



Using Operations Research to Strengthen Programmes for Encouraging Abandonment of Female Genital Cutting

FRONTIERS IN REPRODUCTIVE HEALTH
POPULATION COUNCIL



Using Operations Research to Strengthen Programmes for Encouraging Abandonment of Female Genital Cutting

Report of the Consultative Meeting on Methodological Issues for FGC Research*

April 9 – 11 2002, Nairobi, Kenya

** This meeting was supported by the United States Agency for International Development (USAID) through the FRONTIERS Program of the Population Council*

Contents

Introduction	1
Behaviour change models and FGC interventions.....	2
Designing intervention research studies	7
Managing unanticipated outcomes	12
Designing appropriate interventions.....	14
Indicators for monitoring and evaluating behaviour change.....	15
Sampling	21
Collecting data to measure indicators	23
Understanding the intervention process.....	28
Utilising research findings	29
Conclusions	31

Meeting Participants

Dr Ian Askew, Associate Director for Africa, Frontiers in Reproductive Health Programme, Population Council, Nairobi, Kenya

Dr Inge Baumgarten, Project Coordinator, Supra Regional Sector Project: Promotion of Initiatives to End FGM, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Eschborn, Germany

Dr Jane Chege, Program Associate, Frontiers in Reproductive Health Programme, Population Council, Johannesburg, South Africa

Dr Nafissatou Diop, Program Associate, Frontiers in Reproductive Health Programme, Population Council, Dakar, Senegal

Dr Hadi Eltahir, Washington, Senior Technical Advisor on FGC, USAID, Washington, USA

Ms Salma Elbeblawi, Research Coordinator, Reproductive Health, Population Council, Cairo, Egypt

Ms Michelle Folsom, Director, PATH, Nairobi, Kenya

Dr Henri van den Hombergh, Health Sector Coordinator, GTZ Kenya, Nairobi, Kenya

Ms Elizabeth Jackson, Research Coordinator, Policy Research Division, Population Council, New York, USA

Ms Jane Kamau, Project Co-ordinator, MOH- GTZ FGM Project, GTZ, Kenya

Ms Joanne Lewa, Administrative Assistant, Frontiers in Reproductive Health Programme, Population Council, Nairobi, Kenya

Dr Linda Morison, Lecturer, Medical Statistics and Epidemiology Division, Tropical Epidemiology Group, London School of Hygiene and Tropical Medicine, London, UK

Ms Agnes McAntony, Co-ordinator, National Focal Point of FGM, Northern Aid, Nairobi, Kenya

Dr Dupe Oduwole, Coordinator, Sagamu Community Centre, Sagamu, Ogun State, Nigeria

Mr Samson Radeny, Senior Program Officer, PATH, Nairobi, Kenya

Dr Karin Ringheim, Senior Program Officer, PATH, Washington, USA

Ms Evelyne Sakeah, Research Assistant, Navrongo Health Research Centre, Navrongo, Ghana

Dr Cheikh Touré, Resident Medical Adviser, INTRAH/PRIME II, Bamako, Mali

Ms Rikka Trangsrud, Deputy Director, PATH, Nairobi, Kenya

Ms Monica Wanjiru, Communications Specialist, Frontiers in Reproductive Health Programme, Population Council, Nairobi, Kenya

Introduction

Female genital cutting (FGC), also known as female circumcision (FC) and female genital mutilation (FGM), is widely practised in Africa, and to a lesser extent in some Asian countries. Defined as “all procedures involving partial or total removal of the external female genitalia or other injury to the female genital organs whether for cultural or other non-therapeutic reasons”¹, the practice has been condemned as a harmful traditional behaviour, not only because it can pose risk to the health and well being of the women and girls who are cut, but also because it violates internationally accepted human rights.

A number of interventions to encourage individuals, families and communities to abandon the practice have been undertaken over the past 70 years. These began with educational campaigns by colonial governments and Christian missionaries in east Africa that highlighted the health problems associated with the practice. Promoting messages that focus primarily on the adverse health outcomes of genital cutting has subsequently been the predominant approach used. There is little evidence, however, of there being much effect on the practice, and highlighting adverse health outcomes may even have contributed to medicalization of the practice as communities seek ways to maintain the tradition while reducing the likelihood of doing harm. There have also been efforts to stop traditional practitioners and health care providers who do the cutting, but this approach does not seem to have much effect on motivation to continue the practice. More recently, organizations working at community, national and international levels have developed and supported a range of interventions that utilize community-based behaviour change strategies (some of which are described in the next section).

For the most part, these interventions have been implemented with little attempt to document how they work, or to evaluate their impact on knowledge, beliefs, attitudes and behaviour. This is probably because most of the organizations have been small scale and working with limited budgets. Several of the larger international development assistance organizations and philanthropic foundations are now supporting such interventions with higher levels of funding, offering the opportunity to document and evaluate the interventions more systematically. Moreover, with increased funding also comes greater expectations that empirical evidence be collected to demonstrate whether or not interventions work, how and why they work, and what effect they have in ultimately influencing support for the practice. It is only recently, however, that serious efforts have been made to adapt and develop research and evaluation methods to generate quality information concerning FGC.

¹ WHO. 1997. Female Genital Mutilation: A Joint WHO/UNICEF/UNFPA Statement, World Health Organization: Geneva, Switzerland.

A review of FGC-related research issues by WHO notes that “the most neglected area is that of applied or operational research on how to design interventions that would convince individuals and communities to stop the practice. Methodologies for monitoring and evaluating different interventions are also lacking.”² To address these methodological gaps concerning intervention research, the Population Council’s Frontiers in Reproductive Health programme, with funding from USAID, organised a consultative meeting in Nairobi, Kenya in April 2002. The meeting brought together a small group of researchers and programme managers who are actively undertaking operations research and systematic programme evaluations to review the ‘state of the art’ concerning intervention research design and measurement issues. This report highlights the key issues discussed, and offers some guidance to those interested in undertaking similar research in the future.

Behaviour change models and FGC interventions

To undertake research on interventions that seek to change behaviour, researchers (as well as those implementing the interventions) should have a clear understanding of the theory underlying why and how the intervention is expected to cause such a change. **Karin Ringheim** presented several behaviour change models that can be used when designing and evaluating FGC interventions.

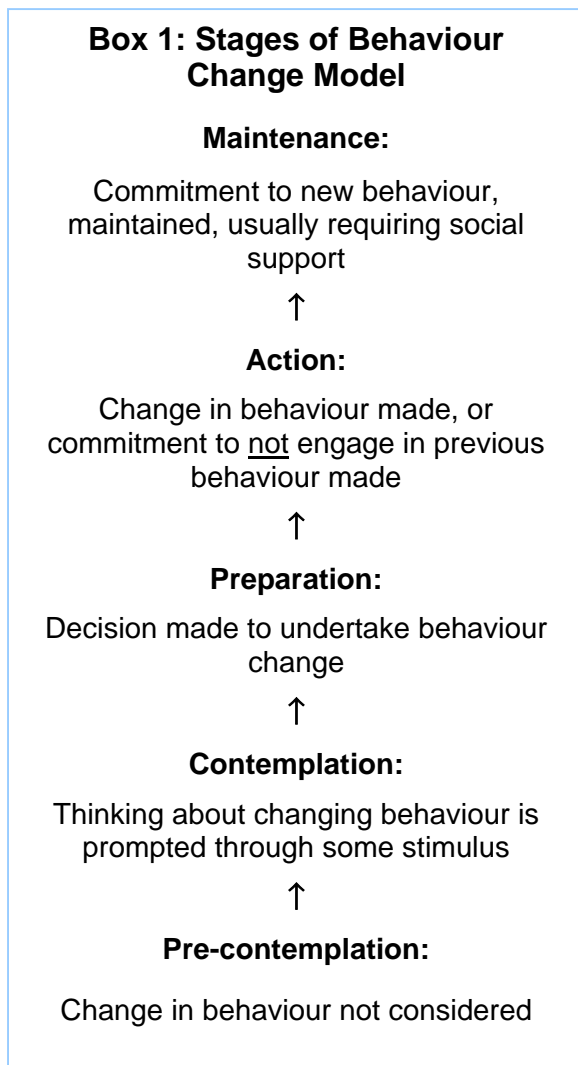
The ‘*Diffusion of Innovations*’ model³ proposes that new ideas and behaviours are not adopted by all persons at a single point in time, but rather are adopted first by “innovators”. Diffusion occurs gradually within a community until a “critical mass” of adopters has been reached, at which time more diffusion may accelerate⁴. Over time, this model has been adapted to recognize that interventions implemented in a participatory way are likely to lead to more rapid diffusion of behaviour changes. Another model is the *Theory of ‘Reasoned Action’*, which suggests that decisions are made on a rational basis following a comparison of the benefits and drawbacks of adopting a new behaviour. A limitation of this model is that people are not always rational in their beliefs and behaviours. The ‘*Stages of Behaviour Change*’ model (see Box 1), builds on the diffusion model to propose a sequence of stages that a person or community needs to pass through for a behaviour change to be made and

² WHO. 1998. Female Genital Mutilation: An Overview, World Health Organization: Geneva, Switzerland.

³ Rogers, E. M. 1962. Diffusion of Innovations, The Free Press: New York, USA.

⁴ A social diffusion model has been used to explain why a decline in FGC has only recently started in Egypt, see El-Gibaly, O. *et al.* 2002. “The decline of female circumcision in Egypt: evidence and interpretation” *Social Science and Medicine*, 54: 205-220.

then sustained. A fuller description of this model and its application to FGC interventions can be found elsewhere⁵.



