Campbell Systematic Reviews
2012:9
First published: 28 June, 2012
Last updated: 25 April, 2012
Search executed: 03 March, 2011

Interventions to reduce the prevalence of female genital mutilation/cutting in African countries

Rigmor C Berg, Eva Denison

THE CAMPBELL COLLABORATION

Colophon

Title	Interventions to reduce the prevalence of female genital mutilation/cutting in African countries
Institution	The Campbell Collaboration
Authors	Berg, Rigmor C Denison, Eva
DOI	10.4073/csr.2012.9
No. of pages	156
Last updated	25 April, 2012
Citation	Berg RC, Denison E. Interventions to reduce the prevalence of female genital mutilation/cutting in African countries. Campbell Systematic Reviews 2012:9 DOI: 10.4073/csr.2012.9

Contributions	The authors are grateful for outstanding support from Sari Ormstad and Jan Odgaard-Jensen: The search strategy was developed and executed by Sari Ormstad (SO) and Jan Odgaard-Jensen (JOJ) assisted with statistical analyses. Rigmor C Berg (RB) and Eva Denison (ED) selected literature, performed methodological quality assessment, and extracted data. Berg synthesised the data and wrote this review. Rigmor C Berg will be responsible for updating this review as additional evidence accumulates and as funding becomes available.
Editors for this review	Editor: William Turner Managing editors: Jane Dennis, Hugh Waddington, Martina Vojtkova
Support/Funding	3ie. The authors are indepted to 3ie for both financial support and intellectual encouragements.
Potential Conflicts of Interest	The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

Corresponding author	Rigmor C Berg Norwegian Knowledge Centre for the Health Services P.O. Box 7004 St. Olavs plass 0130 Oslo, Norway E-mail: rigmor.berg@nokc.no

Campbell Systematic Reviews

The Campbell Collaboration (C2) was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. C2 offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of C2's editors, librarians, methodologists and external peer-reviewers contribute.

The Campbell Collaboration
P.O. Box 7004 St. Olavs plass
0130 Oslo, Norway
www.campbellcollaboration.org

Table of contents

Key messages

This systematic review aimed to answer the following research questions:

1. What is the effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting (FGM/C) compared to no or other active intervention?
2. How do factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions designed to reduce the prevalence of FGM/C?

We used an integrative evidence approach, whereby analyses of effectiveness data and context data were completed in separate streams, but where we in the final step integrated the results from the two sets of data in a realist synthesis approach.

We included eight effectiveness studies (research question 1). All employed a controlled before-and-after study design. The quality assessment resulted in a final decision of weak study quality for all eight studies, which involved 7,042 participants residing in seven different African countries: Burkina Faso, Egypt, Ethiopia, Somalia/Kenya, Mali, Nigeria, and Senegal. We could perform four meta-analyses but there was doubt about the validity of all results. The results showed that the effectiveness of the included interventions was limited, although they pointed to potential advantageous developments, such as positive changes in attitudes and knowledge regarding FGM/C, as a result of the FGM/C abandonment interventions.

We included 27 context studies (research question 2) from the countries where an effectiveness study had taken place (we did not identify any studies from Ethiopia). The synthesis of context studies showed that the factors related to the continuance and discontinuance of FGM/C varied across contexts, but the main factors that supported FGM/C were tradition, religion, and reduction of women's sexual desire. The main factors that hindered FGM/C were medical complications and prevention of sexual satisfaction.

Our ability to conclude regarding both the question of effectiveness of FGM/C abandonment interventions and how factors related to FGM/C help explain the effectiveness of interventions was hampered by a general lack of information. However, through the realist synthesis we identified that all of the interventions were based on a theory that dissemination of information improves cognitions about FGM/C, but the interventions' success was contingent upon a range of contextual factors.

Executive summary/Abstract

BACKGROUND

Female Genital Mutilation/ Cutting (FGM/C) is a traditional practice that involves partial or total removal of the external female genitalia or other injury to the female genital organs for non-medical reasons. FGM/C is mainly practiced on prepubescent girls in 28 countries in Africa. Recent figures for African countries show a prevalence of FGM/C of more than 70% in Burkina Faso, Djibouti, Egypt, Eritrea, Ethiopia, Guinea, Mali, Mauritania, Northern Sudan, Sierra Leone, and Somalia. However, there is considerable variation in prevalence between and within countries, reflecting ethnicity and tradition. The practice is also rooted in religio-social beliefs within a frame of psycho-sexual and personal reasons such as control of women's sexuality and family honour, which is enforced by community mechanisms. Girls exposed to FGM/C are at risk of immediate physical consequences, such as severe pain, bleeding, and shock, difficulty in passing urine and faeces, and infections. Long term consequences can include chronic pain and infections. For many girls and women, undergoing FGM/C is a traumatic experience that may adversely affect their mental health.

OBJECTIVES

We aimed to review the empirical research on the effectiveness of interventions to reduce the prevalence of FGM/C in African countries (question 1), and the empirical research on contextual factors that may help explain the effectiveness, or lack thereof, of such interventions (question 2).

SEARCH STRATEGY

The primary method of study identification was electronic searches in 13 international databases up to March 2011: African Index Medicus, Anthropology Plus, British Nursing Index and Archive, the Cochrane Library (CENTRAL, CDR, DARE), EMBASE, EPOC, MEDLINE, PILOTS, POPLINE, PsychINFO, Social Services Abstracts, Sociological Abstracts, and WHOLIS. Databases for grey

literature included the Demographic and Health Surveys, British Library for Development Studies, IDEAS, JOLIS, Google Scholar, and Google. The electronic searches were supplemented with searches in databases of six international organizations that are engaged in projects regarding FGM/C, communication with FGM/C experts, forward citation tracking through ISI Web of Knowledge, and bibliographic back-referencing. Lastly, we conducted a manual search of the journal Social Science & Medicine covering the years 1993 to 2003.

SELECTION CRITERIA

The inclusion criteria for question 1 were: Population: girls and/or young women at risk of FGM/C and other members of communities practicing FGM/C. Intervention: any intervention intended to prevent, or reduce the prevalence of, FGM/C. Comparison: no FGM/C intervention, wait list, or other active FGM/C intervention. Outcomes: rates of FGM/C as well as behaviours, intentions, attitudes, beliefs, and knowledge of FGM/C, and awareness of rights. Study designs: randomized controlled trials, quasi-randomized trials, controlled before-and-after studies, and interrupted time series designs.

The inclusion criteria for question 2 were: Population: members of communities practicing FGM/C. Interest: factors related to the continuance and discontinuance of FGM/C. Context: geographical context was African countries in which controlled studies of interventions to reduce the prevalence of FGM/C had been carried out. Historical context meant that the context studies had to have collected data no more than five years prior to the intervention to be relevant. Study designs: cross-sectional quantitative study designs, qualitative study designs, or a combination of the two (mixed-methods studies).

DATA COLLECTION AND ANALYSIS

The data collection and analysis processes proceeded in several steps. We used an integrative evidence approach, whereby data extraction and analyses of effectiveness data and context data were completed in separate streams. Two reviewers extracted data from the included sources using a pre-designed data extraction form for each of the two review questions. The same reviewers assessed the quality of studies, using appropriate checklists for included study designs.

For effectiveness data, we estimated effects of interventions by the adjusted absolute risk difference and the relative risk and 95% confidence interval. For continuous outcomes, we calculated mean difference and 95% confidence interval. We used Mantel-Haenszel random effects meta-analysis for dichotomous outcomes and inverse-variance random effects meta-analysis for continuous outcomes. Analysis of context data was based on published examples and guidelines from the Evidence for Policy and Practice Information and Co-ordinating Centre. For quantitative context

studies we used a generic inverse variance approach. We had planned to use thematic analysis for qualitative context data but only one qualitative study was included. In the final step of the integrative evidence approach, the results from the effectiveness studies were integrated with the results from the context studies in a realist synthesis. This realist synthesis approach addressed context-mechanisms-outcome configurations that underlie interventions.

RESULTS

We included eight effectiveness studies (research question 1) and 27 context studies (research question 2). Regarding the effectiveness studies, all employed a controlled before-and-after study design. The quality assessment resulted in a final decision of weak study quality for all eight studies, which involved 7,042 participants residing in seven different African countries: Burkina Faso, Egypt, Ethiopia, Somalia/Kenya, Mali, Nigeria, and Senegal. We could perform four meta-analyses, each with two studies synthesised, but there was doubt about the validity of all results. The results showed that the effectiveness of the included interventions was limited but they pointed to potential advantageous developments as a result of the FGM/C abandonment interventions.

Once we had identified that controlled effectiveness studies had taken place in seven contexts, we searched for and included a total of 27 eligible studies from these contexts (we did not identify any studies from Ethiopia). We judged that nine of the context studies had high methodological quality, 12 had moderate and six had low methodological quality. The synthesis of context studies showed that the factors related to the continuance and discontinuance of FGM/C varied across contexts, but the main factors that supported FGM/C were tradition, religion, and reduction of women's sexual desire. The main factors that hindered FGM/C were medical complications and prevention of sexual satisfaction. The extent to which we could conclude regarding how factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions was limited. However, based on the context-mechanism-outcome configurations we identified that all of the interventions were based on a theory that dissemination of information improves cognitions about FGM/C, but the interventions' success was contingent upon a range of contextual factors. For example, in contexts where FGM/C and Islam were closely related, the failure to involve religious leaders and base the program on the beneficiary community's needs and wants triggered low attendance and program drop-out.

AUTHORS' CONCLUSIONS

Our ability to conclude regarding both the question of effectiveness of FGM/C abandonment interventions and how factors related to FGM/C help explain the

effectiveness of interventions was hampered by a general lack of information. The findings show that much work remains to be conducted regarding the evaluation of FGM/C abandonment efforts. There is a need to conduct methodologically rigorous intervention evaluations. Such studies should address the local communities' enforcement systems that support FGM/C and be based on a sound theory of behaviour change.

1 Background

1.1 DEFINITION OF FGM/C

Female genital mutilation/cutting (FGM/C) [1] is a traditional practice that involves the partial or total removal or other injury to the female genital organs for non-medical reasons (WHO, 2008).

The WHO's (1997) current classification describes four types of FGM/C: Type 1 (clitoridectomy) involves partial or total removal of the clitoris and/or the prepuce. Type 2 (excision) involves partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora. Type 3 (infibulations) involves narrowing of the vaginal orifice with creation of a covering seal by cutting and appositioning the labia minora and/or the labia majora, with or without excision of the clitoris. Infibulation is considered the most invasive type of FGM/C. Defibulation, opening of the covering seal, is often necessary prior to childbirth. Reinfibulation refers to the recreation of an infibulation after defibulation. Type 4 (other), involves all other harmful procedures to the female genitalia for non-medical purposes, for example: pricking, piercing, incising, scraping and cauterizing (WHO, 1997). An explanation of terms is found in appendix 1.

1.2 PREVALENCE OF FGM/C

FGM/C is practised in more than 28 countries in Africa, usually on prepubescent girls, and in some countries in the Middle East and Asia. FGM/C is also practised by immigrant communities in a number of other countries, including Australia, Canada, France, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States (HRP, 2006).

