel, EpiInfo, Stata, SAS or SPSS. An important question in evaluation research is how to distinguish between changes brought about by the intervention and changes due to other factors or chance.

An awareness of concepts, such as statistical significance and confounding, is therefore critical. Statistical significance allows differences that are meaningful to be distinguished from differences that are not meaningful and brought about by chance or small sample sizes. Adjusting for confounders (i.e. influences on the outcome of interest other than the intervention) is important as these may confuse, distort or mask true associations.

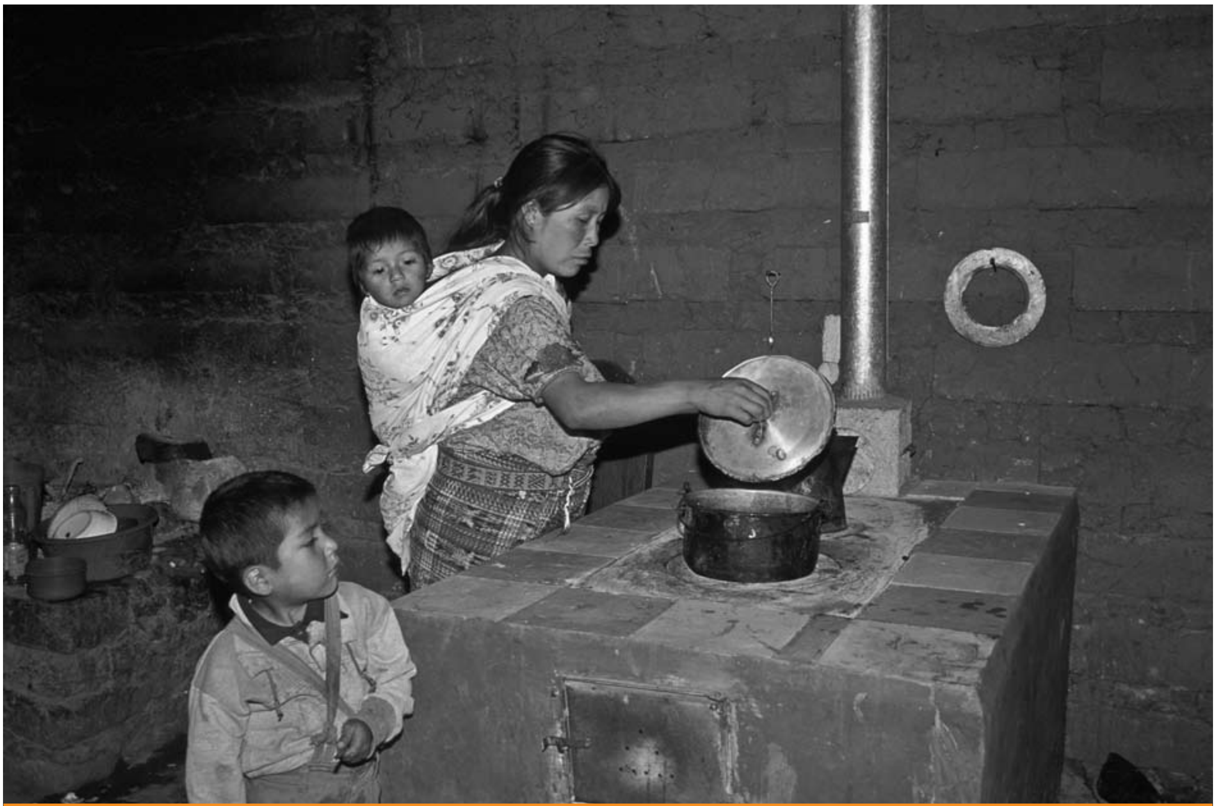
Some organizations may not have the skills to analyse the data they are well-equipped to collect. In this case, it is necessary either to develop the skills within the organization through specialist training (e.g. courses in statistics) or partner with organizations that already have the required knowledge (e.g. universities).

Reporting evaluation results

Communicating the results of an evaluation is critical. Reporting difficulties may help others avoid repeating the same mistakes. Sharing successes may enable others to replicate these in different settings. Important communication channels include organization-specific reports, articles published through organizations engaged with household energy and health monitoring (some of which are listed in Chapter 7) and peer-reviewed articles in the scientific literature.

As with any report, the description of an evaluation should be as detailed as necessary while as concise as possible. Even technical experts enjoy reading an interesting report, and findings should therefore be communicated in a clear and simple way. Usually it is not necessary to report every finding: key messages should be selected based on the target audience for the report.

Many evaluations will combine qualitative and quantitative methods including IAP monitoring, stove tests, questionnaires, discussions and observations. Data presentation should reflect this diversity and include the use of text, graphs, tables, quotations, photographs and even sketches.