For those interventions that have been designed explicitly to encourage and enable communities and individuals to move between stages, an evaluation should explicitly recognize these stages and develop a study design, indicators and data collection methods that can assess the effectiveness of the intervention in moving the population on to the next stage. In most cases, however, FGC interventions have not been designed with reference to a theoretical model, but develop in response to a particular situation (e.g. uncut girls needing an alternative ritual), or programmatic experience (e.g. a functional literacy programme leading to discussions about ending FGC), or simply intuition (e.g. converting traditional practitioners).

Lack of an underlying theoretical model can make it difficult to identify appropriate indicators for evaluation, however, because the cause and effect relationships between intervention activities and expected outcomes are not always clear. Consequently, there should be more explicit consideration of a

theoretical, or at least a ‘logical’, model when designing interventions, with an emphasis on models of behaviour change⁶. For example, in developing its forthcoming project to support anti-FGC activities in Kenya, GTZ brought together key stakeholders and then used the Logical Framework model to develop the outline for the project design.

As is clear in the model described in Box 1, a behaviour change process does not end with taking a single decision, but requires sustaining over time, especially if it is to diffuse widely so that not cutting girls becomes the social norm. This implies that when evaluating FGC

⁵ Izett, S. and N. Toubia. 1999. Learning About Social Change: A Research and Evaluation Guidebook Using Female Circumcision as a Case Study, Rainb♀: New York, USA.

⁶ For an analysis of the extent to which the effectiveness of alternative FGC interventions can be explained by sociological theories see G. Mackie. 2000. “Female Genital Cutting: The Beginning of the End”, in Shell-Duncan, B. and Y. Hernlund (eds) Female “Circumcision” in African Culture, Controversy and Change, Lynne Rienner Publishers: Boulder, USA.

interventions, a statement or declaration that a change has occurred is a necessary step and should be measured, but it cannot be taken as an indicator that behaviour has changed permanently. Measuring sustained or permanent change can only be done through a longitudinal study spanning several years, and because of resource limitations, most intervention evaluations are obliged to measure an intention to sustain the change (see section on indicators).

Most FGC interventions to date have focused on the ‘contemplation’ stage, that is, finding ways to prompt individuals or communities to start thinking about behaviour change. Community education or sensitisation campaigns have been the predominant means for ‘prompting’ contemplation, initially through messages about the health effects, but more recently with a focus on the violation of human rights that FGC represents. The relative efficacy of different messages has not been systematically assessed, and most interventions now include a combination of messages. **Jane Chege** described studies being implemented in Awash, Ethiopia, East Bara, Sudan, and the Dadaab refugee camps for displaced Somalis in northern Kenya within the framework of CARE International’s Primary Health Care programmes. Different approaches to community education are being compared in terms of their effectiveness in promoting attitudinal changes in these extremely conservative and traditional societies.

Other interventions that seek to provoke contemplation of change within communities are using individuals who can influence opinions, although there are often mixed results depending on whether these leaders are themselves considering a change in behaviour. The CARE International project in the Dadaab refugee camps, for example, has sought to involve the Islamic leaders, but they were sufficiently divided over whether or not to allow the project to be implemented that one of the three camps had to be withdrawn from the study, although the project is continuing in the other two camps. **Samson Radeny** gave a historical account of the role that Christian churches have played in FGC activities in Kenya, where some denominations, notably the Seventh Day Adventists, have consistently condemned the practice, whereas others, such as the Catholics, have had a more ambiguous position because of fear of losing members who wanted to continue to circumcise. Research findings⁷ do suggest, however, that where the church and religious leaders are actively against FGC, a behaviour change is more likely to occur.

Salma Elbeblawi presented two approaches addressing key actors in the community that influence decision-making about FGC, in Egypt. The first approach uses key community individuals, in which the minority of families or individuals in a community who have

⁷ Chege, J., I. Askew and J. Liku. 2001. An Assessment of the Alternative Rites Approach for Encouraging Abandonment of Female Genital Mutilation in Kenya, Population Council: Nairobi, Kenya

already changed their behaviour are encouraged and enabled to advocate for behaviour change. Developed and introduced by the American NGO CEDPA, these individuals are known as ‘positive deviants’ because they have deviated from the social norms by questioning these norms and refusing to follow them. Examples include parents who have refused to circumcise their daughters, a young girl who convinced her parents not to circumcise her or her sister; leaders who speak out openly against the practice; and traditional midwives who have stopped genital cutting. This intervention appears to have met with some success⁸, although it has yet to be systematically evaluated.

The second approach was implemented by the Egyptian NGO CEOSS, where community members reached a collective agreement to fight the practice of FGC and signed a document articulating their decision. Community members who participated in this decision included parents of girls within the ages of circumcision, religious and community leaders, traditional birth attendants and licensed health providers. This intervention was assessed seven years after the public declaration⁹, and revealed promising outcomes in terms of positive change in attitudes as well as behaviour. The rationale for this approach is that if those groups can be brought to a public declaration against FGC, it will provide an environment of social support and pressure that will speed up the change of attitudes and behaviours community-wide.

A more contentious approach that focuses on using key individuals as proponents of attitudinal and behaviour change is to work with healthcare providers. **Cheick Touré** described an intervention targeting providers that the PRIME II Project (implemented by Intrah) is currently testing in Mali. Drawing from experiences gained through an earlier operations research study¹⁰, the study is seeking to increase providers’ knowledge and awareness of FGC, to improve their skills in identification, treatment and referral for FGC complications, and to increase providers’ capability to educate and counsel clients and the community through a combination of training and supportive supervision. The PRIME II Project is about to start a similar project in Kenya, implemented by PATH.

This approach is complicated by the fact that many of the healthcare providers may support the practice themselves, and in some cases may actually be circumcising girls. For the providers to be able to prompt a contemplation of change among their clients and the broader community, therefore, they themselves first need to go through a process of change, at least to a stage of ‘preparation’ for change if not action itself. Indeed, both the Mali and Kenya studies will focus primarily on evaluating the behaviour change process among the providers;

⁸ Abdel-Tawab, N. and S. Hegazi. 2000. Critical Analysis of Interventions Against FGC in Egypt, Population Council: Cairo, Egypt

⁹ Abdel Hadi, A. 1997. We Are Decided: The Struggle of an Egyptian Village to Eradicate Female Circumcision, Cairo Institute for Human Rights: Cairo, Egypt.

¹⁰ Diop, N. et al. 1998. Etude de l’Efficacité de la Formation du personnel Socio-Sanitaire dans l’Education des Clientes sur l’Excision au Mali, Population Council: Bamako, Mali

change at the community level will not be evaluated within these projects as it is expected to take longer than the two years available.

PATH and Maendeleo ya Wanawake found that after their community education activities in Kenya had moved families from ‘pre-contemplation’ through ‘contemplation’ to the ‘preparation’ stage, transition to the ‘action’ stage required an explicit intervention to facilitate social acceptance of girls who were uncircumcised, which led to development of the Alternative Rites approach. **Jane Chege** discussed whether this intervention is sufficient to also move families on to the ‘maintenance’ stage, and concluded that this is a question that requires further research. She also emphasized that this intervention cannot be implemented unless families or communities are already at the ‘preparation’ stage, as it requires that an attitudinal change has already occurred to be successful. Moreover, this intervention is not always needed for a transition to be made to the ‘action’ stage, for example in communities where FGC is not an explicit rite of passage, or where sufficient social support has been developed during the contemplation and preparation stages.

Nafissatou Diop described a behaviour change intervention that is currently receiving much

Box 2 Tostan’s Village Empowerment Programme

1. **Pre-contemplation → contemplation:** A core group of women receive functional literacy education, including FGC messages. Women share knowledge with others within the village.
2. **Contemplation → preparation:** Public discussions organized during which the minority convince the majority to change behaviour
3. **Preparation → action:** Groups of villages are brought together to make public declarations against the practice and to promote a normative change
4. **Action → maintenance:** Tostan organizes media campaigns to diffuse and sustain public opinion against the practice

attention, the ‘*Village Empowerment Programme*’ developed by the Senegalese NGO Tostan. This approach attempts to support the entire behaviour change process, aiming to move whole communities from ‘pre-contemplation’ through to ‘maintenance’ using a four-step model (see Box 2). The intervention is currently the subject of two operations research studies, in Senegal and Burkina Faso, which are systematically documenting the implementation process and evaluating its effectiveness. Evidence collected by ethnographic researchers based in the study sites suggests that the intervention has been successful in achieving all four steps in

Senegal, and at least the first two steps in Burkina Faso. As noted by Diop, however, not all villages are following these steps in the same order, and so the research will ascertain

whether alternative processes influence the efficacy of the programme in bring about a behaviour change¹¹.

A major challenge when evaluating FGC interventions being implemented in areas where attitudinal and behaviour change are already occurring (e.g. many parts of Kenya) is to distinguish the impact of an intervention that seeks to accelerate these changes from the ‘natural’ process of change that is already occurring. Research from a behaviour change perspective that can explain the natural process would help to understand how the intervention could speed up this process. Evaluating interventions in such situations is best undertaken with a study design that includes a control or comparison group as well as ‘pre’ and ‘post’ intervention measures, so that the rate of change stimulated by the intervention can be compared with the natural rate of change (see section on research design).

How long it takes individuals, families and communities to go through a complete process of change is not known, but can be assumed to vary considerably depending on the strength of belief in the practice at which the community is starting, the forces working to maintain the tradition, and the nature of the intervention process itself. Intervention research needs also to ascertain the amount of time necessary to facilitate a behaviour change.