Recent figures for African countries show a prevalence of FGM/C of more than 70% in Burkina Faso, Djibouti, Egypt, Eritrea, Ethiopia, Guinea, Mali, Mauritania, Northern Sudan, Sierra Leone, and Somalia (Yoder & Kahn, 2008). However, there is considerable variation in prevalence between and within countries, reflecting ethnicity and tradition. Therefore, UNICEF (2005a) has proposed that countries be

[1] We adopt the official terminology used by UNICEF and UNFPA 'female genital mutilation/cutting' (FGM/C) throughout this report. When quoting published studies we sometimes adopt the terminology used by the authors.

categorised in three groups according to FGM/C prevalence rates: *Group 1*, 80% or higher prevalence, e.g. Ethiopia and Somalia, *Group 2*, 25-79% prevalence, e.g. Senegal and Kenya, and *Group 3*, 1-24% prevalence, e.g. Nigeria (table 1).

Table 1. Prevalence of FGM/C between and within selected countries

Country	Total prevalence[1]	2 lowest [2]	2 highest[2]	Group
Burkina Faso	76.6	41.5	86.9	1
Egypt	97.0	-	-	1
Ethiopia	79.9	0.8	100	1
Kenya	32.2	0.7	96.3	2
Mali	91.6	53.1	98.0	1
Nigeria	19.0	0.5	52.9	3
Senegal	28.2	1.6	78.2	2
Somalia[3]	88.0[3]	-	-	1

Legend: 1=data from UNICEF (2005a). 2="2 lowest" refers to the two ethnic groups in the country with the lowest FGM/C prevalence; "2 highest" refers to the two ethnic groups in the country with the highest FGM/C prevalence. 3=data from Yoder & Kahn (2008).

The table shows that the wide range in FGM/C distribution within a country is closely associated with ethnicity. The practice is intertwined with ethnic identity, serving as an ethnic marker throughout the lifespan (UNICEF, 2005a). The practice is also rooted in religio-social beliefs within a frame of psycho-sexual and personal reasons such as control of women's sexuality and family honour, which is enforced by community mechanisms (WHO, 1999). While reasons for the practice vary across cultural groups, religious reasons rest on the belief that it is a religious requirement and socio-economic reasons include beliefs that FGM/C is a prerequisite for marriage or an economic necessity in cases where women are largely dependent on men (UNFPA, 2007).

1.3 CONSEQUENCES OF FGM/C

Girls exposed to FGM/C are at risk of immediate physical consequences, such as severe pain, bleeding, and shock, difficulty in passing urine and faeces, and infections. Long term consequences can include chronic pain and infections (WHO, 2008). A review of the health complications of FGM/C (WHO, 2000) identified a range of obstetrical problems, the most common being prolonged labour and/or obstruction, episiotomies and perineal tears, post partum haemorrhage, and maternal and foetal death. A recent study investigating 28,393 women attending obstetric centres in several African countries (WHO study group, 2006) concluded that women with FGM/C were significantly more likely than those without to have adverse obstetric outcomes such as a caesarean, postpartum blood loss ≥500 mL, extended maternal hospital stay, birth weight <2500 g, infant resuscitation, and

inpatient perinatal death. The authors also concluded that the risks seemed to be greater with more extensive FGM/C. More recently, a systematic review on the sexual consequences of FGM/C determined that women with FGM/C were twice as likely not to experience sexual desire, 1.5 times more likely to have pain during intercourse, and they experiences les sexual satisfaction (Berg & Denison, 2011).

For many girls and women, undergoing FGM/C is a traumatic experience that may adversely affect their mental health. In fact, several psychological and psychosomatic disorders such as disordered eating and sleeping habits have been attributed to FGM/C (HRP, 2006). There are also reports of posttraumatic stress disorder, anxiety, and depression associated with FGM/C (WHO, 2008). Data from a systematic review of the psychological consequences following FGM/C showed that women with FGM/C may be more likely to experience psychological disturbances, including anxiety, somatisation, low self-esteem, and to have a psychiatric diagnosis (Berg et al., 2010a).

Lastly, given FGM/C is a deeply entrenched tradition among some ethnic groups it carries consequences both when it is and when it is not practised. When girls and families conform to the practice they acquire social status, respect, and community membership (UNICEF, 2005b). In some societies, the link between FGM/C and value is explicit: girls who undergo FGM/C often receive rewards in the form of celebrations and gifts, and the bride price for a girl who has been cut is much higher than that for one who has not (Wheeler, 2003). Conversely, failure to conform can lead to difficulty in finding a husband for the girl, shame, stigmatization, as well as loss of social status, honour and protection, resulting in the family's social exclusion in the community (UNICEF, 2005b).

1.4 INTERVENTIONS TO REDUCE THE PREVALENCE OF FGM/C

Efforts to abandon the practice of FGM/C in Africa have used several different approaches. These approaches include those based on human rights frameworks, legal mechanisms, health risks, alternative rites, positive deviance, training health workers as change agents, training and converting circumcisers, and the use of comprehensive social development processes. Interventions based on these approaches have targeted stakeholders at individual, interpersonal, community, and national levels (Muteshi & Sass, 2005).

In 2007, the Population Reference Bureau (PRB) published their results of an extensive survey of current intervention projects taking place in African countries (Feldman-Jacobs & Ryniak, 2007). In total, the survey identified 92 projects, 27 of which were evaluated, mostly by observational designs. Only four of the 27 evaluated projects (15%) used a controlled before-and-after design. While contributing valuable understanding about the range of interventions initiated to reduce the

prevalence of FGM/C, this was not a systematic review and it did not reach any conclusions about the effectiveness of interventions.

More recently, the authors of the present systematic review specifically examined the effectiveness of interventions to reduce the prevalence of FGM/C in a systematic review (Denison et al., 2009). Through our literature search of February 2009 we identified a total of seven controlled studies, six of which could be obtained in full text. All six studies were controlled before-and-after studies carried out in African countries. In contrast to the PRB overview (Feldman-Jacobs & Ryniak, 2007), we included only controlled studies, i.e. studies with reference to a non-intervention comparison group, and we concluded that while the evidence base was insufficient to draw definite conclusions, there are possible advantageous developments as a result of these anti-FGM/C efforts. Notably, our review highlighted the uncertainties regarding relevance of the interventions (e.g. regarding objectives, intervention targets, activities). That is, since it was not a focus of the systematic review, we were unable to provide any assessment of the degree to which the interventions were appropriate responses to the populations' needs with respect to FGM/C, including the degree to which factors that contribute to the perpetuation of the practice were taken into account in the interventions. It is apparent that the degree of relevance of the intervention exerts a considerable influence on an intervention's effectiveness in reaching its designated goals, and may to a large extent help explain variation in behavioural and other outcomes among members of groups.

In sum, two recent publications have examined aspects of interventions designed to reduce the prevalence of FGM/C. One is a systematic review examining the effectiveness of interventions designed to reduce the prevalence of FGM/C (Denison et al., 2009). The present systematic review follows the same standard steps as far as systematically reviewing the evidence. However, the literature search is updated and expanded. Moreover, behaviour change techniques are identified and the effectiveness of interventions within a perspective of context is assessed and a realist synthesis carried out, which allows an examination of factors that facilitate and hamper the success of interventions.

1.4.1 Contextual factors related to the continuance or discontinuance of FGM/C

FGM/C is a long-standing tradition that has become inseparable from ethnic and social identity among many groups (UNICEF, 2005b). Disaggregation of data from the Demographic and Health Surveys (DHS) shows that the practice of FGM/C varies considerably by demographic variables such as age, urban-rural residence, and region or province, and also by variables such as education, ethnicity, and religion (Yoder et al., 2004). Further analysis of DHS data by UNICEF (2005a) suggests that educational attainment, a woman's own circumcision status, and ethnicity have the greatest influence in explaining support or opposition to the practice. Thus, programmes designed to reduce the prevalence of FGM/C should be country specific and adapted to reflect regional, ethnic, and socio-economic

variances while also taking into account the diverse reasons why FGM/C is practised among a given ethnic or cultural group.

1.5 OBJECTIVES

In line with the goals of the Campbell Collaboration we aimed to review the empirical research on the effectiveness of interventions to reduce the prevalence of FGM/C, and the empirical research on contextual factors that may help explain the effectiveness, or lack thereof, of such interventions. Specifically, the review aimed to answer the following research questions:

1. What is the effectiveness of interventions designed to reduce the prevalence of FGM/C compared to no or other active intervention?
2. How do factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions designed to reduce the prevalence of FGM/C?

The review summarized data relating to 1) key intervention program features, targeted participants, main outcomes, and estimates of intervention effectiveness and 2) contextual factors related to the continuance and discontinuance of FGM/C in areas where an included intervention had taken place, such as demographic factors, the frequency and strength of various stakeholders' cognitions and behaviours related to FGM/C, and stakeholders' lived understanding of the persistence of the practice. This allowed us to assess not only effectiveness of interventions but also what facilitates the success of these interventions. Explaining interventions' success and failure involves exposing their theories of why they would work and their utility in different contexts. Examining relevance in particular – i.e. the extent to which intervention programs have heeded and built upon factors related to the continuance and discontinuance of FGM/C, the extent to which interventions have been provided to the most appropriate stakeholder groups, and which forces have been overlooked as critical program elements – allowed us to learn about the conditions that are necessary for the FGM/C abandonment programs to work.

2 Methods

We conducted the review according to the criteria that the Campbell Collaboration applies to systematic reviews (Petrosino et al., 2001), while also adhering to other systematic review guidelines (Higgins & Green, 2009; Petticrew & Roberts, 2006). Having completed the Campbell-style review, examining both effectiveness and factors related to FGM/C, we aimed to gain insights into the underlying mechanisms that underpinned prevention programs and contexts in which they were implemented. Thus, we performed a realist approach to each of the interventions in the review, applying standards as described by Pawson (2006) and Pawson and colleagues (Pawson et al., 2004; 2005).

2.1 INCLUSION AND EXCLUSION CRITERIA

2.1.1 Criteria for inclusion and exclusion in the review

Research question 1: What is the effectiveness of interventions designed to reduce the prevalence of FGM/C compared to no or any other intervention?

Eligible for inclusion:

- Population: Girls and/or young women at risk of FGM/C; members of communities practicing FGM/C such as women, men, traditional circumcisers, religious leaders, educators, community elders, youth, government officials, health workers.
- Intervention: Any intervention intended to prevent, or reduce the prevalence of, FGM/C including: Legislation against FGM/C; education about health risks associated with FGM/C; training health workers as change agents; training and converting circumcisers; alternative rites; positive deviance; comprehensive social development. There were neither restrictions regarding intervention setting nor implementation year(s).
- Comparison: No FGM/C intervention, wait list, or other active FGM/C intervention.
- Primary outcomes: Rates of FGM/C; behaviours related to FGM/C, such as encouraging others not to cut their daughters.

- Secondary outcomes: Intentions regarding FGM/C; attitudes towards FGM/C; beliefs related to FGM/C; knowledge of adverse effects of FGM/C; awareness of rights.
- Study designs: Randomized controlled trials, quasi-randomized trials, controlled before-and-after studies, and interrupted time series designs.
- Languages: Publications in all languages were accepted and publications in languages not mastered by the review team were translated into English when considered to meet inclusion criteria.

Research question 2: How do factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions designed to reduce the prevalence of FGM/C?