NIGEL BRUCE

6. Looking ahead

In recent years, access to modern cooking energy has attracted growing attention in national and international development agendas. One of the reasons for this trend is a better understanding of the serious health consequences of indoor air pollution from solid fuel use. As the profile of household energy rises, so does the pressure on implementers to adequately measure and report the various impacts of their work. To date, however, many of the impacts of interventions designed to reduce indoor air pollution remain poorly tested.

This catalogue of methods is intended to help organizations examine whether a given intervention has been well-received, adopted and retained by the users, and to assess the impacts on indoor air pollution, health, socio-economic conditions and the environment. Monitoring and evaluation can help ensure lessons are learned within and across projects and programmes. Most importantly, rigorous quantitative measures of impact can be used to build the evidence required to make the case for household energy, indoor air pollution and health with governments and donors.

We hope that you feel inspired and better-equipped to undertake monitoring and evaluation of your work.

7. Further reading

General resources

GTZ Household Energy Programme (HERA). *Measuring successes and setbacks: how to monitor and evaluate household energy projects*. Eschborn, GTZ/HERA,1995. (included in CD-ROM)

International Nutrition Foundation for Developing Countries. *A manual for the use of focus groups. Methods for social research in disease*. Boston MA, INFDC, 1993. Available at: www.unu.edu/unupress/food2/UIN03E/UIN03E00.HTM

Practical Action. *'How to' guide to household smoke monitoring*. Available at: www.hedon.info/goto.php/view/407/forum.htm

Practical Action. *Smoke, health and household energy. Volume 1. Participatory methods for design, installation, monitoring and assessment of smoke alleviation technologies*. Bourton-on-Dunsmore, Practical Action, 2005. Available at: http://practicalaction.org/?id=smoke_health_household_energy

World Health Organization. *Indoor air pollution and household energy monitoring: workshop resources*. Geneva, WHO, 2005. Available at: www.who.int/indoorair/publications/workshopresources/ (included in CD-ROM)

Specific resources

Avis J. *Cooking up carbon credits: can carbon trading aid the successful dissemination of improved cooking stoves?* Oxford, Environmental Change Institute, 2004.

Baldwin SF. *Biomass stoves: engineering design, development and dissemination*. Princeton University, Enterprise Works/Volunteers in Technical Assistance, 1987.

Bond T, Venkataraman C, Masera O. 2004. Global atmospheric impacts of residential fuels. *Energy for Sustainable Development* 8(3):54–66.

Smith KR, Uma R, Kishore VVN, Zhang J, Joshi V, Khalil MAK. 2000. Greenhouse implications of household stoves: an analysis for India. *Annual Review of Energy and the Environment* 25:741–63.

Smith KR, Dutta K, Chengappa C, Gusain PPS, Masera O, Berrueta V, et al. 2007. Monitoring and evaluation of improved biomass cookstove programs for indoor air quality and stove performance: conclusions from the Household Energy and Health Project. *Energy for Sustainable Development* 11(2):5–18.

Still D et al. Comparing cooking stoves. Aprovecho Research Center, Shell Foundation and Partnership for Clean Indoor Air, in press.

Organizations

HEDON Household Energy Network
http://www.hedon.info/

Partnership for Clean Indoor Air
www.pciaonline.org/

University of California at Berkeley, Center for Entrepreneurship in International Health and Development
http://ceihd.berkeley.edu/heh.htm

University of California at Berkeley, Environmental health sciences, Household environmental monitoring
http://ehs.sph.berkeley.edu/hem/page.asp?id=1

University of Liverpool, Department of Public Health, Household energy, health and sustainable development
www.liv.ac.uk/hehevaluation

World Health Organization, Capacity-building for indoor air pollution
www.who.int/indoorair/interventions/training/

In recent years, the need to provide access to modern cooking energy has attracted growing attention. One of the reasons for this trend is a better understanding of the serious health consequences of indoor air pollution from solid fuel use. As the profile of household energy rises, so does the pressure on implementers to adequately measure and report the various impacts of their work. To date, however, many of the impacts of interventions designed to reduce indoor air pollution remain poorly tested.

This catalogue of methods discusses evaluation options in the areas of Adoption, Market development, Performance, Pollution levels and personal exposure, Health and safety, Time, socio-economic and other impacts and Environmental impacts. It provides methods that range from simple questionnaires to complex monitoring techniques, and outlines practical issues related to study design, ethical considerations, data analysis and reporting. Ultimately, this catalogue of methods is intended to help governmental agencies, non-governmental organizations and universities involved with household energy and health interventions develop an evaluation strategy appropriate to their goals and organizational capacities.

ISBN 978 92 4 159691 6