Designing intervention research studies

Most research concerning FGC to date has described and explained the prevalence of the practice, the socio-cultural and psychological context in which it is practiced, and the physical consequences of genital cutting. Such research is usually undertaken through questionnaire surveys, qualitative research, anthropological studies, and medical case histories. Research into FGC interventions is relatively recent, but the approaches used have also been varied. Some intervention research is purely descriptive, such as case studies of on-going or completed projects. Descriptive studies of interventions, which analyse and explain the way in which the intervention works, have also been undertaken. For example, a FRONTIERS study in Egypt¹² first undertook case studies of several different interventions and then synthesized the findings to provide a comparative analysis. Another study by FRONTIERS collected data from and then compared the characteristics of those families that did and did not participate in the alternative rites programme in Kenya¹³.

¹¹ For a useful discussion of the ‘convention theory’, which underlies this intervention, and an explanation of the role that a public declaration can play in encouraging abandonment of the practice, see G. Mackie *op. cit.*

¹² Abdel-Tawab, N. and S. Hegazi *op. cit.*

¹³ Chege, J. *et al. op. cit.*

To evaluate the effectiveness of interventions requires, however, a research design that follows the principles of experimental research. **Ian Askew** gave an overview of these principles and described the different types of research designs that are being used by some of the projects presented at the meeting (see Box 3). More detailed information about research designs for operations research in general can be found elsewhere¹⁴.

Study design	Example
Case study	Meta-analysis of projects, Egypt
Post-intervention assessment	Alternative rites study, Kenya
Pre- post- intervention evaluation	CARE refugee camp study, Kenya
Pre- post- control evaluation	Tostan model, Senegal
Pre- post- control OR	Tostan model replication, Burkina Faso
Two-intervention control evaluation	CARE studies, Sudan and Ethiopia
Randomised control trial	Navrongo, Ghana

All designs for experimental research have two basic features:

- There is some control over implementation of the intervention and timing of the data collection;
- The key variables can be measured before as well as after the intervention is introduced.

Using a design in which the key variables are measured before and after the intervention is introduced is the most basic design possible for testing an FGC intervention, but can only demonstrate that an FGC intervention may have had an effect if changes in attitudes or behaviour are detected between the ‘pre’ and ‘post’ measures. This design has many weaknesses however¹⁵, and so is not recommended; indeed, none of the studies reviewed here have used this design. In situations where it is known for sure that there is no ‘natural’ change in attitudes towards FGC due to development, or where there is no likelihood of other organizations carrying out interventions that influence attitudes towards FGC, it could be argued that it is only necessary to measure changes in those communities receiving the intervention, and that comparison communities are not needed to evaluate effectiveness.

Such situations are extremely rare, however, and so studies that can also make a comparison between communities receiving the intervention and those not receiving it are preferable,

¹⁴ Fisher, A., and J. Foreit, with J. Laing, J. Stoeckel, and J. Townsend. 2002. Designing HIV/AIDS Intervention Studies: An Operations Research Handbook, Population Council: New York, USA.

¹⁵ Fisher *et al.*, *op. cit.*

however, because they can measure the extent to which the intervention has caused a change over and above any other possible factors that may lead to an attitudinal or behavioural change within the same population. This type of design is being used by the CARE International studies in Ethiopia and Sudan, and by the FRONTIERS studies in Senegal and Burkina Faso. In all four studies the process of ‘matching’ was used to select the communities forming the comparison group, that is, the communities were selected because they are (or are believed to be) very similar to the intervention communities in terms of their socio-cultural and other characteristics that may influence FGC attitudes and practices.

Matching comparison with intervention communities means that, in principle, the situations are virtually equivalent, and so any attitudinal or behavioural changes found in the intervention communities that are not found in the comparison communities can be attributed to the intervention itself. Using a research design that allows this is important, because it also enables the researcher to better understand and manage other FGC-related factors that may be influencing people’s attitudes and behaviour. This is discussed in more detail in the section on ‘managing unanticipated outcomes’.

In some studies, a comparison is made between communities receiving one type of FGC intervention with communities receiving another type of FGC intervention. Studies that include more than one intervention can be more useful as they compare alternative ways of addressing the same issue among the same population. An example of this design is the CARE International study in two Somali refugee camps at Dadaab in northern Kenya, which is testing and comparing two interventions:

- A targeted education activity, through which health and social workers in the camps discuss FGC in both group and individual interactions.
- An advocacy activity, through which community leaders are educated on FGC and then trained to sensitise community members on the negative aspects of FGC through public pronouncements and discussions.

The study is designed to compare the effect of the targeted education activity in one camp with that of the two interventions implemented together in the other camp, with the expectation that the effect of the combined interventions will be greater than the single intervention. The effect of each approach can be compared with the existing situation because ‘pre’ intervention measures of attitudes were taken through a baseline survey, and ‘post’ intervention measures will be taken through an endline survey in late 2002.

An alternative way of creating intervention and comparison communities for a study is to firstly select a single group of similar communities, and to then randomly assign each of the communities to either receive or not receive the intervention. This is an especially strong research design when the group of similar communities are completely equivalent, because it ensures that the groups of intervention and comparison communities are directly comparable.

It is extremely difficult to create in real life situations, however; to date this design has only been used once in an FGC intervention study, at the Navrongo field station in northern Ghana. Rates of circumcision were not equivalent among the culturally alike, geographically contiguous villages in the Navrongo experiment. Because of this, it was necessary during analysis to control for differences in background characteristics at the individual level and baseline FGC prevalence at the cluster level.

The Navrongo study was designed to test two intervention strategies in four paramount chiefdoms¹⁶ in a high prevalence area of the Kassena-Nankana District. The two smaller paramuncies were combined into one area to form three areas of approximately equal size. The two interventions are implemented jointly in one area and separately in the other two, so that the relative effectiveness of each of the three possible combinations of interventions can be tested and compared. Each of the three paramuncy groups was randomly assigned to receive one of the three intervention strategies. Groups were ranked according to numbers assigned from a conventional random number table to determine which intervention strategy they would receive.

A "stepped wedge design" was used to implement the intervention in stages, beginning with a pilot area. During analysis, each area is treated as a comparison area prior to the time it begins to receive the intervention. Because of lower than expected FGC incidence rates, however, implementation occurred faster than is appropriate for the statistical evaluation of results. For this reason, the study area population has been increased by the addition of two neighbouring paramuncies, which serve as comparison areas. Because these additional comparison areas were not under observation at the start of the study three years ago, it is necessary to collect retrospective data on FGC incidence in these areas.

Although including comparison communities is strongly recommended, the resources needed for such a design may be greater than some funding agencies are willing to contribute. Traditionally, funding within project budgets for evaluation has been limited to an end of project assessment only, and many interventions are still being introduced without a plan or budget for evaluating their effectiveness through a systematic design. **Inge Baumgarten** described how GTZ requires all of the organizations it supports to undertake, at the very least, a baseline survey prior to developing and introducing interventions so as to be able to implement a monitoring system with a minimum set of indicators to assess their effectiveness

¹⁶ Lineages have chiefs or sub-chiefs that define communities. Groups of communities have paramount chiefs. In most of northern Ghana, around 1900, paramount chiefs were appointed by the Colonial government, as large chieftaincy units were non-existent, and paramuncies were needed for indirect rule. Chiefs were also new concepts at that time. Instead, "Tendana" were traditional leaders responsible for land distribution and religious functions. For this reason, community cohesion in northern Ghana, and Kassena-Nankana in particular, often lacks the vitality of community organization in southern Ghana.

at some later point in time. Given that the costs of adding comparison communities also would almost double the evaluation budget, it is understandable why funding agencies that do not normally support strong evaluation components may be reluctant to do so for FGC projects. The onus is on researchers, therefore, to demonstrate that the benefits to be gained by using a stronger design outweigh the additional costs. USAID, as described by **Hadi Eltahir**, is one of the few donors willing to support the use of such intervention research designs (although the CARE projects are funded through the CARE Africa Fund and private individual donors as well as USAID), and so to maximize this support FRONTIERS is actively forming partnerships with organizations that are interested in evaluating their interventions in an experimental framework.

One aspect of research into FGC (and particularly in intervention research), that can often be downplayed or sometimes simply ignored, are the ethical principles underlying the way the study is designed and the data collected. This issue is the focus of a paper by Abdel-Tawab, presented at a conference in Bellagio, Italy. Put briefly¹⁷, she highlights the need to ensure that any intervention study should follow the principles of:

1. *Beneficence*: the study maximises the potential benefits it can produce for individuals and society;
2. *Non-maleficence* the study minimizes any potential harms (including physical, social, psychological) to individuals and society;
3. *Respect for autonomy*: the study respects the rights of individuals and groups to make decisions for themselves, and protects persons with diminished autonomy (e.g. children);
4. *Justice*: the study treats all subjects equally and disadvantaged community members do not bear a disproportionate burden of a study from which all members of a community will benefit.

¹⁷ For a detailed discussion of the application of these ethical principles to FGC intervention research see: Abdel-Tawab, N. 2002. "Ethical Issues in conducting interventions research on FGC", Paper presented at the Conference on 'Advancing Research on Female Genital Cutting', Bellagio, Italy, April 29th – May 3rd.

Managing unanticipated outcomes

The primary reason for using one of the research designs described above is to be able to control other activities that may be influencing the outcomes that the intervention itself is expected to influence (which in this context are the norms, attitudes and behaviour concerning FGC). Community level behaviour change interventions cannot be implemented in total isolation from other factors that can also influence these outcomes, however, and so a strong intervention study needs to be able to manage and explain outcomes that were unanticipated when the study was designed.