Eligible for inclusion:

- Population: Members of communities practicing FGM/C such as women, men, traditional circumcisers, religious leaders, educators, community elders, youth, government officials, health workers.
- Interest: Factors related to the continuance and discontinuance of FGM/C, such as demographic factors, the frequency and strength of various stakeholders' cognitions related to FGM/C, and stakeholders' lived understanding of the practice.
- Context: Geographical context was countries in which controlled studies of interventions to reduce the prevalence of FGM/C had been carried out. Historical context meant that the context studies had to have collected data no more than five years prior to the intervention started to be relevant. Data collected after the intervention were not accepted.
- Study designs: Cross-sectional quantitative study designs, qualitative study designs, or a combination of the two (mixed-methods studies). Specifically, we included, first, any type of cross-sectional study design reporting quantitative data. Second, qualitatively-based studies had to have used either individual interviews or focus group interviews to collect data about FGM/C and used qualitative data analysis methods, such as thematic analysis, to be eligible for inclusion.[2] Third, mixed-methods studies that incorporated both quantitative and qualitative components where the research design matched the nominated study designs were included. Both the quantitative and the qualitative components of the study were subjected to the same inclusion criteria as the mono-methods studies and the study was only included when the inclusion criteria were met.
- Study quality: Study quality was assessed by checklists appropriate to study design, i.e. cross-sectional and qualitative. Studies that did not meet any

[2] We define qualitative evidence as "papers had to report results of qualitative (i.e. text-based and interpretive) analysis based on qualitative methods of data collection" (Smith et al., 2005 p826).

quality criteria at all were excluded. Qualitative studies of the highest level of quality were given priority in the analysis of qualitative studies, and studies with weak methodological quality (as determined by quality assessment) contributed less to our conclusions.

- Languages: Publications in all languages were included and publications in languages not mastered by the review team were translated into English when considered to meet inclusion criteria.

2.2 SCREENING AND SELECTION OF LITERATURE

Selection of primary studies was based on the inclusion criteria described above. The Reference Manager database containing the search results was used to keep track of references identified through the electronic database search.

Screening of literature proceeded in two steps. First, for level 1 screening, two reviewers (RB and ED) independently performed an assessment of the identified records by reading the title, and when available, abstract. The pre-developed inclusion questions for level 1 were based on the inclusion criteria described above (questions available in the protocol, Denison et al., 2011). The reviewers answered each question 'Yes' (=promote), or 'Can't tell' (=promote), or 'No' (=exclude, do not promote). The reviewers then compared and discussed their judgments. Differences in opinion at level 1 screening were resolved by promoting the record to level 2 screening. Records that unmistakably failed to meet the inclusion criteria were excluded, such as editorials and commentaries. Records not excluded at level 1 were promoted to level 2 screening, and ordered in full text.

Second, for level 2 screening, the same two reviewers (RB and ED) independently evaluated the full text of each record promoted from level 1 screening for inclusion, in accordance with Cochrane guidelines. They again used a pre-developed inclusion form (available in Denison et al., 2011) based on the inclusion criteria described above. There was a separate set of screening questions for records describing the effect of interventions and records reporting on reasons for the perpetuation of FGM/C. The reviewers answered each question 'Yes' (=Include), 'No' (=Exclude) or 'Can't tell' (=Discuss). The reviewers then compared and discussed their assessments. Differences in assessment at level 2 screening were discussed until consensus was reached. If consensus could not be reached, a third review member would have been asked to resolve disagreements, but this was not necessary as RB and ED arrived at an agreement for all assessed studies. The reference was included when the reviewers agreed to score 'Yes' to all questions. The reference was excluded when the reviewers agreed to score 'No' to any one inclusion question. If either one of the reviewers scored 'Can't tell' to any one inclusion question, the question was resolved by re-reading of the text, discussion and consensus. The main reason for exclusion at this stage was recorded for each record, and a list of excluded records,

with reasons, was created. These steps are in accordance with the Cochrane Handbook (Higgins & Green, 2009)

In the interest of time, the reviewers were not at any screening level blinded to the authors or other information about the record when assessing the studies. When there was more than one record of the same study, we included all records meeting the inclusion criteria, but used the most relevant one, i.e. the publication containing the most complete data set, as the main record.

Once included, we grouped all included studies according to their methodological focus into three main study types: 1) effectiveness studies, 2) quantitative context studies, 3) qualitative context studies.

2.3 LITERATURE SEARCH STRATEGY

2.3.1 Searches in electronic databases

The primary method of study identification was electronic searches, as advised by the Cochrane Handbook (Higgins & Green, 2009). We searched systematically for relevant literature up to March 2011 in 13 international databases:

- African Index Medicus
- Anthropology Plus
- British Nursing Index and Archive
- The Cochrane Library: CENTRAL, CDR, DARE
- EMBASE
- EPOC
- MEDLINE
- PILOTS
- POPLINE
- PsychINFO
- Social Services Abstracts
- Sociological Abstracts
- WHOLIS.

Under the guidance of one reviewer (ED) a research librarian performed the searches using a strategy incorporating subject headings (e.g. MeSH terms in MEDLINE, see http://www.nlm.nih.gov/mesh/) and text words (in title and abstract) relating to FGM/C and the four classifications thereof. No method filters were applied as we were more concerned about sensitivity than specificity and prepared to screen a large number of references. We did not restrict our searches by country, language or year. The MEDLINE search strategy served as the model for the other database searches using appropriate controlled vocabulary as applicable. The search strategy for each database is shown in appendix 2.

We included searches in databases for 'grey' literature (defined here as "reports that are produced by all levels of government, academics, business and industry in print and electronic formats but that are not controlled by commercial publishers" (Higgins & Green 2009)): Demographic and Health Surveys , British Library for Development Studies, IDEAS, JOLIS, Google Scholar, and Google. Theses and dissertations were considered for inclusion on the same basis as peer-reviewed articles.

2.3.2 Additional search strategies

The electronic search was supplemented with searches in databases of six international organizations that are engaged in projects regarding FGM/C: Centre for Development and Population Activities, Population Council, Population Reference Bureau (PRB), The United Nations Children's Fund (UNICEF), The United Nations Population Fund, and The World Health Organization (WHO). Notably, POPLINE contains ministerial reports on FGM/C.

We also asked for suggestions for literature we could have missed from FGM/C experts, e.g. the research director of Population Council. Forward citation tracking was performed through ISI Web of Knowledge in order to identify further studies. We also performed bibliographic back-referencing to identify new leads. Lastly, we conducted a manual search of Social Science & Medicine covering the years 1993 to 2003. These eleven years include the time span for eligible context studies (as stated in the inclusion criteria for research question 2, to be relevant, data for context studies should be collected no more than five years prior to the intervention started and not after the intervention had ended).

2.4 DATA EXTRACTION

Two authors (RB and ED) independently extracted data from the published sources using a pre-designed data extraction form, as recommended in the Cochrane Handbook (Higgins & Green, 2009).

Research question 1. We extracted the following types of data from studies included for answering research question 1: publication identification details (author, year, and source), study design, study setting, population, intervention details, comparisons, outcomes, follow-up details, attrition and missing data details, outcome data (for dichotomous data: number of events, number of persons in the groups, and p-values; for continuous data: means, standard deviations, and p-values), and effect estimates. Intervention details were classified according to a taxonomy of behaviour change techniques used in interventions, as proposed by Abraham and Michie (2008). These theory-linked behaviour change techniques were used to characterise and differentiate between intervention components. The data extraction form for research question 1 can be found in Denison and colleagues (2011).

Research question 2. Data extracted from studies included for answering research question 2 were: publication identification details (author, year, and source), study design, study setting, population, quantitative or qualitative data related to the continuance and discontinuance of FGM/C, i.e. demographic factors, the frequency and strength of various stakeholders' cognitions and behaviours related to FGM/C, and stakeholders' lived understanding of the practice. With respect to qualitative-based studies, the reviewers read the texts independently of each other and extracted all text data related to views about FGM/C in light of stakeholders' reasoning for its continuance and discontinuance. Our extraction of data was inclusive (Briggs & Flemming, 2007). We copied all findings in the form of sentences, phrases or text units appearing to deal with reasons for FGM/C's continuance and discontinuance verbatim onto our pre-designed data extraction form. The data extraction form for research question 2 is available in Denison and colleagues (2011).

For the realist approach, which tied the effectiveness and context review parts together, outcome patterns were established at the conclusion of the effectiveness review (research question 1). Information concerning mechanisms that were assumed to underpin the intervention was collected from the effectiveness reports and their related publications. We searched these documents for and extracted all statements which addressed mechanism ('why') issues: why would the intervention work or why did it not work? We copied all findings in the form of sentences, phrases or text units appearing to deal with mechanisms verbatim onto a form. Concerning context, we considered the data on context provided in the effectiveness studies as well as the studies included for question 2.

After independent extraction of data, the two reviewers compared their extraction results, and resolved any discrepancies by consulting the relevant texts and discussing.

2.5 QUALITY ASSESSMENT

Two authors (RB and ED) independently assessed the quality of studies, using appropriate checklists for included study designs (see below). Next, they compared and discussed their assessments. If consensus could not be reached, a third person would have been asked to resolve disagreements, but this was not necessary.

Research question 1. Studies included for answering research question 1 were assessed by the Quality Assessment Tool for Quantitative Studies, developed by the McMaster University Effective Public Health Practice Project (Thomas, no date). The Cochrane Handbook recommends this tool for reviews in the public health domain (Higgins & Green, 2009: Chapter 21 Reviews in public health and health promotion). The tool is presented in appendix 4. It includes the following sections: selection bias, study design, confounders, blinding, data collection methods, withdrawals and drop-outs, intervention integrity, and analyses. The first six

sections are each given a rating of strong, moderate or weak according to pre-specified criteria. The last two sections include questions regarding intervention integrity and analyses but are not given a rating of strong, moderate or weak. The reviewers independently and later jointly assessed the studies with respect to each quality component. In accordance with the quality tool assessment instructions, a final decision of strong, moderate or weak methodological quality was agreed upon by the reviewers after discussing whether there was a discrepancy with respect to the component ratings.

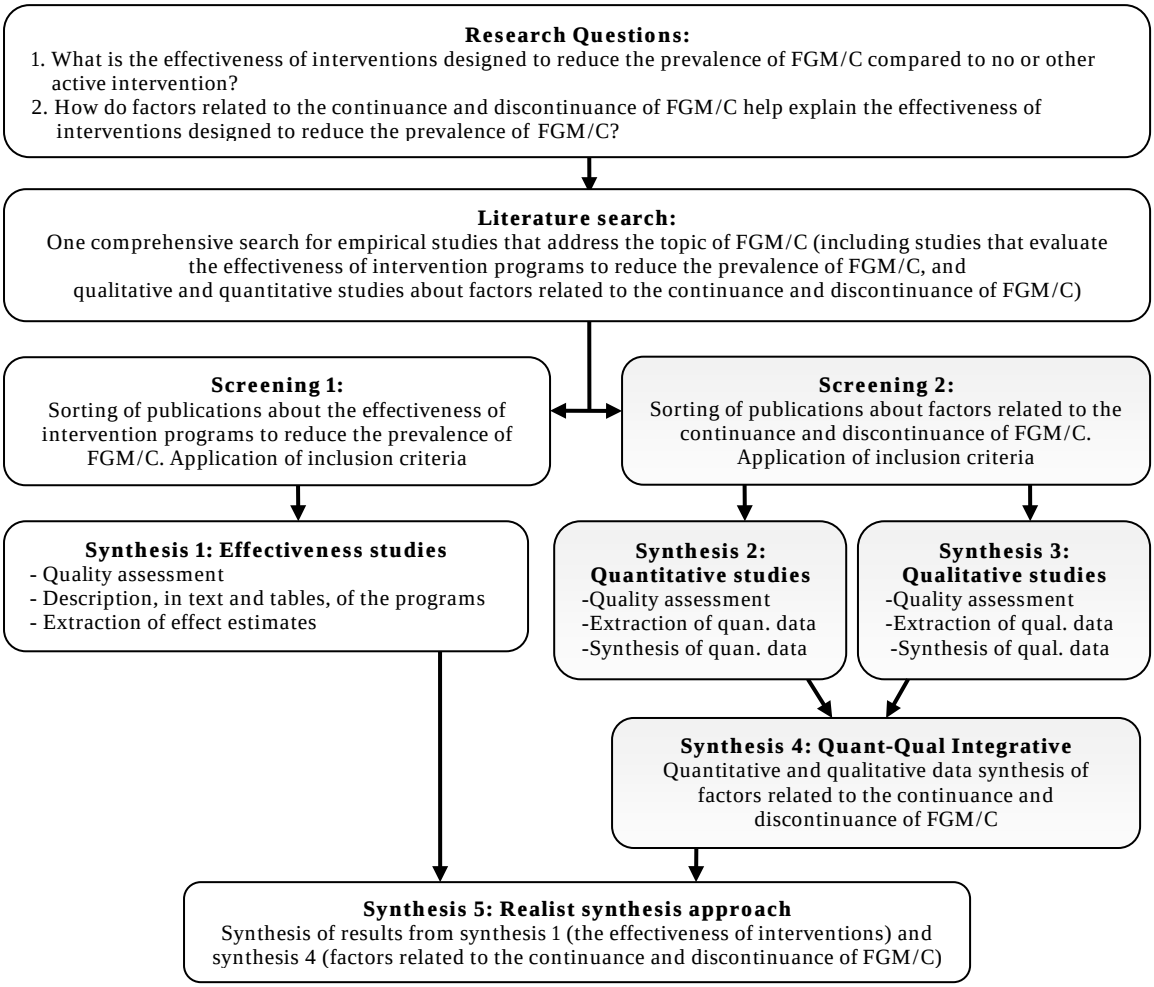
Research question 2. Quantitative studies included for answering research question 2 were assessed according to guidelines for evaluating cross-sectional studies (Boyle, 1998). The guidelines cover sampling-, measurement-, and analysis issues which have been compiled into a checklist comprising seven quality criteria by researchers at the Norwegian Knowledge Centre for the Health Services (appendix 4). The reviewers independently and later jointly assessed the studies with respect to the seven quality criteria. The reviewers agreed on a final decision of high, moderate or low methodological quality (or exclude), after discussing whether there was a discrepancy with respect to the seven criteria.

Qualitative studies included for answering research question 2 were assessed by a tool designed by the Critical Appraisal Skills Programme (CASP, 2006). The tool comprises 10 questions addressing the rigour, credibility, and relevance of qualitative studies. Two questions are screening questions to determine whether it is worth proceeding with the remaining questions. The remaining eight questions cover research design, sampling, data collection, reflexivity, ethical issues, data analysis, findings, and value of the research. The tool does not contain explicit guidance as how to judge the quality of studies. We applied the same procedure as for the assessment of quantitative cross-sectional studies. We do not have permission to reproduce the tool; it is therefore not included in the appendix.

2.6 DATA ANALYSES

The data analysis process proceeded in several steps. We used an integrative evidence approach, whereby data extraction and analyses of effectiveness data and context data were completed in separate streams (figure 1). We detail the process below.

Figure 1. Data synthesis approach



Given the separate streams, readers who are interested in only reading the results of the effectiveness review can read synthesis one, while those more interested in the contexts of FGM/C in the intervention areas can read synthesis two to four. The realist synthesis is detailed last, in synthesis five.

2.6.1 Analysis of effectiveness data

For the synthesis of effectiveness evidence (research question 1), we grouped the studies by intervention type. Because of the low number of studies identified, no sensitivity analyses were used to examine the stability of the effect estimates in relation to quality of studies. We described key intervention features, outcomes and effect estimates for each study in text and tables.

With respect to statistical analyses, we present dichotomous data for the outcomes listed in the inclusion criteria in results tables when pre- and post scores for both intervention and comparison groups were reported by study authors, allowing comparison. Based on recommendations by our statistical expert (JOJ), we estimated effects of interventions in two ways. First, we estimated effect by the adjusted absolute risk difference (ARD) in which the pre-post change score (in

percentage points) in the comparison group was subtracted from the pre-post change score (in percentage points) in the intervention group. Whether ARD is deemed to be large is a judgment whereby we also took baseline difference into account. Second, we estimated effect by the relative risk (RR) and 95% confidence interval (95%CI) based on post-intervention data. For continuous outcomes, we calculated mean difference (MD) and 95%CI. We adjusted the standard errors of the studies which compared clusters. The benefit of using ARD was that effect could be estimated even though prognostic similarity at baseline could not be assumed due to non-randomisation. Because none of the effect studies that we identified were randomised it was necessary to use this approach. Had randomised studies been identified in the update search (March 2011) and included in the review, we would have synthesised these separately. An obvious drawback to our analysis was that some element of judgment regarding baseline similarity was present when conclusions were drawn from the results. Based on suggestions by several agencies (PRB, 2008; UNICEF, 2005a; WHO, 2008), we used the three prognostic indicators education, prevalence of FGM/C, and religious affiliation to judge baseline similarity.

The a priori decision was made that, if studies were sufficiently similar, we would pool those that could be grouped together and use the statistical technique of meta-analysis to estimate effect, with RevMan v5.1 (Cochrane Collaboration meta-analysis software). To be pooled, the same outcome had to be assessed in a similar manner in similar populations across similar intervention studies. Baseline similarity of the outcome was a prerequisite. As is also standard, we used Mantel-Haenszel random effects meta-analysis for dichotomous outcomes and inverse-variance random effects meta-analysis for continuous outcomes. Compared to fixed effects meta-analysis, random effects meta-analysis is more conservative, which is reflected in the broader 95%CIs usually observed. We examined between-study heterogeneity with the Chi-square (Chi^2) and I-square (I^2) tests. We had planned to use subgroup analyses or moderator analyses to explore sources of heterogeneity, but because there were only two studies included in each meta-analysis it was not possible to perform moderator analyses.

2.6.2 Analysis of data from context studies

The contribution of quantitative and qualitative data describing contextual factors pertained primarily to the understanding of results from the intervention studies, and to help define interventions more precisely in relation to their contexts (research question 2).

Data extraction and analyses of quantitative and qualitative evidence were largely completed in separate streams (see figure 1 presented above). Data from cross-sectional survey studies (quantitative data) were combined with data from studies which examined various stakeholders' perspective of factors influencing the continuance and discontinuance of FGM/C (qualitative context studies). Our integrative evidence approach was largely based on published examples and guidelines from the Evidence for Policy and Practice Information and Co-ordinating

Centre (EPPI) (for worked examples see e.g., Harden et al., 2004; Shepherd et al., 2006; Thomas & Harden, 2008). It was the same approach as the one we used in Berg and colleagues (2010b). The synthesis was aggregative (Dixon-Woods et al., 2006) and focused on summarising data by pooling conceptually similar data from the quantitative studies and the qualitative studies. First, we analysed the two sets of evidence separately.

With respect to the quantitative studies (synthesis 2 in figure 1), we used a generic inverse variance approach (Higgins & Green, 2009) to synthesise the extracted quantitative data across studies. Further, for each study, we reviewed the units extracted and categorize them according to our predefined categories, such as expressed reasons for continuing FGM/C. We then determined the frequencies of these categories in order to create a ranked list of factors, ending up with one list for each country/local region. Thus, we were able to both synthesise data across studies (settings) and describe local variation.

With respect to qualitative evidence (synthesis 3 in figure 1), we planned to use a thematic analysis which drew on published EPPI studies (e.g., Harden et al., 2004; Shepherd et al., 2006). Thematic analysis involves identifying prominent or recurring themes in the literature and summarizing the findings of the different studies under thematic headings (Dixon-Woods, 2005). Further, we had planned to use results from the quantitative data set as organizing principles for the qualitative data analysis (synthesis 4 in figure 1), and given the small number of qualitative studies included (n=1), this occurred logically. Because only one qualitative study was included, we report all extracted results from this study. The qualitative results were subsumed under the quantitative results and the qualitative results extended the results from the quantitative analysis. For a full description of the planned thematic analysis, see Denison and colleagues' protocol (2011).

2.6.3 Realist synthesis approach

In the final step of the analysis (synthesis 5 in figure 1) the results from the effectiveness studies were integrated with the results from the context studies in a realist synthesis approach. The integration aimed to address context-mechanisms-outcomes (CMO) configurations that underlie interventions (Pawson, 2006; Pawson et al., 2004; Pawson et al., 2005).

Realist synthesis is concerned with review of complex social interventions (Pawson, 2006; Pawson et al., 2004; Pawson et al., 2005). It follows the same 'standard' sequence as systematic reviews, i.e. identifying the research question, searching for primary studies, quality appraisal, extracting the data, synthesising the data, and dissemination of the results. The basic components of a realist explanation are outcome patterns, generative mechanisms, and contextual conditions. Generative mechanisms are the engine behind behaviour (what is on offer in the program that may persuade participants to change) and the pivot around which realist synthesis revolves. The context is important because the action of mechanisms to some extent depends on the realities of the context in which they are used. Specifically, Pawson

explains that “interventions offer resources which trigger choice mechanisms (M) which are taken up selectively according to the characteristics and circumstances of subjects (C), resulting in a varied pattern of impact (O)” (Pawson, 2006 p25). For example, an abstinence-only program would likely evidence quite different trajectories in a rural, Catholic girls-only school and an inner-city, co-educational school because the students’ perceived threat and severity of STIs and unwanted pregnancies would be different. Similarly, oral pain medication will work quite differently in a person who is gluten intolerant, because uptake would be poor, versus one who is not. Pawson (2006) and colleagues (2005) further detail that the realist approach thus aims to differentiate program participants and their contexts in respect to how they might respond to the program. As a result, a realist synthesis attempts to explain how efficacy (outcome patterns) of an intervention varies depending on the particular configuration of its constituent mechanisms and contexts. The approach is hypothesis generating, the result of which leads to tentative recommendations meant to influence the design of new programs.

In our proposed synthesis (synthesis 5 in figure 1) the effect estimates extracted from the effectiveness studies provided the outcome patterns. Information concerning social and behavioural mechanisms that were assumed to be at work when the intervention was implemented was collected from the effectiveness reports and their related publications. As recommended by Pawson and colleagues (2005), we searched these articles for statements which addressed mechanism (‘why’) issues: why would the intervention work or why did it not work? We searched both for behavioural and social ‘cogs and wheels’ of the intervention and administrative and related mechanisms. These data were used to identify, or generate hypotheses about, the change theories underlying the interventions. Over multiple rounds of discussions regarding the mechanism statements, the two review authors arrived at a common understanding. Concerning context, we not only considered the data on context provided in the effectiveness studies, but completed a separate search and synthesis of context studies, depicted as synthesis 4 in figure 1. Recent examples of realist syntheses that served as models for our approach are Greenhalg and colleagues (2007), Dieleman and colleagues (2009), Kane and colleagues (2010), van der Knaap and colleagues (2008).