For studies that include comparison groups, the major challenge is to manage possible ‘contamination’ of the comparison populations, that is, if they also become exposed to the intervention. As was described by **Jane Chege**, although it was known that the nomadic lifestyles of the Afar community in Ethiopia and the East Bara communities in Sudan would present challenges to controlling exactly who is and who is not being reached and with what messages, the high level of interactions between groups assigned to receiving the intervention or not means that the original design will probably have to be abandoned. However, this situation does present the opportunity to examine the way in which the intervention’s behaviour change messages are diffused. In addition to monitoring and documenting what has been taking place in both the experimental and control sites, qualitative information will be collected to map the migratory patterns and community interactions. Further, questions on exposure to intervention messages and sources of information were included in the endline survey instruments to assess the level of diffusion of FGC behaviour change messages.

Participants gave examples of other factors that have produced unanticipated outcomes. **Samson Radeny** described how, since the beginning of colonial administration through to the present day, many church groups in Kenya have been valuable partners for FGC eradication interventions. PATH has over the years worked in close collaboration with church groups. For example, the Seventh Day Adventist (SDA) church has appreciated working with an organization that can systematically design and implement an anti-FGC intervention, and they are able to mobilize much larger audiences for these interventions than PATH or any other community-based institutions. However, because the SDA is such a large and disparate organization, it is extremely difficult for PATH to identify the relative effects of the interventions it does jointly with the church from its own interventions and from the anti-FGC activities that the church undertakes without PATH. This poses a challenge not only in drawing conclusions about effectiveness of interventions, but also what constituted a particular intervention.

Karin Ringheim discussed medicalisation of FGC. Although it was an unanticipated and unwanted outcome of efforts to educate communities about the health complications of FGC in Kenya, it can be understood as a rational response on the part of parents who intend to

continue the practice but want to minimize potential harm to their daughters. Having the procedure done by medical practitioners increases the likelihood that the procedure is undertaken in sanitary conditions, and that antiseptics, painkillers and anti-tetanus drugs may be used. Potentially, medical providers may see to it that there is less tissue cut and therefore less trauma to the girl. However, medicalising the procedure still constitutes a violation of a girl's right to bodily integrity and does not address long-term sexual, reproductive and mental health complications that may result from FGC. Medicalisation provides a financial incentive for health providers to take up the practice, making it more difficult to eradicate. WHO, USAID, PATH and other international organisations are unequivocal in their opposition to medical providers engaging in the practice, and some are working to engage health care providers in opposing medical participation in FGC. It remains, however, a contentious and complex issue¹⁸.

Experience from Navrongo, as presented by **Elizabeth Jackson** and **Evelyne Sakeah** showed that both researchers and implementers should always be prepared to meet unanticipated outcomes that could result in changing the study design. Following the baseline survey, it was found that the prevalence of FGC was considerably lower than anticipated among younger women, and it became necessary to increase the number of girls in the control area by interviewing in an adjacent area to collect information about retrospective FGC incidence. Researchers have to accept that pre-planned designs can “breakdown”, even after following a participatory planning process.

In many countries it is becoming increasingly important to control for national legislation or presidential decrees against FGC. **Jane Kamau** described the situation in Kenya where there have been a number of presidential decrees banning FGC, and legislation has recently been passed which now outlaws the practice. Because such factors can affect the whole country, and can greatly influence the way in which people answer questions about FGC, it is hard to manage or account for their influence without using a comparison group.

It is important, therefore, that wherever possible a study includes comparison groups of communities because changes in attitude and behaviour can be due to so many factors outside the intervention itself, either due to planned interventions such as the stance of a church or national legislation, or to ‘natural’ changes that are occurring as levels of education and awareness of other cultures increase. It is also important that an intervention design is appropriate for a particular socio-cultural or ethnic group and that their reaction can be

¹⁸ See Shell-Duncan, B. 2001. ‘The Medicalization of female ‘circumcision’: harm reduction or promotion of a dangerous practice?’ *Social Science & Medicine* 52: 1013-1028; and Mandara, M. ‘Female Genital Cutting in Nigeria: Views of Nigerian Doctors on the Medicalization Debate’, and Shell-Duncan, B. *et al.* ‘Women without Choices: The Debate over Medicalization of Female Genital Cutting and its Impact on a Northern Kenyan Community’ in Shell-Duncan, B. and Y. Herlund (eds). 2000. Female “Circumcision” in Africa: Culture, Controversy and Change. Lynne Rienner Publishers: Boulder, USA.

anticipated. **Dupe Oduwole** gave an example from Nigeria where qualitative research has proved necessary among different sub-groups of the Yoruba ethnic group, because not all Yoruba hold similar beliefs or practices regarding genital cutting.

Designing appropriate interventions

The organization introducing an intervention must understand as fully as possible the context in which it is operating. Consequently, descriptive (or formative) research should always be undertaken first, and this step must be included within a study plan. **Jane Kamau** gave a vivid example of the importance of this when she presented results from a baseline survey (that collected both quantitative and qualitative data) among the Kalenjin community in Rift Valley Province, Kenya. This survey indicated an unexpected and apparently drastic decline in the practice among the youngest generation, thereby convincing GTZ not to continue with introducing the planned intervention.

Given the need to understand community values and practices when designing an intervention, and to design an intervention that is fully acceptable to the community whose behaviour it is intended to change, it is strongly recommended that a participatory learning approach (PLA) be used before starting an experimental study design. Using PLA to generate behaviour change makes community members more confident in solving their own problems because they become part of the problem-solving process. **Dupe Oduwole** gave an overview of the key principles of PLA and how they will be applied in a forthcoming project to encourage abandonment of FGC in southern Nigeria. **Evelyne Sakeah** described how PLA was used first to understand the values and motivations underlying FGC in northern Ghana, and then to gather suggestions from a wide range of community representatives, which led to the development of two interventions: a series of health education activities and activities to develop the autonomy of girls and women. **Jane Chege** then described how CARE International used PLA to identify the key components of the interventions that they have introduced in Ethiopia, Sudan and northern Kenya.

Indicators for monitoring and evaluating behaviour change

All FGC interventions have the same ultimate goal – total and sustained abandonment of the practice. The wide variety of FGC interventions demonstrates, however, that there may be many different ways, in principle at least, to achieve this goal. However, one of the obstacles to the evaluation of specific interventions, and subsequent systematic assessment of characteristics associated with success, has been difficulty in identifying valid and feasible impact indicators. RAINB♀ is in the process of completing its Review, Evaluation and Monitoring (REM) project, which reviewed a range of indicators being used to evaluate FGC interventions. The aim of the REM project was to propose indicators for evaluation of FGC interventions that would generate the information needed to assess which approaches and what characteristics of implementation are important for success. Whether there is a need for standardised indicators, especially as abandonment of FGC is so context specific, was debated. It was emphasised that the proposed indicators could serve as a guide to the type of information that would be useful for comparing different approaches and would have to be adapted to specific contexts. On behalf of the REM Project, **Linda Morison** presented the preliminary findings of the review¹⁹.

The presentation started with a review of problems surrounding measurement of what is considered the primary indicator of an intervention's impact: the incidence of FGC. This type of data is normally gathered during a survey in interviews with mothers about recent circumcision status of their daughters. In some cases where FGC is performed at an older age the girls themselves may be asked for this information (as is the case in Navrongo). As is described later, studies have shown that the validity of such reports can vary, however, as high levels of agreement between reported status and that found on clinical examination have been found in rural Gambia and Egypt. But at Navrongo in northern Ghana, the context of recent legislation against the practice and an active FGC intervention led to over 60 percent of women denying being circumcised after previously reporting that they had been circumcised. Clinical examination of the genital area would be required, therefore, if inaccurate self-reporting is anticipated, but this is obviously unacceptable in a survey situation (although inviting parents to bring girls for voluntary confirmation might be possible).

Linking indicators to the Stages of Behaviour Change Model (see Box 1) and ensuring project objectives were feasible within a particular timeframe was emphasized. For example,

¹⁹ The full set of indicators is described in the forthcoming report: RAINB♀. 2002. [A Framework for Design, Monitoring and Evaluating Anti-FGM Programs: A report of the Female Genital Mutilation Review, Evaluation and Monitoring Project 2001-2002](#), RAINB♀, London, UK.

in a community at a predominantly “pre-contemplative” stage a feasible objective might be to move them to the “contemplative” or “preparation” stage, rather than the later stages.

Box 4 Indicators for non-FGC outcomes of the village empowerment intervention

Knowledge:

Recall messages on body physiology, reproductive health and human rights;

Understanding consequences of harmful practices and the importance of good reproductive health practices;

Awareness of discriminatory practices against women

Awareness of gender-based violence

Attitudes:

Recognition of the relative advantages of using reproductive health services;

Communication with spouse and friends on family size and reproductive health;

Acceptability of women heading mixed community groups;

Propose strategies for reducing discriminatory practices against women;

Propose strategies for reducing gender-based violence.

Action:

Discussion of reproductive health and human rights in public;

Testimony of experiences in public;

Advocate ideas to others;

Join or create a movement against violence;

Make a public declaration against FGC.

A crucial stage in the process of sustained abandonment of FGC is the maintenance against possible opposition of a decision against FGC (for example a mother deciding not to circumcise a daughter). In situations where whole villages (and in some cases whole intermarrying areas) have publicly declared against the practice, it is expected that opposition to an individual decision against FGC will be minimised. Holding a public declaration is likely to represent a shift in actual practice, and would thus be a good indicator of reduction in incidence. The need for long-term evaluations of FGC interventions was stressed, because circumcision can be performed over a wide age-range and a short-term evaluation might only be detecting a delay in age at circumcision rather than a circumcision averted.