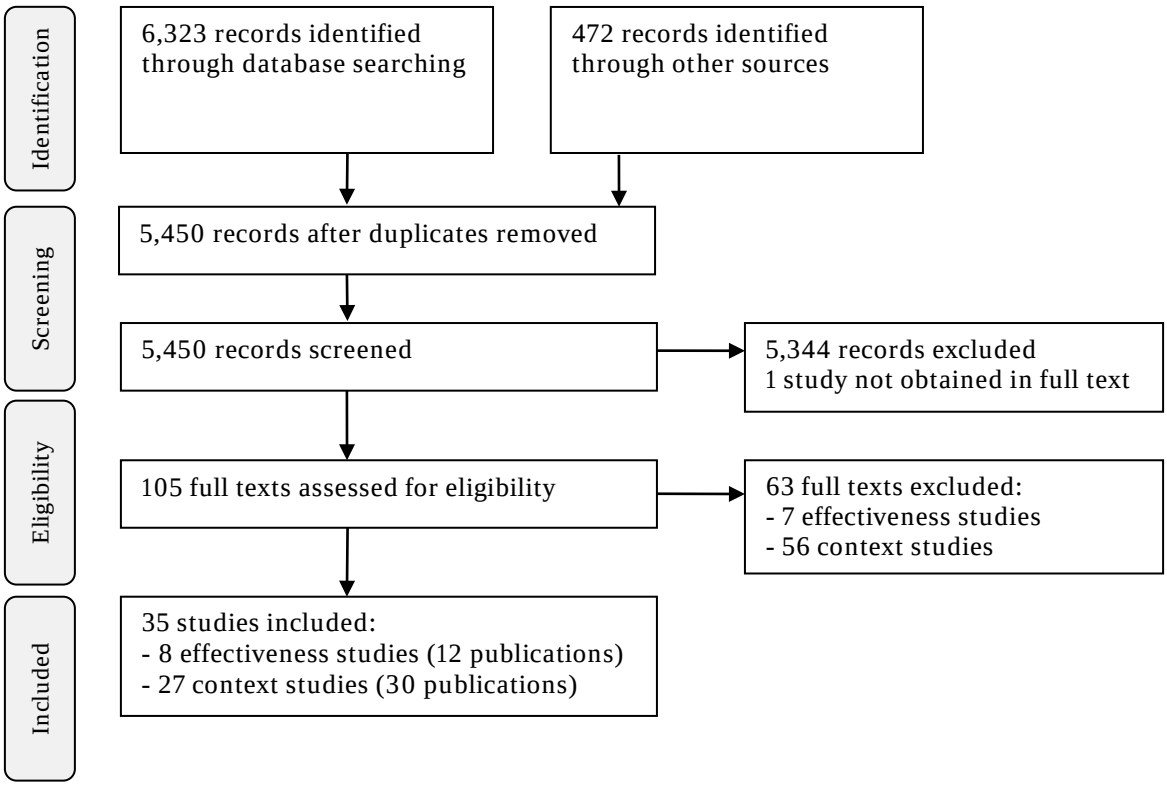
3 Results

3.1 RESULTS OF THE SEARCH

The electronic searches resulted in over 6,000 individual records and the manual search in 472 relevant records (figure 2). After removal of duplicates, the reviewers screened titles and abstracts, excluding 5,344 publications. Thus, 106 potentially relevant records were promoted to full text screening, but one potentially relevant study could not be obtained in full text (Akweongo et al., forthcoming).

Of the 105 studies read in full text, we excluded 63 publications (appendix 3) and included 35 studies. Eight of these were effectiveness studies (research question 1) while 27 were context studies (research question 2). Four publications in languages not mastered by the review team were translated into English.

Figure 2. PRISMA flow diagram of literature reviewing process



3.2 EFFECTIVENESS STUDIES

3.2.1 Quality assessment of effectiveness studies

Application of the McMaster quality assessment tool resulted in a final decision of weak study quality for all eight included effectiveness studies (appendix 4). All of the studies employed a controlled before-and-after study design. None were randomised. Six studies appear to have used cross-sectional and independent samples at baseline and endline (Babalola et al., 2006; Chege et al., 2004; Diop et al., 2004; Ouoba et al., 2004. We are uncertain about Easton et al., 2002). Six studies were dissimilar regarding prognostic factors in the intervention and comparison groups at baseline (Babalola et al., 2006; Chege et al., 2004; Diop et al., 1998; Easton et al., 2002; Ouoba et al., 2004). Data collection tools were neither shown to be valid nor reliable. Regarding intervention integrity, there was likely contamination due to population movements into and out of the intervention and comparison areas in three studies (Chege et al., 2004; Diop & Askew, 2009), and due to another FGM/C awareness-raising program in two intervention areas (Diop and Askew, 2009; Easton et al., 2002).

3.2.2 Synthesis 1: Descriptions and results of the interventions

Eight controlled before-and-after studies, described in 12 publications, met the inclusion criteria (Babalola et al., 2006; Chege et al., 2004; Diop et al., 1998; Diop et al., 2003; Diop et al., 2004; Diop et al., 2008a; Diop & Askew, 2009; Easton et al., 2002; Monkman et al., 2007; Mounir et al., 2003; Ouoba et al., 2004; UNICEF, 2008). They are briefly presented in table 2.

Table 2. Description of included effectiveness studies (N=8)

Author	Population	Intervention	Comparison	Outcomes
Diop 1998	N=108. Mali. Health personnel	Training, supervision	No intervention	Beliefs, attitudes, knowledge
Mounir 2003	N=682. Egypt. Female university students	Education	No intervention	Knowledge
Babalola 2006	N=957. Nigeria. Community members	Multimedia communication	No intervention	Intentions, beliefs, attitudes, knowledge
Chege 2004a	N=1,440. Somali refugees in Kenya	Outreach, advocacy	Education	Intentions, beliefs, attitudes, knowledge
Chege 2004b	N=819. Ethiopia. Community members	Outreach, advocacy	No intervention	Intentions, beliefs, attitudes, knowledge
Easton 2002	N=239. Mali. Community members	Tostan educ. program	No intervention	Beliefs
Diop 2004	N=1,332. Senegal. Community members	Tostan educ. program	No intervention	Prevalence, intentions, attitudes, beliefs, knowledge

Author	Population	Intervention	Comparison	Outcomes
Ouoba 2004	N=1,465. Burkina Faso. Community members	Tostan educ. program	No intervention	Prevalence, behaviors, intentions, attitudes, beliefs, knowledge

Note: The study in Kenya and the study in Ethiopia were reported in the same publication (Chege et al., 2004). Here, Chege 2004a refers to the intervention set in Kenya, Chege 2004b refers to the intervention set in Ethiopia.

Table 2 shows that the eight included studies involved a total of 7,042 participants at entry from different African countries: Burkina Faso, Egypt, Ethiopia, Somalia/Kenya, Mali, Nigeria, and Senegal. (The number of participants reported here are those directly receiving an intervention or serving as comparison and from whom baseline data were collected. The number of people receiving (elements of) the intervention may have been higher). With the exception of the study among Somalis, only one category of comparison was used: No intervention. None of the studies collected biological data. Rather, all information was self-reported prevalence of FGM/C, behaviours, intentions, beliefs/attitudes, and knowledge that was collected from the participants through face-to-face structured interviews or paper-and-pencil questionnaires.

Two interventions were delivered at the individual level (Diop et al., 1998; Mounir et al., 2003) while in the other six included studies, the intervention was provided to communities and involved a bundle of activities and behaviour change techniques. We identified five broad categories of interventions in the studies we reviewed. Training, which involved the provision of knowledge and skills-based training to health clinic staff and included practice sessions (Mounir et al., 2003). Formal classroom education with one-directional teaching to students but also some discussion and role play (Mounir et al., 2003). Media communication focusing on awareness and dialogue regarding FGM/C in the community (Babalola et al., 2006). The intervention in two areas consisted on outreach and advocacy (Chege et al ., 2004) and the last category of interventions was informal adult education (Easton et al., 2002; Diop et al., 2004; Ouoba et al., 2004). We present the results for each intervention below, grouped by category.

3.2.2.1 Training of health personnel in Mali

One study evaluated the effectiveness of training and supervising Mali health personnel on skills caring for women with FGM/C, various beliefs about FGM/C, and knowledge of FGM/C, such as complications following the practice (Diop et al., 1998). Health clinic staff (n=59) at eight sites were selected to receive the intervention while staff (n=49) at six health clinics in the same region served as the comparison group and received no training. The 108 study participants consisted of obstetricians, gynaecologists, and family planning providers from the districts of Bamako and Bla (Segou region, south-central Mali). The median age was 38 (range 15-52), 96% were Muslim and the majority ethnic group was Bambara (34%).

3.2.2.1.1 The intervention

The individually-based intervention lasted for two months and appears to have taken place in 1997. Health clinic staff at eight sites were selected to receive 1) three group training sessions about FGM/C, its health-related complications, and IEC (information, education, communication) activities to use for holding health talks at the clinics, and 2) supervision during clinic duty. Application of Abraham and Michie's (2008) taxonomy of behaviour change techniques showed that six techniques were used (techniques 1, 2, 5, 21, 26, 30):

- Provide information on consequences of behaviour in general
- Provide information on consequences of behaviour to the individual
- Goal setting (behaviour)
- Provide instruction on how to perform the behaviour
- Prompt practice
- Prompt identification as role model/ position advocate.

In appendix 5, we explain the judgements, including text from the publication (Diop et al., 1998), regarding the behaviour change techniques identified.

3.2.2.1.2 Mechanisms assumed to be at work

We searched and extracted information concerning mechanisms that were assumed to underpin the intervention from the effectiveness report (Diop et al., 1998) only, because we did not locate any other publications describing this intervention. There were few statements addressing mechanisms of the intervention, and only one behaviour change theory postulated, which was: Training leads to improved knowledge, attitudes, and skills. While the effectiveness report offered no further reasons of why the training would increase health staff's knowledge, attitudes and skills regarding FGM/C, it postulated one process factor about why it may not have worked optimally, which was that the time span used for training was too short (Diop et al., 1998). In appendix 6, we show the text from the publication (Diop et al., 1998), regarding the mechanism identified.

3.2.2.1.3 Results of the intervention

We extracted four outcomes (table 3). All were secondary outcomes. There was no significant difference between the groups regarding any of the four outcomes at endline.

Table 3. Study outcomes and effect estimates for intervention training health personnel in Mali

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
could name any type of FGM/C [1]	76	95	47	81	-15	1.12 (0.97, 1.30)
could name at least 3 types of long-term	50	72	61	73	10	0.99 (0.79, 1.26)

FGM/C complications [1]						
believed FGM/C pose no health risk if done in hygienic environment [1]	29	44		9	29	-5	1.54 (0.91, 2.60)
wished to play role in educating health clinic clients about FGM/C	93	91		89	86	-1	1.07 (0.93, 1.23)

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups.

3.2.2.2 Education of female students in Egypt

The educational intervention for female students at Alexandria University, Egypt, aimed to examine the effectiveness of two educational sessions on participants' beliefs and knowledge regarding reproductive health aspects, including FGM/C (Mounir et al., 2003). The study participants (n=682) were students living in university hostels (university on-campus residence halls). Students (n=354) living in one hostel (Ezbet Saad) were selected to receive the intervention while students (n=328) living in another hostel (El-Shatby) served as the comparison group and did not receive any intervention. The mean age was 19 and about half of the students (47%) were from low social class families.