The need also to examine and measure the influence of anti-FGC interventions on women’s empowerment was emphasized. Some programmes accept women’s lower social status as being too difficult to influence, and primarily target men to bring about an end to the practice,

but this it can be argued that this is unethical as it perpetuates women’s disempowerment. Consequently, developing and measuring indicators that can measure the broader impact of anti-FGC interventions (for example, on gender relations, women’s rights and community development) is important. **Nafissatou Diop** described some of the difficulties in developing and using such indicators from the on-going studies to test the Village Empowerment Programme in Senegal and Burkina Faso. Box 4 describes some of the indicators being used in these two studies.

As discussed by **Karin Ringheim**, it is essential that age-specific incidence and prevalence rates be calculated for measuring the impact of interventions on FGC. Using data from 1993 baseline and 1999 endline surveys of PATH FGC projects in Kenya, Ringheim showed that even though a statistically significant decline in FGC from 90 to 82 percent was found in the overall female population aged 14 – 60 years, this decline was much more pronounced when the women were disaggregated by age group. Since more than 95 percent of women over age 30 in the project areas had already been circumcised, these women could no longer be affected by an intervention to eliminate FGC. Thus age-specific rates for being cut should be used when evaluating interventions. Moreover, if resources are limited, only the younger age ranges need be measured at the endline survey to determine changes in the prevalence of cutting.

Calculating age-specific rates and comparing them over time requires, however, that age at cutting be recorded for both surveys, and that sufficiently large sample sizes be used for each age range and in both rounds of surveys. Calculating age specific rates is difficult in countries such as Mali, where genital cutting remains nearly universal and is increasingly taking place in the first year of life. Ringheim’s analysis of Kenya data was constrained by the fact that age specific rates were not reported in the baseline. However, it was known that 78 percent of girls had been circumcised before age 20 in 1993, whereas by 1999, the percentage of girls under age 20 who had been cut had fallen to 56 percent.

Box 5 Observed and expected prevalence rates by age group

Age	Observed	Expected
14	30%	52%
15-17	47%	71%
18-22	74%	80%
23-30	84%	NA
31-35	95%	NA
35-45	91%	NA
46-60	97%	NA

To estimate the decline in FGC among younger girls who had some likelihood of being affected by the intervention, she compared the observed rates by age range from 1999 with ‘expected’ rates for 1999 (see Box 5). ‘Expected’ in this case is defined as the proportion of girls in that age range that would be expected to be circumcised if the age at which women were circumcised still prevailed and if 82 percent of the total sample were eventually to be circumcised, as was the case in 1999.

Controlling for age when analysing the impact of exposure to interventions is now becoming a key feature in analyses of behavioural changes and the role of interventions. For example, Chege and colleagues²⁰ controlled for age when comparing the reported cutting of daughters

²⁰ Chege, J. *et al.* p38.

by parents exposed to sensitisation activities in Kenya, and El-Gibaly and colleagues²¹ did the same while comparing self-reported cutting among girls of different ages whose parents had been exposed to media publicity about FGC in Egypt.

A further example was given by **Elizabeth Jackson** for the longitudinal study in Navrongo, Ghana, which recruited a cohort of girls who were aged 12 - 19 years during the baseline survey in 1999, 12 years being the age that they are first at risk of being cut. This is an “open” cohort in that each year of surveillance other girls aged 12 are added. The impact of the two arms of the intervention implemented separately and together will be compared in terms of the proportion of the observed decline in FGC incidence that can be attributed to each of the three possible strategies using a discrete time hazard logit model (see Box 6). This decline will be measured in terms of the difference between the incidence rates in

Box 6 Discrete Time Hazard Logit Model

The relative effect of each arm of the experiment on incidence of FGC is calculated using a discrete time hazard logit model. Data are ordered by individual and by year, and observations are censored in the dataset after circumcision has taken place or when a girl is lost to observation. Logistic regression analysis is then performed on the dataset in order to calculate the yearly probability of incidence of FGC by experimental exposure, relative to incidence in the comparison area. Possible covariates are controlled for, including age, time, education, marriage, religion and baseline FGC prevalence in each cluster.

communities exposed to each intervention compared with the incidence rates in communities not exposed. Other outcome variables through which impact of the interventions will be measured include changes in the level of awareness of the effects of FGC and public declarations by mothers, girls, circumcisers and lineage heads that they will no longer support or take part in the practice.

In the studies to evaluate the Village Empowerment Programme in Senegal

and Burkina Faso, the key dependent variable is also whether or not it will influence the cutting of young girls who are currently not cut. As described by **Nafissatou Diop**, the indicator being used is the proportion of daughter(s) aged 0 – 10 years whose mother reports that they are cut (in these countries, 95 percent of girls are cut by age 10 years). The assumption is that if the intervention has an impact, parents of girls aged 0 to 10 years will be more willing not to cut their daughters, and this will be reflected in a lower probability that mothers will report having cut their daughters by the age of 10 years.

This study is, by necessity, following a panel of women over time, and not recruiting cohorts during the baseline and endline surveys because it is tracking women exposed to the intervention. By using a pre-post control group design it is relying on three assumptions to test the hypothesis: that the proportions of girls reported to be cut will be statistically equal in

²¹ El-Gibaly, O. *et al. op cit.*

the experimental and control groups during the baseline; that the proportions of girls reported to be cut in the experimental group will be statistically lower than in the control group during the endline; and that the proportion of girls reported to be cut in the control group will increase from the baseline to endline. The change over time referred to is, therefore, between the two groups and not within each group.

A life table analysis will be used to measure any change in the probabilities of being cut between the baseline and endline times for both groups, with the expectation of no change in the probability of being cut in the experimental villages and an increase in the control villages. A life table analysis is the most appropriate way to calculate this indicator because the experience of some of the girls in these samples will be ‘censored’, that is, they are not yet circumcised but might be in the future. The basic idea of the life table is to subdivide the period of observation (i.e. 0-10 years) into smaller time intervals (i.e. years). For each interval, all girls who have been observed are used to calculate the probability of being cut occurring during that interval. The probabilities estimated from each of the intervals are then used to estimate the overall probability of the event occurring at different time points.

Samson Radeny highlighted the difficulties in not only implementing an approach that explicitly promotes the idea of FGC as a contravention of basic human rights to bodily integrity, but also in evaluating it. Implementation required simplifying the terms and concepts used in legal documents so that they could be used by fieldworkers and understood by communities who may never have discussed such issues before. Evaluating changes in people’s awareness of and agreement with such rights has proved challenging in questionnaire surveys, because the concepts often require extensive explanation and discussion before questions on awareness, understanding and approval can be asked.

Most interventions are implemented for a limited period of time and, as described earlier, may only be able to influence transitions to the early stages of behaviour change.

Consequently the indicators that should be used during an evaluation need to be developed carefully so as to measure the degree of change that can reasonably be attributed to the intervention. For example, statements of intention not to cut future daughters are commonly used as indicators of behaviour change, both during individual interviews and as community or group declarations and pledges. Whether these statements should be accepted as indicators of behavioural change, i.e. having reached the ‘action’ stage, or are more appropriate as indicators of attitude change, i.e. the ‘preparation’ stage, was debated without agreement. Whereas a public pledge not to cut future girls is likely to bind families into a new social norm because of the strong meaning attached to public declarations in many African cultures, making such a statement during an individual and confidential interview may reflect a courtesy bias to the interviewer rather than a genuine change in the person’s intention.

Interventions that influence the role that health staff can play, both in managing women presenting with FGC-related complications and as behaviour change agents among their clients and community, require somewhat different indicators than those interventions that seek to influence behaviour change directly among community members. Health staff are usually from the local community themselves, and so it cannot be assumed that they would be personally against the practice and/or willing to speak out against it with their clients or in public. Consequently, it is important to measure the extent to which a provider-oriented intervention is able to influence their attitudes and behaviour. **Cheick Touré** presented the indicators and data collection methods that Intrah is using in its study with the Ministry of Health in Mali (see Box 7).

Box 7 Indicators for measuring effectiveness of health staff interventions in Mali

Indicators	Data collection
<p>Management of complications</p> <ul style="list-style-type: none"> • Number of service providers whose scores on FGC management, counseling and education tests met standards • Number of clients receiving counselling related to FGC 	<p>Management of complications</p> <ul style="list-style-type: none"> • Written test regarding FGC complications • Written questionnaire on provider perceptions and FGC-related attitudes • Review of clinic registers and medical files for data on FGC management • Case studies to collect data on current provider FGC management practices according to management procedures • Observation of service providers in real situation: Welcome and history taking with a patient in pre-natal consultation
<p>Behaviour change communication</p> <ul style="list-style-type: none"> • Number of health talks given on the subject of FGC • Number of community sensitisation / education sessions on FGC targeting men and youth, in which the provider served as a resource person 	<p>Behaviour change communication</p> <ul style="list-style-type: none"> • Simulation: Counselling role plays specific to FGC • Observation in real situation: Reproductive health education talk • Provider diary tracking number of clients counselled on FGC and number of FGC talks given

Sampling

Calculating the appropriate sample size and developing a strategy for identifying the respondents for a sample are extremely important components of operations research, and can negate the findings of a study if they are not well planned and implemented²². The most important issue to bear in mind when developing the sampling plan for intervention studies is that the principles on which the sample size are calculated are different from those used for calculating samples in descriptive studies. This is because the sample size for a descriptive study is based on the need to measure the frequency with which a variable (for example, prevalence of FGC) occurs within a population, whereas the sample size for an intervention study has to enable a researcher to compare the frequency of a variable at two or more times or situations (i.e. in the baseline and endline intervention surveys, and/or in the intervention and control groups), and to be able to detect statistically a change and whether or not there is a meaningful difference between them.

More details about these different approaches to calculating sample sizes can be found elsewhere²³, but it is essential that researchers and programme staff fully understand this difference when designing and implementing an FGC intervention study. **Jane Chege** described situations where baseline surveys had been undertaken with sample sizes calculated for descriptive rather than intervention-testing purposes, and the consequent problems in trying to detect statistically significant differences between intervention and control groups, as well with the endline survey measures.

In calculating sample sizes for evaluating interventions, it is important to do at least the following two things:

- Identify the single most important key variable of behaviour change with which you would like to judge the success of the intervention. You also need to either know or be able to make an accurate estimate of the current level of that variable in the target population.
- Decide on the magnitude of change that you think the intervention can realistically achieve from this current estimate within the time available, and which would be judged as a success by those interested in this intervention.

Another statistical consideration arises because FGC interventions are usually implemented at the community level, or among groups (or ‘clusters’) within a community. Sample size calculations must therefore take into account the fact that people within the same community

²² See earlier discussion of problems faced in Navrongo when it was found that the incidence of new FGC cases was much less than had been estimated.

²³ For example, Fisher *et al.*, *op. cit*

or cluster tend to be more similar than people from different communities or clusters. Because of these possible differences between the clusters, sample size calculations for evaluating community-level interventions should take into account this ‘design effect’ as it is known. Failure to include this can result in underestimating the sample sizes needed and consequently the results produced may not be statistically valid. Simple methods of including

Box 8 Examples of sample size calculations

The studies in Senegal and Burkina Faso that test the Village Empowerment intervention are both using the key indicator of ‘proportion of women supporting the practice’ to evaluate its impact. This education programme is expected to create a major shift in attitudes, and so a reduction by half in the proportion of women supporting the practice in the experimental group is felt to be an acceptable indicator of success.

Previous studies showed that in 1999 68% of women in Senegal supported the practice while this figure was only 32% in Burkina Faso. Because the pre-intervention measures for this indicator are different in each country, the sample size needed to detect a significant difference will necessarily be different. The table shows sample sizes for 90% power to detect a difference at the 5% level of significance.

Country	Sample size ignoring the design effect	Sample size incorporating the design effect
Senegal	49	80
Burkina Faso	160	300

Sample sizes are bigger for Burkina Faso because the difference between the control and intervention villages is expected to be smaller (i.e. 32% to 16%) compared with Senegal (68% to 34%).

Moreover, the sample sizes that do incorporate the design effect for a comparison between 20 intervention and 20 control villages are much larger than if these ‘between village’ differences had been ignored, showing that an under-estimation of the sample size needed to produce statistically valid results will occur if this factor is ignored.

this ‘clustering’ consideration in calculating sample sizes for community-level interventions have been developed²⁴, although a statistician should always be consulted.

There are other statistical considerations to be considered when calculating the sample size (e.g. confidence and significance levels), but these are the main considerations and ones that will require much discussion and negotiation between those doing the research and those implementing the intervention to get agreement on them. Box 8 describes an example of how these considerations lead to substantially different sample sizes for studies that are testing the same intervention and using the same impact indicator, but under different circumstances.

The implementation of the CARE study in two countries faced a great challenge in getting agreement on sample sizes with program managers and consultants hired to conduct the

baseline survey. Most of the programme managers and consultants are more familiar with the sample size calculation for a descriptive study and could not understand the justification for the approach used in calculating sample sizes for evaluating interventions. Indeed, the

²⁴ For example, Hayes, R. and S. Bennett. 1999. “Simple sample size calculation for cluster-randomized trials” *International Journal of Epidemiology*, 28: 319-326.

endline survey for one of these studies has subsequently had to be abandoned because the baseline survey sample size was not large enough to allow meaningful pre- and post-comparisons to be made.

It is also essential that sample sizes be calculated separately for each of the sub-groups within a population whose behaviour the intervention is intended to change. The decision as to which sub-groups (e.g. adult women, adolescent girls, fathers, etc) to sample will be determined by the causal logic underlying the behaviour change intervention, the indicators chosen, and of course by the budget available. If resources are limited, then it will become necessary for these sub-groups to be prioritised in terms of the relative impact of the intervention.

Collecting data to measure indicators

Because of the sensitive nature of FGC, a number of problems can arise when measuring indicators describing the practice. First and foremost is whether or not a woman or girl is actually circumcised, and if so, with what type of cut. Most research into FGC-related behaviour and interventions is undertaken using individual or group interviews, through which respondents are expected to verbally describe their own or their daughter's status. It is impossible, however, to validate the person's response given without actually observing their genitalia, which obviously poses huge ethical and logistical concerns. On the four occasions when this has been done (in Egypt, the Gambia and twice in Nigeria)²⁵, a significant difference between the woman's self-reported status and her observed status was only found in urban Nigeria, suggesting that questioning individuals may be a valid way of measuring this indicator. Two presentations illustrated situations, however, in which it is unlikely that self-reported status does accurately reflect actual status (although neither study was able to use observation for validation).

From the longitudinal FGC study being undertaken at Navrongo in northern Ghana, **Elizabeth Jackson** reported that while longitudinal observation gives researchers a unique

²⁵ Huntington, D. *et al.* 1996. Clinic-based Investigation of the Typology and Self-reporting of FGM in Egypt, Population Council: Cairo, Egypt.

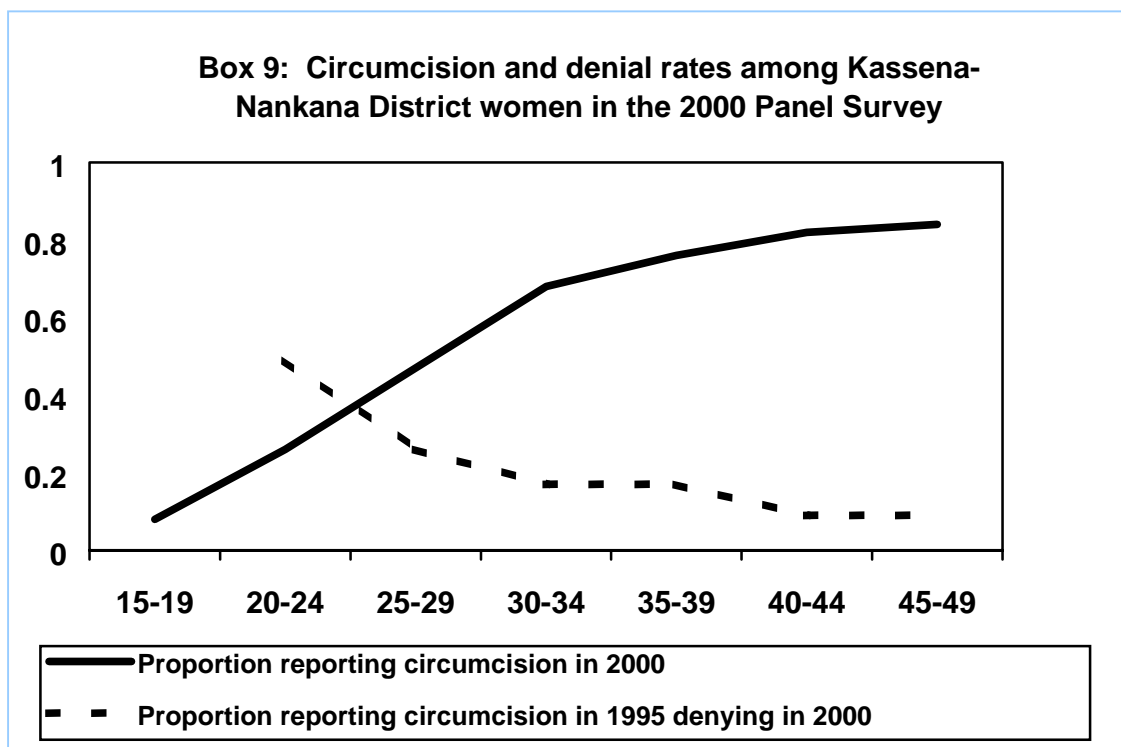
Morison, L. *et al.* 2001. "The long-term health consequences of female genital cutting in rural Gambia: a community-based survey", *Tropical Medicine and International Health*, 6: 643-653.

Larsen, U. and F. Okonofua. 2002. "Female circumcision and obstetric complications", *International Journal of Gynecology & Obstetrics*, 77: 255-265.

Snow, R. *et al.* 2002. "Female genital cutting in southern urban and peri-urban Nigeria: self reported validity, social determinants and secular decline", *Tropical Medicine and International Health*, 7,1: 91-100.

opportunity to assess the determinants of FGC on an individual level, repeated observation exposes the fact that some women report their circumcision status inaccurately or inconsistently. Findings from questionnaire surveys suggest that two factors, recent legislation against FGC and declining prevalence associated with changing attitudes about FGC, may cause women who have previously reported being cut to subsequently report not being cut. Qualitative research suggests this inconsistent reporting is likely to be denial of being cut during the second interview because of a reluctance to reveal being circumcised to interviewers (rather than falsely reporting being cut during the first interview and accurately reporting not being cut during the second interview).