3.2.2.2.1 The intervention

The study appears to have taken place in 2001. Students living in a university hostel were selected to receive information about reproductive health, such as FGM/C, through two 1-hour sessions of health talks, group discussion, role play, and use of educational aids. One behaviour change technique was used (technique 2): Provide information on consequences of behavior to the individual. In appendix 5, we explain the judgement, including text from the evaluation publication (Mounir et al., 2003), regarding the behaviour change technique identified.

3.2.2.2.2 Mechanisms assumed to be at work

We searched and extracted information concerning mechanisms that were assumed to underpin the intervention from Mounir and colleagues (2003). These are shown in appendix 6. Similar to the evaluation about health personnel in Mali, there were few statements addressing mechanisms of the intervention and the authors offered only one theory for the intervention: Education leads to improved knowledge and attitudes. The authors of the effectiveness report offered no comments regarding process factors that may have enhanced or reduced the effectiveness of the intervention.

3.2.2.2.3 Results of the intervention

There was only one outcome for which we could calculate an effect estimate. This was a secondary outcome concerning knowledge. The mean knowledge score about dangers of FGM/C increased 0.47 points in the intervention group and 0.03 points in the comparison group. The mean gain difference at endline was significantly different (MD= 0.44 points on a 0-3 scale, 95%CI= 0.14, 0.74).

3.2.2.3 Communication program in Nigeria

The intervention in Nigeria, which aimed towards "changing FGC-related attitudes and promoting the intention not to perform FGC", took place in the south-eastern part of the country (Babalola et al., 2006). Female and male community members (n=484) in three local government areas in Enugu state received the intervention while residents (n=473) of three local government areas in an adjacent state, Ebonyi, served as the comparison group. The participants from Ebonyi did not receive any intervention. The mean age of the participants was 34 (range 18-59), 66% had education less than secondary level, and 58% considered themselves Protestant.

3.2.2.3.1 The intervention

The communication intervention in Nigeria, called Ndukaku (Igbo word for 'health is better than wealth'), was delivered at three community levels: hamlet ('village') level, local government area level, and state level. It consisted of multimedia activities (e.g. newspaper columns, radio call-in shows), development of action plans to improve women's situation, and community meetings. The intervention appears to have lasted about 12 months and took place in 2003-2004. Application of the taxonomy of behaviour change techniques (Abraham & Michie, 2008) showed that five techniques were used (techniques 1, 2, 5, 6, 29):

- Provide information on consequences of behaviour in general
- Provide information on consequences of behaviour to the individual
- Goal setting (behaviour)
- Goal setting (outcome)
- Plan social support/ social change.

We explain the judgements in appendix 5.

3.2.2.3.2 Mechanisms assumed to be at work

There was one publication which described the Nigerian study and therefore served as our only source for theories postulated about the mechanisms underlying the intervention (Babalola et al., 2006). Details about the mechanisms are provided in appendix 6. We found that the study authors suggested that the change theory for the intervention was convention theory. According to this theory, conventions lie behind the stability of institutions and tradition, and can also explain rapid change. Specific to FGM/C, Mackie (1996) posits that FGM/C emerged as a strategy to secure marriage by signaling fidelity and that it spread to become a prerequisite for marriage. The theory predicts that change in the practice of FGM/C results from coordinated abandonment in intramarrying groups so as to preserve a marriage market for girls not cut. Babalola and colleagues (2006) further specified that the intervention program would lead to increased awareness, which would lead to self-examination of beliefs and values, which would trigger ways of thinking and value orientations. They also specified that the program would lead to dialogue and

group/social interactions and advocacy, which in turn would improve self-efficacy and perceived social support.

The effectiveness report further postulated that high degree of program exposure, mainly through the radio, improved FGM/C-related ideation, and that program exposure through both mass media and community activities affected change more so than exposure through either one alone. However, 36.6% of the respondents were not exposed to any program components.

3.2.2.3.3 Results of the intervention

The results for males and females were reported separately. We extracted six outcomes for female participants (table 4). Only one was a primary outcome (had encouraged someone not to perform FGM/C on daughter). Five outcomes were significantly different between the groups at endline, but there was a large baseline difference between the groups regarding three of these outcomes.

Table 4. Study outcomes and effect estimates for communication intervention in Nigeria (women)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
had encouraged someone not to perform FGM/C on daughter	16	24	11	9	10	2.68 (1.76, 4.08)*
no intention to perform FGM/C on daughter	59	76	64	67	14	1.13 (1.02, 1.26)*
did not believe there were benefits to FGM/C[1]	58	75	66	72	11	1.04 (0.95, 1.15)
disapproved of FGM/C[1]	63	88	71	73	23	1.21 (1.11, 1.13)
believed most community members favored discontinuation of FGM/C[1]	36	49	21	14	20	3.50 (2.58, 4.76)
believed had self-efficacy to resist spousal pressure to perform FGM/C on daughter[1]	57	72	40	42	13	1.71 (1.47, 1.99)

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups.*=RR significant at the p<.05 level and baseline similarity between groups.

We extracted six outcomes for male participants (table 5). Only one was a primary outcome (had encouraged someone not to perform FGM/C on daughter). Four outcomes were significantly different between the groups at endline, but there was a large baseline difference between the groups regarding two of these outcomes.

Table 5. Study outcomes and effect estimates for communication intervention in Nigeria (men)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
had encouraged someone not to perform	12	14	10	12	0	1.19 (0.71, 2.01)

FGM/C on daughter						
no intention to perform FGM/C on daughter [1]	53	73	64	66	18	1.11 (0.97, 1.27)
did not believe there were benefits to FGM/C	53	76	58	65	16	1.17 (1.02, 1.33)*
disapproved of FGM/C [1]	67	77	30	30	10	2.57 (2.06, 3.20)
believed most community members favored discontinuation of FGM/C	24	35	25	20	16	1.76 (1.25, 2.47)*
believed had self-efficacy to resist spousal pressure to perform FGM/C on daughter [1]	72	88	51	40	27	2.20 (1.85, 2.61)

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

3.2.2.4 Outreach and advocacy in Kenya and Ethiopia

The outreach and advocacy intervention took place in 2001-2002 among Somalis in refugee camps in Kenya and among the Afar people in Ethiopia.

The intervention was first implemented in the geographic area of Kenya and was provided to Somalis (n=1,440), who resided in refugee camps close to the Somali border (Chege et al., 2004). The community members residing in one camp, Ifo, were selected to receive the intervention (n=720), while residents of an adjacent refugee camp, Hagadera, received educational activities and served as the comparison group (n=720). There was no information about the educational activities provided in the report, but we note that this is the only study with a comparison group that received some form of service. The selected study camps in Dadaab in the north-eastern province of Kenya had been in operation since the 1990s and had a variety of public services, including hospitals, clinics, and schools. The study population consisted of both women and men, 15-60 years (mean age 31), most were Muslim, 74% were unemployed, and 61% had less than primary school education.

Next, the intervention was implemented in Ethiopia, near Awash town in the northeast of Ethiopia (Chege et al., 2004). The intervention was provided to female and male community members in six villages (n=407) while residents in six other villages served as the comparison group and received no intervention (n=419). It was targeted to the Afar people. The mean age of the participants was 26 (range 8-60), 82% had education below primary level, and the majority of the participants were Muslim.

3.2.2.4.1 The intervention

As described in Chege and colleagues (2004), in Kenya, the outreach and advocacy intervention lasted for 18 months and took place from January 2001 through June 2002. In Ethiopia, it lasted for 15 months from August 2001 through October 2002. While the same intervention was used in both sites, the study authors explain that the approaches to address human rights and gender differences were not the same

between the two study sites because of the different cultural contexts. The intervention had two components: 1) community-level information and education activities using behaviour communication change approaches, and 2) community-level advocacy. Activies included educational events, community meetings, theatre group performances, video sessions, mass media activities, and support of advocacy activities. It consisted of four behaviour change techniques (2, 21, 29, 30):

- Provide information on consequences of behavior to the individual
- Provide instruction on how to perform the behavior
- Plan social support/social change
- Prompt identification as role model/ position advocate.

We explain the judgements in appendix 5.

3.2.2.4.2 Mechanisms assumed to be at work

Because we did not identify any other publications related to the outreach and advocacy intervention, we searched and extracted information concerning mechanisms only from the effectiveness report (Chege et al., 2004). In appendix 6, we show the data related to mechanisms in both study areas. Based on the effectiveness report, we identified four change theories that were assumed to underpin the intervention in both areas:

- Education leads to increased awareness
- Training and education trigger advocacy ("speaking out against FGC")
- IEC (information, education, communication) activities affect intentions
- IEC activities lead to individuals' improved knowledge and attitudes, which lead to groups' increased mutual understanding and agreement, which translates into collective action, which in turn shapes social norms.

From statements about behavioural, social, administrative, and related factors that may have affected the results of the intervention in Chege and colleagues (2004), we identified several factors postulated to explain its success in both intervention sites: 1) Program changes were possible because it was a community-based intervention, 2) FGM/C was addressed as part of a larger set of reproductive health issues, 3) the intervention organization appeared 'neutral' by not imposing its values on the community, and 4) project staff were trusted. Additionally, in the Ethiopian site, the intervention was successful in mobilizing religious leaders.

We similarly identified two process factors which may have reduced the effectiveness of the intervention in the two sites: community members objected to the intervention, and there was insufficient exposure to the intervention. In the Kenyan site, four additional factors were brought forth: 1) planned work with religious leaders did not take place and they subsequently gave mixed messages, 2) some program messages were not recalled (messages that FGM/C violates girls' human rights were not recalled, best recalled was messages that FGM/C is harmful

to health), 3) enforcement of a Kenyan law criminalising FGM/C during the intervention may have led to a backlash, and 4) the intervention was implemented by a Christian group, which may have raised distrust among the predominantly Islamic group receiving the intervention.

3.2.2.4.3 Results of the intervention

We extracted four outcomes from the outreach and advocacy intervention in Kenya for which there were both pre– and posttest data available (table 6). All were secondary outcomes. There was one significant difference between the groups at endline: a greater proportion of individuals in the comparison group than the intervention group believed that FGM/C compromised the human rights of women.

Table 6. Study outcomes and effect estimates for outreach and advocacy intervention in Kenya

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
no intention to perform FGM/C on daughter	14	17	10	18	-5	0.94 (0.75, 1.17)
supported abandonment of FGM/C in own community [1]	23	23	11	19	-9	1.21 (0.99, 1.48)
believed FGM/C compromised human rights of women	25	31	27	40	-7	0.77 (0.67, 0.89)*
knew of harmful consequences of FGM/C [1]	57	91	71	89	16	1.02 (0.99, 1.06)

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

We extracted four outcomes concerning the intervention taking place in Ethiopia (table 7). These were all were secondary outcomes. All four outcomes were significantly different between the groups at endline, but there was a large baseline difference between the groups regarding one outcome.

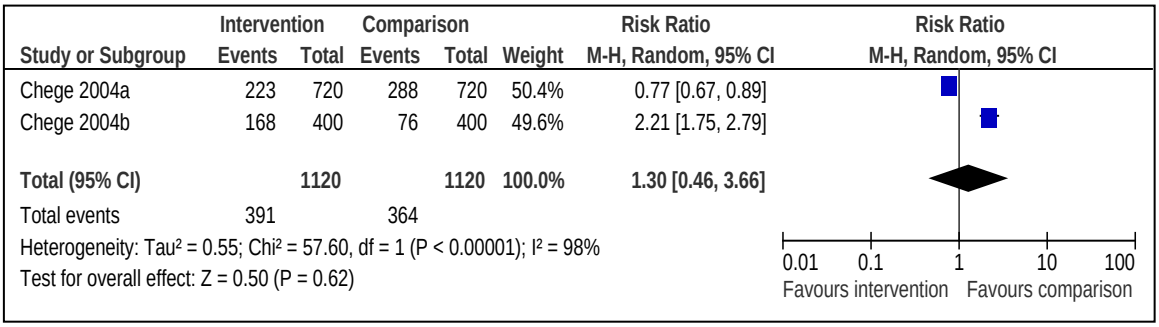
Table 7. Study outcomes and effect estimates for outreach and advocacy intervention in Ethiopia

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
no intention to perform FGM/C on daughter	8	34	14	13	27	2.62 (1.96, 3.49)*
supported abandonment of FGM/C in own community [1]	22	54	15	25	22	2.16 (1.78, 2.62)
believed FGM/C compromised human rights of women	8	42	10	19	25	2.21 (1.75, 2.79)*
knew of harmful consequences of FGM/C	33	87	36	71	19	1.37 (1.26, 1.49)*

Note: Pre-, and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

There was one outcome from these two studies for which we could conduct meta-analysis. For the outcome 'belief that FGM/C compromise the human rights of women', the result showed no significant difference between the groups at endline (RR=1.30, 95% CI= 0.46, 3.66). Further, as shown in figure 3, there was high heterogeneity (I^2=98%). The study-level result was in opposite directions in the two studies: In Kenya, a higher proportion of comparison participants believed that FGM/C compromised the human rights of women, while in Ethiopia a higher proportion of intervention participants believed this at endline. This finding is despite the fact that "in Ethiopia, there were no direct educational messages speaking to human rights, just actions that would promote gender equity" (Chege et al., 2004, p24).

Figure 3. Forest plot, belief that FGM/C compromise human rights of women



3.2.2.5 Tostan education program in Mali, Senegal, and Burkina Faso

Three of the eight interventions identified consisted of a community-based education program, developed by the non-profit organization Tostan (Gillespie & Melching, 2010). The four module education program took place in Mali, Senegal, and Burkina Faso. The modules covered hygiene, problem solving, women's health, and human rights. They were delivered by trained facilitators, in the local languages of the sites, during two-hour sessions, held 2-3 days per week for 6-8 months at some point between 2000-2002 (Easton et al., 2002; Diop et al., 2004; Ouoba et al., 2004). Gillespie and Melching (2010, p478) describe the intervention as a "nonformal adult education programme" and the terminology used for the program in the three intervention countries was "basic education program" in Senegal (Diop et al., 2004) and "village empowerment program" in Burkina Faso and Mali (Easton et al., 2002; Ouoba et al., 2004). We describe the details of each of these studies below before we summarise them.

3.2.2.5.1 Tostan education program in Mali

The goal of the program in southern Mali was "to test the Tostan Village Empowerment program in the Malian context as a participatory approach to stopping FGC" (Easton et al., 2002). It took place in five villages in the region of Kati (n=132). Four other villages were chosen in the same geographic area, some distance from the intervention villages, to serve as comparison. The residents of these comparison villages did not receive any intervention (n=107). The female and male

study participants were mainly Bambara, the mean age was 36, most considered themselves Muslim, and 69-87% had no education.

The three educational modules hygiene, problem solving, and human rights were provided through two-hour sessions, three days per week for a total of 36 sessions between March and May 2000. The fourth module, which covered women's health (including FGM/C), was offered as a three day intensive workshop (36 hours) in June and July 2000. In total, the intervention duration was six months. We identified four behavior change techniques for the Tostan education program in Mali (techniques 2, 6, 26, 30):

- Provide information on consequences of behavior to the individual
- Goal setting (outcome)
- Prompt practice
- Prompt identification as role model/ position advocate.

In appendix 5, we explain the judgements, including text from the publication (Easton et al., 2002), regarding the behaviour change techniques identified.

Based on three reports about the intervention (Easton et al., 2002; Gillespie & Melching, 2010; Monkman et al., 2007) we identified four behaviour change theories (appendix 6), that were assumed to underpin the intervention:

- Education leads to increased knowledge and understanding (critical awareness), which foster confidence and empowerment, which affect a sense of activism/motivation to act
- Education affects intentions, attitudes, and skills
- Education leads to public discussions
- Education increases empowerment, which affects attitudes and behaviors.

A range of comments regarding process factors that may have enhanced or reduced the effectiveness of the intervention were offered by the authors of the intervention report (Easton et al., 2002) and authors of two other publications commenting on the intervention (Monkman et al., 2007; Gillespie & Melching, 2010). We show these in appendix 6. The five process factors which may have enhanced the effectiveness of the intervention included: 1) the program was 'grounded' in the local context, 2) the program included both genders and men's participation was valuable, 3) there were separate women's circles which "were important reinforcement mechanisms", 4) the human rights framework was meaningful to the participants, 5) many participants received information about FGM/C during the intervention period from other sources than the intervention program.

Conversely, the three reports pointed to factors that may have negatively affected the intervention, including: 1) there was drop-out, especially by men, 2) there was insufficient pre-service facilitator training, which meant that some facilitators were less effective at engaging participants and increasing participation, 3) there were

disagreements and a lack of mutual expectations among the various entities involved in realising the program, 4) there were implementation problems, in particular regarding the reproductive health and FGM/C module, including a lack of clarity about how the program could be adjusted to fit the local context.

Despite the stated program goal of stopping FGM/C, the evaluation report for Mali included only one outcome result with respect to FGM/C: participants' position on FGM/C (support, mixed, oppose). The proportion of intervention participants who was opposed to FGM/C increased to 82% (from 8%) and the proportion of comparison participants who was opposed to FGM/C increased to 28% (from 25%). Thus, the difference between the groups was significantly different at endline (adjARD= 21, RR= 2.95, 95%CI= 1.62, 5.34). However, as shown, there was a difference between the groups already at baseline.

3.2.2.5.2 Tostan education program in Senegal

The educational intervention in Senegal took place in the region of Kolda, southern Senegal. It should be mentioned that Tostan's activities in this region started in 1988, with return in 1996 and in 1998. However, as far as we were able to ascertain, the report by Diop and colleagues (2004) describes the first controlled evaluation of the Tostan education program in this area. The "basic education program" (Diop et al., 2004 pi) was delivered to community members in 20 villages (n=949) while residents of 20 similar villages (n=383) served as comparison and received no intervention. Prior to launching the educational intervention, the study organizers undertook a "social mapping study" of 90 villages regarding their main language, health area, department, and size of population, the results of which appear to have been used to match intervention and comparison villages (Diop et al., 2004 p5). The female and male study participants in these 40 villages were mainly of the Pulaar and Mandingo ethnic groups, the mean age was 35, 78% had no education and 99% were Muslim.

The education program was delivered over three sessions per week from January to June 2001, for a total of about 70 sessions over six months. The study evaluation showed that two behaviour change techniques were used (techniques 2, 30):

- Provide information on consequences of behavior to the individual
- Prompt identification as role model/ position advocate (the judgments are provided in appendix 5).

We searched and extracted information concerning mechanisms that were assumed to underpin the intervention from the effectiveness report (Diop et al., 2004) as well as related reports regarding this intervention (Diop et al., 2008a; Diop and Askew, 2009; Gillespie and Melching, 2010; UNICEF, 2008). The identified mechanisms for Senegal were almost identical to the ones for Mali:

- Education leads to increased knowledge and understanding
- Education leads to improved attitudes and skills

- Education leads to public discussions/social interactions (mobilization) which leads to public commitment
- Education empowers people (appendix 6).

Although Diop and Askew (2009 p309) briefly note that "the rational for the TOSTAN strategy as a model for changing behavior concerning FGM/C has been explained by reference to social convention theory", this is not mentioned in any of the other reports concerning Tostan, indicating it may be an *ex post facto* reflection. We found that the data did not support placing convention theory on the list of mechanisms supposed to underpin the program. In the reports (Diop et al., 2004; Diop et al., 2008a; Diop & Askew, 2009; Gillespie & Melching, 2010; UNICEF, 2008), we found a long list of factors that may have impacted the effectiveness of the intervention. The factors that may have impacted the intervention's effectiveness positively were: 1) the participants were satisfied with the educational program, 2) "Only the most motivated women participated in the complete program", 2) over 50% of the participants had received information about FGM/C before the intervention started, 3) about 40% of the participants received information about FGM/C during the program from other sources than the intervention program, and 4) the selected intervention villages met certain criteria (had strong leaders, the ability to support a facilitator) and many were willing to abandon FGM/C (villages requested the program and believed that a condition to receive it was their willingness to abandon FGM/C).

There were even more factors that may have negatively impacted the Tostan education intervention in Senegal: 1) a religious leader openly expressed favoring the continuation of FGM/C, which may have influenced participants, 2) many villagers who had expressed interest in participating in the educational program did not attend, 3) there was inconsistent attendance and drop-out (especially among men, who were also less likely than women to share the educational information with others), 4) some community leaders and other members objected to the intervention, believing that it had come to fight against their traditions (the head of one village ordered for the program to be stopped), 5) making participants contribute in events about harmful practices such as FGM/C was more difficult compared to other events, 6) some of the village committees set up to support social action encountered problems. The texts from the publications showing these factors are provided in appendix 6.

Diop and colleagues (2004) reported the results for females and males separately. We extracted seven outcomes for female participants in the Tostan study in Senegal for which both pre- and posttest data were available (table 8). One of the seven outcomes was a primary outcome (cutting). All seven outcomes were significantly different between the groups at endline, but there was a baseline difference between the groups regarding five outcomes.

Table 8. Study outcomes and effect estimates for Tostan intervention in Senegal (women)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
0-10 year old girls who had been cut	54	40	52	52	-14	0.77 (0.64, 0.93)*
no intention to perform FGM/C on daughter [1]	29	88	11	46	24	1.91 (1.64, 2.23)
regretted having had daughter cut [1]	66	90	52	58	18	1.55 (1.37, 1.76)
disapproved of FGM/C [1]	28	84	11	40	27	2.10 (1.76, 2.51)
believed that husband disapproved of FGM/C [1]	35	86	14	46	19	1.87 (1.60, 2.18)
believed that FGM/C was unnecessary [1]	30	85	12	39	28	2.18 (1.82, 2.61)
knew at least two consequences of FGM/C	7	73	3	25	44	2.92 (2.28, 3.74)*

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

We extracted three outcomes, all of which were secondary outcomes, for male participants in the Tostan intervention in Senegal (table 9). All outcomes were significantly different between the groups at endline, but there was a baseline difference between the groups regarding two outcomes.

Table 9. Study outcomes and effect estimates for Tostan intervention in Senegal (men)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
no intention to perform FGM/C on daughter [1]	34	87	22	44	31	1.97 (1.65, 2.36)
believed FGM/C was supported by religion [1]	14	34	25	48	-3	0.71 (0.51, 0.99)
knew at least two consequences of FGM/C	11	66	14	21	48	3.10 (2.28, 4.23)*

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

3.2.2.5.3 Tostan education program in Burkina Faso

The intervention study in Burkina Faso was provided in villages in south-central Burkina Faso in 2001-2002 (Ouoba et al., 2004). Residents (n=1012) of 23 villages in Zoundwéogo province received the education program while residents (n=453) in 23 villages in the geographically remote province of Bazega did not receive any intervention and hence served as the comparison group. The female and male study participants were of the Mossi ethnic group, the mean age was 36, 90% had no education, and 45% were Muslim.