In the first set of surveys, interviews were carried out with women of reproductive age in 1995 and again five years later. Fifty percent of women aged 20-24 who reported that they were circumcised in 1995 reported that they were not circumcised in 2000. A much smaller proportion of women aged 45-49 (10 percent) denied their status (see Box 9). Women who denied being circumcised after earlier indicating that they had been cut were not only younger, but also more likely to be educated, and less likely to practice traditional religion than circumcised women who did not subsequently deny their status. Women who denied were less educated and more likely to practice traditional religion than women who consistently stated they had not been cut.



A second set of surveys interviewed adolescent girls aged 12-19 in 1999, 2000, and 2001. Denial among this group appeared to increase over time and with exposure to anti-FGC intervention activities. A qualitative study is currently underway to improve response accuracy on this sensitive issue. Jackson concludes that in situations where the prevalence of FGC is decreasing due to changes in social norms, where there is legislation against FGC, or where an intervention to discourage the practice is underway, it is particularly important to assess whether or not inaccurate reporting of FGC status biases results.

A similar situation exists in Burkina Faso, where FGC has been illegal since mid-1996, and this fact is well known – three quarters of a sample of people interviewed in rural Burkina Faso during a 1998 survey knew that FGC was illegal. Clearly this may influence responses to questions concerning the practice, and indeed data from 1996 and 1998 panel surveys undertaken by the Population Council support this concern. In the 1998 survey, 33 percent of those who did not know that the practice was illegal said that they would cut their daughter(s) in the future, whereas among those who did know that FGC was illegal, only 10 percent indicated an intention to cut their daughters.

In the context of the Burkina Faso study to test the Village Empowerment Programme, there is uncertainty about the validity of those respondents who state that their daughters are uncut, or who do not want to cut their daughters in the future. The illegal status of the act means that it is not known for sure whether they are replying truthfully or not. Consequently, the proportion of parents who state that their daughters are cut is a preferable indicator because it is more likely that these persons are telling the truth, whether or not they know that it is illegal. During analysis of data, however, care will be taken to control for factors such as knowledge of the legality of FGC, age of daughter, cut status of other sisters, etc.

Jane Kamau described a situation from the Rift Valley Province in Kenya where a recent survey (among 382 adult women and 312 girls aged 15 – 20 years) reported that prevalence had declined from 65 percent among the adults to two percent among the adolescents. Although a decline in the practice could be explained in terms of an apparent change in social norms (as was reflected in responses to questions on attitudes and intentions), which was attributed to the influence of the church and improvements in education, the virtual eradication of the practice within one generation is not credible. Qualitative data from in-depth interviews and group discussions revealed two alternative explanations, however. First, key informants suggested that there is underreporting because a proportion of women are cut after completing schooling to improve their marriageability, and some even after marriage, and so these women would not be reflected in the survey. The mean age at cutting is 14 years, however, suggesting that these examples of delayed cutting are likely to be small in number. A more plausible explanation given is that the survey was undertaken in Koibatek District, which is the home of the Kenyan President who has been a vocal and sustained

opponent of FGC since 1982. Consequently, although there is certainly a marked change in social norms, those continuing the practice are probably unwilling to admit it.

Problems were also identified around measuring age at circumcision, type of cutting, and adverse outcomes associated with circumcision. All three indicators are normally used in descriptive studies of the practice or to assess changes in the nature of the practice over time, rather than as outcome indicators of interventions. Changes in these indicators can be useful, however, to measure an intervention's (usually unintended) effect on the nature of the practice. For example, it may be important to assess the extent to which the practice has been 'medicalized', or less severe types of cutting used, as a result of exposure to health-oriented anti-FGC messages, or when the age of cutting has declined to reduce the likelihood of possible opposition when the girl is older.

Ian Askew described a study carried out in a sample of clinics in Burkina Faso and Mali²⁶, in which trained service providers observed and recorded the type of cut and presence of any longer term gynaecological and obstetric complications associated with genital cutting among reproductive health and delivery clients who had a pelvic exam for any reason during their consultation. Although this study provided much useful information, there were a number of methodological issues that have been subsequently addressed in other studies. For example, sampling among women attending clinics for reproductive health services and who have a pelvic exam may be a biased representation of the general population, and so a community-based survey would be more representative²⁷. Complications associated with delivery need to be analysed independently, and may also be more related to how the delivery was managed than to genital cutting, and so it is important to include such factors in studies²⁸.

²⁶ Jones, H. *et al.* 1999. "Female genital cutting practices in Burkina Faso and Mali and their negative health outcomes", *Studies in Family Planning*, 30(3): 219-230.

²⁷ For the only known example of a community-based survey of FGC-related health complications, see Morison *et al. op cit.*

²⁸ Slanger, T., R. Snow and F. Okonofua. 2002. "The impact of female genital cutting on first delivery in southwest Nigeria, *Studies in Family Planning*, 33(2): 173-184.
Larsen, U. and F. Okonofua *op cit.*

Samson Radeny highlighted the many challenges in measuring the psychosocial and sexual complications associated with genital cutting (see Box 10). He described some of the attempts to measure these effects, including confidential self-reporting in clinical contexts. PATH has tried to establish the proportion of women reporting psychosocial and sexual effects, but most of these attempts have not been successful, especially in qualitative studies.

Box 10: Problems with measuring psychosexual effects

- Culturally, it is improper for people to volunteer information regarding their sexuality.
- Psychosexual problems are complex, and may not be easily identified by clients. Some cultures have no words to express psychosocial problems.
- Most circumcised women do not have the opportunity to visit psychologists and therefore even if they had psychosexual problems, these will go undetected and unreported.
- There are many other potential causes of psychosexual problems facing women in Africa, such as sexual violence.

A study conducted by PATH and the Seventh Day Adventist Church in western Kenya²⁹ did not find any psychological trauma or anxiety associated with the cutting, but this may be due to the way the study was undertaken. Radeny asserted that it is probably easier for women to report sexual problems than psychosocial effects because of the conflict that may be present between social pressure to support the practice and the personal problems being experienced. He also called attention to the lack of interventions designed to explicitly to reduce these types of complications³⁰.

Whether self-reporting of FGC status is a valid measure appears to depend, therefore, on the context in which the questions are being asked.

If FGC is widespread, socially acceptable and

there are no well-publicised interventions causing people to question its acceptability and legality (as was the case in Egypt, the Gambia and Nigeria), then self-reporting is likely to be highly valid. If there are reasons why it would not be attractive for respondents to declare that they are cut (as was the case in Ghana and Kenya), then self-reported measures should be questioned and ways sought to confirm the results. Clearly it is not normally going to be possible to validate self-reporting through observation, but the suggestion was made that observations of actual status (and type of cut and complications, if appropriate) among a sample of clinic clients that is representative of the wider population should be considered when it is feasible and ethically acceptable to do so at low cost. While observing the physical status of adolescent girls is extremely difficult because they rarely attend clinics, such observations could possibly be done during routine antenatal care examinations for first time

²⁹ Mohamud, A., S. Abwao and E. Omwega. 1996. "Community perspectives on FGM: A qualitative research report in Nyamira", PATH: Nairobi, Kenya.

³⁰ WHO covers the identification and management of psychosocial and sexual problems following FGC in its recently published training materials: WHO. 2001. Female Genital Mutilation: A Teacher's Guide, WHO: Geneva, Switzerland.

deliveries, as most women in Africa still have their first delivery at a relatively young age and the vast majority of pregnant women attend for antenatal care. Alternatively, as was done in Kenya and Ghana, qualitative research methods should be used.

Understanding the intervention process

Developing indicators and collecting data to measure them are central components of any study to evaluate interventions. However, for behaviour change interventions, and particularly those that are community based, it is also critical that sufficient attention and resources be paid also to understanding the process through which the intervention is implemented. To do so effectively it is necessary to document the implementation process as it happens, including the challenges encountered at each stage and the lessons learnt. Most community-level interventions are never implemented as planned, and descriptions of how activities were actually undertaken in real-life situations is crucial if a successful intervention is to be replicated elsewhere. In particular, being able to understand how individuals and families reach decisions about changing their beliefs and behaviours, and especially which messages or information most influenced them, can be difficult to do solely through baseline and endline surveys.

Process documentation provides additional information that is not usually captured in baseline and endline surveys and helps explain how and why changes have occurred or not occurred). Planning for sufficient resources and identifying appropriate research methods for collecting this information are critical when initially designing a study. **Nafissatou Diop** highlighted the importance of documenting the process of implementing the Village Empowerment Programme, especially in Senegal where the programme was being implemented in 90 villages and with some variation between them. Moreover, because many communities have links through marital and other social relationships, to villages that are not in close proximity, the intervention also appears to be having impact beyond the study sites themselves. Consequently, it is proving essential to use ethnographic methods to understand the social networks and communication patterns so as to map and document these ‘ripple’ effects, whereby those living in the study villages can be having an influence in many other villages.

Collecting this type of information is particularly important when working in unusual or unique situations (such as refugee camps or with nomadic populations), as responses by such communities are often unpredictable. **Jane Chege** described experiences from the projects being implemented by CARE International in Ethiopia where the traditional ‘*dagu*’ communication system of the nomadic Afar community has strongly influenced the way in

which the messages from the IEC intervention are given and passed on to communities beyond the intervention sites. Conversely, the implementation of interventions in research field stations, such as Navrongo, are usually extremely well documented to such an extent that the population may suffer from being over-researched and sometimes respondents give the answers they think the researchers want to hear. As a result, although field stations are excellent for testing the effectiveness of medical interventions that are measured through physical indicators of bodily health, they may be too artificial a setting for testing interventions that are measured primarily through reporting of social and behavioural change.

Utilising research findings

The role of intervention research is to generate empirical findings that can guide the design and implementation of new approaches to addressing FGC, or the expansion of existing approaches. Presentations by **Hadi Eltahir** on USAID and **Inge Baumgarten** on GTZ described perspectives from the two leading donors FGC-related research. Although both agencies see research as an integral component of their strategy for contributing to the overall abandonment of the practice, each agency's perspective is somewhat different. USAID supports research to better understand the various interventions that appear to be gaining momentum, and is keen to document and assess how research findings have been translated into action, and how best to maximise collective efforts in combating the practice. GTZ puts particular emphasis on a participatory research approach being included in all its FGC programmes, such that research activities, and particularly findings from diagnostic and baseline research, help a manager to better plan, implement and evaluate their programme.

Another group of key stakeholders who can use research findings are national organizations that coordinate FGC-related activities within a country. Many countries now have such organizations and committees, which may be non-governmental or governmental, and **McAntony Agnes** described the National Focal Point (NFP), a programme within an NGO that fulfils this role in Kenya. Bodies such as these can achieve a great deal by linking, co-ordinating and encouraging collaborations between different players in the field of FGC. It is important to know what everyone is doing to maximise benefits from programmes and research and to avoid replication as well as build capacity. One of their primary functions is giving information to all organizations and individuals interested in FGC, and so they are a major audience for research results, and particularly those that are communicated in non-technical language and oriented to provide guidance for information and service programmes.

A number of issues arose from the discussions around utilisation of research on FGC interventions:

- Research data should be widely available and presented in a way that makes it easy to understand and apply the findings³¹. In particular, the conditions under which the intervention was implemented that led to its success should be clearly documented.
- Donors and international NGOs should make explicit efforts to go beyond communicating results describing what does or does not work and build capacity and provide assistance for local organizations interested in undertaking anti-FGC interventions. **Inge Baumgarten** gave an example of GTZ providing assistance through supporting an annual problem-oriented capacity building workshop on action research offered by the Centre International de Formation en Recherche Action (CIFRA) in Burkina Faso³². Participants at the course acquire skills necessary to investigating their own work environment, developing and implementing appropriate solutions and concepts for the problems identified, and assessing the impacts of their work. When, for example, even documenting an intervention can be a problem for an NGO due to a lack of skills, this problem-oriented research approach is of a high value.
- More needs to be known about the type of evidence and means of communication that are most convincing for programmes and donors. Discussions and negotiations with stakeholders are necessary when writing research proposals to assess the feasibility of the study, interest in the results, and to develop some practical guidelines on how results would be utilised.
- Are findings of ‘pre-post’ studies without control groups sufficiently valid to be used to guide programmes? Funding agencies and programme managers need guidance on interpreting data from operations research studies, and in particular, the validity of findings that each design can and cannot give.
- Funding agencies must be prepared to fund the level of evaluation that will give the “answers” needed, and this can often be much higher than is currently provided.
- Funding agencies and implementers need to be realistic about what can and cannot be changed within the timeframe of a project (i.e. what stage in the behaviour change process people can be moved to) and take this into account when developing objectives and indicators.

³¹ A network linking researchers and others interested in FGC-related research is being developed to promote the use of research findings to realize the end of genital cutting. The network, named INTACT will have a website at: <http://www.intact-network.net>.

³² See the following website for course details: <http://www.gtz.de/action-research/english/service/service.html#1>

Conclusions

As systematic efforts to develop and implement interventions that encourage the abandonment of FGC increase, so do the needs to ensure that their feasibility and effectiveness are known, and the way in which they function is understood. Operations research can provide these types of information, but to date few organizations working towards abandonment of the practice, including those undertaking FGC-related research, have embraced the contributions that experimentally-oriented research studies can add to understanding this particular type of behaviour change. As this workshop has shown, there are now sufficient experiences with FGC-related operations research to be able to start identifying some general issues that should be considered when designing and implementing such studies.

Several challenges remain, however. These include: ensuring that those implementing and funding anti-FGC interventions appreciate how an operations research approach can benefit their efforts; building organizational and individual capacity to undertake operations research on anti-FGC interventions as well as to use the results from such studies; and communicating widely the findings from operations research so that future activities are evidence-based. It is hoped that the deliberations at this workshop will stimulate interest both in undertaking operations research more routinely when programming anti-FGC activities and in furthering the development and application of research methods appropriate for this subject.

Presentations

Behaviour change models and FGC interventions

Dr Karin Ringheim: Models of Behaviours Change and Intervention Research

Dr Jane Chege: FGC Operations Research: Challenges in CARE's Multi-Country FGC Abandonment Project in Ethiopia, Kenya and Sudan.

Dr Jane Chege: Research Design Challenges in Assessing the Impact of MYWO/PATH FGM Intervention Project

Mr Samson Radeny: Accounting for the role of the church: incorporate or control in the study design?

Ms Salma Elbeblawi: Assessing community based approaches to eliminate FGC in Egypt.

Dr Cheick Touré: Preparing service providers to manage FGC complications in Mali.

Dr Nafissatou Diop: Does Changing Reproductive Health and Human Rights Knowledge in Individuals Lead to Changing FGC Behaviours?

Jane Kamau: Preconditions for Alternative Rite of Passage: Experience from Trans Mara District

Designing intervention research studies

Dr Ian Askew: Overview of FGC intervention research

Dr Inge Baumgarten: Promotion of initiatives to end Female Genital Mutilation (GTZ)

Managing unanticipated outcomes

Dr Jane Chege: Controlling or managing preceding / external / unanticipated outcomes

Dr Karin Ringheim: Perceived Positive and Negative Effects of Trends towards Medicalization of FGC

Ms Elizabeth Jackson: Challenges and Benefits of Longitudinal FGC Data Collection

Ms Evelyne Sakeah: Appraisal of Participatory Designed FGC Intervention Strategies

Ms Jane Kamau: Implications of a Presidential Declaration on FGM Study, Koibatek District, Rift Valley Province

Dr Dupe Oduwole: Assessing FGC interventions among populations already exposed to other sensitisation or developmental interventions

Designing appropriate interventions

Dr Dupe Oduwole: How does using PLA techniques to generate behaviour change increase their likelihood of success?

Ms Evelyne Sakeah: Participatory Approach for Designing an FGC Intervention

Indicators for monitoring and evaluating behaviour change

Dr Linda Morison: Indicators for evaluating interventions against FGC

Dr Karin Ringheim: Capturing changes in age-specific FGM rates: problems of incomplete data and early age of circumcision.

Ms Elizabeth Jackson: Modelling and Sampling in a cohort study

Dr Nafissatou Diop: Measurement issues

Mr. Samson Radeny: Documenting and evaluating the impact of messages concerning human rights and bodily integrity

Dr Cheick Touré: Evaluation of Providers' Performance in Identification and Management of FGC Complications in Mali

Sampling

Dr Jane Chege: sample size calculations for hypothesis-testing and descriptive surveys: what is a significant difference?

Collecting data to measure indicators

Ms Elizabeth Jackson: Measurement Issues: Denial

Dr Ian Askew: Measuring gynaecological, obstetrical and psychological consequences of FGC

Mr Samson Radeny: Measuring psychosocial and sexual repercussions of FGC

Understanding the intervention process

Dr Nafissatou Diop: Understanding and documenting an intervention: Importance of a qualitative perspective

Dr Jane Chege: Challenges in Maintaining a Quasi-experimental Design in Nomadic Populations

Utilising research findings

Dr Hadi Eltahir: Critical gaps in FGC programme strategies and interventions

Dr Inge Baumgarten: Action Research: Combining research with practical programme needs

Ms Agnes McAnthony: National Focal Point for the Eradication of Female Genital Mutilation in Kenya

FOR MORE INFORMATION, CONTACT:

FRONTIERS IN REPRODUCTIVE HEALTH
Population Council
4301 Connecticut Avenue, N.W.
Suite 280
Washington, D.C. 20008
USA

Telephone: 202-237-9400
Facsimile: 202-237-8410
E-mail: frontiers@pcdc.org
Website: www.popcouncil.org

AFRICA
Population Council Regional Office
P.O. Box 17643
Nairobi
Kenya

Telephone: 254-2-2713480/1/2/3
Facsimile: 254-2-2713479
E-mail: publications@pcnairobi.org

ASIA AND THE NEAR EAST
Population Council Regional Office
Ground Floor, Zone 5A
India Habitat Center
Lodi Road
New Delhi 110 003
India

Telephone: 91-11-461-0913
Facsimile: 91-11-464\2903
E-mail: frontiers@pcindia.org

LATIN AMERICA AND THE CARIBBEAN
Population Council Regional Office
Escondida 110
Villa Coyoacan
04000 Mexico, D.F.
Mexico

Telephone: 52-5-659-8537
Facsimile: 52-5-554-1226
E-mail: disemina@popcouncil.org.mx



The Population Council is an international, nonprofit, nongovernmental institution that seeks to improve the well-being and reproductive health of current and future generations around the world and to help achieve a humane, equitable, and sustainable balance between people and resources. The Council conducts biomedical, social science, and public health research and helps build research capacities in developing countries. Established in 1952, the Council is governed by an international board of trustees. Its New York headquarters supports global network of regional and country offices.



FRONTIERS is funded by the Office of Population of the UNITED STATES AGENCY FOR INTERNATIONAL DEVELOPMENT (USAID) under the terms of Cooperative Agreement Number HRN-A-00-98-00012-00