As in Mali and Senegal, the educational intervention consisted of modules covering hygiene, problem solving, women's health, and human rights. The classes were

offered two days per week for a total of 63 group sessions. These occurred between February and May 2001, and December 2001 to March 2002, for a total of 8 months. We identified two behaviour change techniques being used in the Burkina Faso educational intervention (techniques 1, 30):

- Provide information on consequences of behavior in general
- Prompt identification as role model/ position advocate.

We explain the judgements regarding the behaviour change techniques in appendix 5.

There were three publications which described the Burkina Faso study and therefore served as our sources for possible mechanisms underlying the intervention (Diop et al., 2003; Gillespie & Melching, 2010; Ouoba et al., 2004). As shown in appendix 6, the identified mechanisms were almost the same as for the Tostan education program in Mali and in Senegal:

- Education leads to increased knowledge and understanding
- Education leads to improved motivation and skills
- Discussions and meetings leads to increased knowledge, engagement, and triggers confidence (in own abilities)
- Education empowers participants.

From statements about behavioural, social, administrative and related factors that may have affected the results of the intervention in Burkina Faso (Diop et al., 2003; Gillespie & Melching, 2010; Ouoba et al., 2004), we identified no factors suggested as having affected the intervention positively but several that may have impacted the intervention negatively: 1) there was a delay in implementing the second part of the program, 2) it was difficult retaining the facilitators in the program, which disrupted normal program progress, 3) there were class attendance problems/irregular attendance, 4) the participants did not share the information they had learned with others in the community, and 5) there was a lack of tangible incentives to motivate program participants. In appendix 6, we provide explanatory text from the publications.

The effectiveness results for females and males were reported separately. We extracted eight outcomes for female participants (table 10). Two of these were primary outcomes of interest (cutting, discussed FGM/C with others). Four outcomes were significantly different between the groups at endline (two others were borderline significant), but there was a baseline difference between the groups regarding one outcome.

Table 10. Study outcomes and effect estimates for Tostan intervention in Burkina Faso (women)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)

0-10 year old girls who had been cut	6	3	4	4	-3	0.74 (0.33, 1.66)
discussed FGM/C with others [1]	54	91	64	65	36	1.40 (1.27, 1.55)
no intention to perform FGM/C on daughter	97	99	96	98	0	1.01 (0.99, 1.03)
regretted having had daughter cut	53	81	50	64	14	1.26 (1.14, 1.40)*
disapproved of FGM/C	89	98	90	94	5	1.04 (1.01, 1.08)*
believed that husband disapproved of FGM/C	97	99	97	96	3	1.03 (1.00, 1.06)
believed that FGM/C was unnecessary	93	99	92	97	1	1.02 (1.00, 1.05)
knew at least two consequences of FGM/C	52	86	57	73	18	1.18 (1.08, 1.29)*

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

We extracted five outcomes for male participants (table 11). One outcome was a primary outcome (discussed FGM/C with others). All outcomes were significantly different between the groups at endline, but there was a baseline difference between the groups regarding one outcome.

Table 11. Study outcomes and effect estimates for Tostan intervention in Burkina Faso (men)

Outcome	Intervention		Comparison		Adj	RR
	Pre	Post	Pre	Post	ARD	(95%CI)
discussed FGM/C with others [1]	62	92	79	75	34	1.22 (1.13, 1.33)
no intention to perform FGM/C on daughter	97	98	98	93	6	1.05 (1.01, 1.09)*
disapproved of FGM/C	90	98	98	93	6	1.10 (1.01, 1.09)*
believed that FGM/C was unnecessary	97	98	97	92	6	1.06 (1.02, 1.11)*
knew at least two consequences of FGM/C	64	88	66	60	30	1.47 (1.31, 1.64)*

Note: Pre- and post scores are in % and reproduced from the study publication. We calculated change scores in %, adjusted absolute risk difference (ARD) and relative risk (RR) with 95% confidence interval (CI). [1]=Baseline difference between the groups. *=RR significant at the p<.05 level and baseline similarity between groups.

3.2.2.5.4 Summary and meta-analyses of Tostan education program in Mali, Senegal, Burkina Faso

The same four-module intervention was provided in three different contexts in neighbouring countries in western Africa in 2000-2002. The Tostan adult education program lasted two months longer in Burkina Faso than the 6 month duration in Mali and Senegal. However, all of the study reports identified prompting identification as role model/position advocate as central behaviour change techniques, and change mechanisms revolved around education affecting knowledge, skills, public discourse, and empowerment.

At study level, across the three sites, of 24 extracted outcomes regarding FGM/C, 10 outcomes which were similar for both the intervention group and comparison group at baseline showed significant differences between the groups at endline, favouring the intervention. We could perform meta-analyses for three of the outcomes but no more than two studies could be synthesised in meta-analysis for any given outcome: prevalence of FGM/C among girls 0-10 years (as reported by mothers), women's knowledge of harmful consequences of FGM/C, and men's knowledge of harmful consequences of FGM/C. These three outcomes were reported in Diop and colleagues (2004) and Ouoba and colleagues (2004). As shown in figure 4, the difference between the groups regarding prevalence among girls was significant (RR= 0.77, 95%CI= 0.64, 0.92), but one study had a higher number of events than the other study and therefore assumed most weight. It therefore contributed disproportionately to the pooled effect size. The results of the meta-analyses for the outcome 'knowledge of harmful consequences of FGM/C' did not show a significant difference between the groups regarding women (RR= 1.85, 95%CI= 0.65, 5.22), but the result did show a significant difference regarding men (RR= 2.11, 95%CI= 1.00, 4.42). As further shown in figures 5 and 6, the results revealed considerable heterogeneity, I^2= 98% and 95%, respectively. For all three meta-analyses results the considerable heterogeneity and/or unequal weight affected the interpretation of effect sizes, raising doubt about the validity of the results.

Figure 4. Forest plot, prevalence of FGM/C among girls 0-10

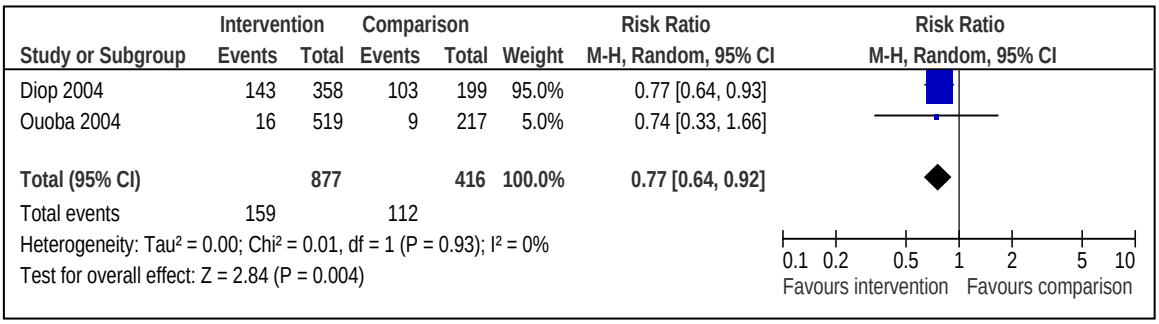


Figure 5. Forest plot, knowledge of harmful consequences of FGM/C (women)

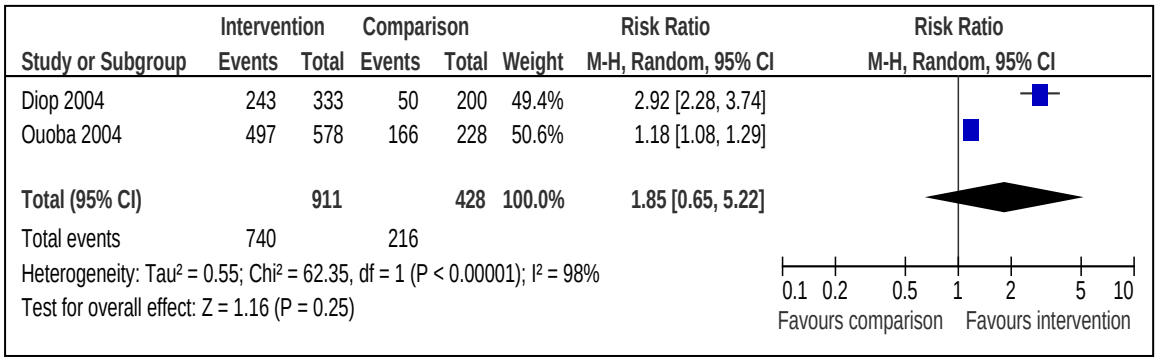
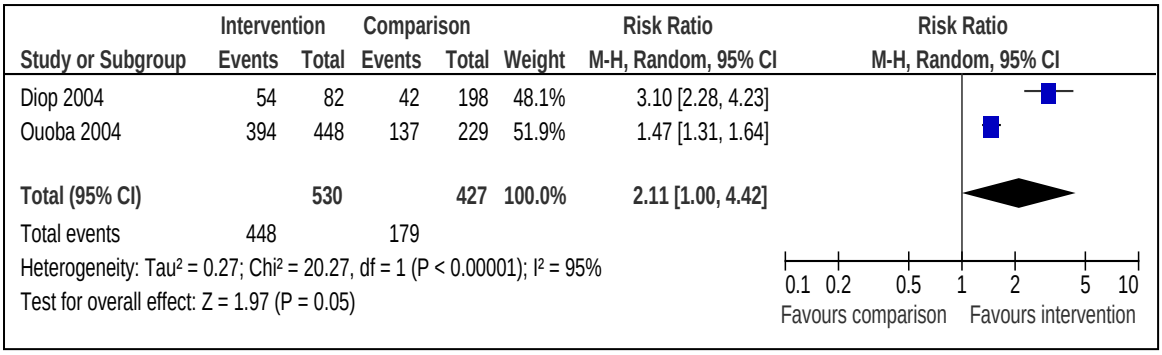


Figure 6. Forest plot, knowledge of harmful consequences of FGM/C (men)



3.2.3 Summary of effectiveness studies

Eight studies evaluating the effectiveness of an FGM/C abandonment program met our inclusion criteria. All employed a controlled before-and-after study design and all were judged to have weak methodological quality.

There were 7,042 participants at entry (range 108 to 1,465) residing in seven different African countries: Burkina Faso, Egypt, Ethiopia, Kenya, Mali, Nigeria, and Senegal. Two interventions were delivered at the individual level while six interventions targeted whole communities, involving multifaceted activities. The five different categories of interventions ranged in duration from two weeks to 18 months and took place in the eight years between 1997 and 2004.

While no behaviour change technique was common for all interventions, technique 2 (provide information on consequences of behaviour to the individual) was found in all but one of the interventions and technique 30 (prompt identification as role model/ position advocate) was found in six of the eight interventions. Given that all interventions at its foundation purported as behaviour change mechanism that providing information about FGM/C would increase knowledge – and all but one that it would improve attitudes – it seems that the proposed engine behind changing behaviours regarding FGM/C was dissemination of information.

Very few studies reported a primary outcome of interest. The primary outcome prevalence among 0-10 year old girls was reported in the Tostan study taking place in Senegal (Diop et al., 2004) and Burkina Faso (Ouoba et al., 2004), the outcome had encouraged someone not to perform FGM/C on their daughter was reported in the study implemented in Nigeria (Babalola et al., 2006), and lastly, the primary outcome had discussed FGM/C with others was reported in the study taking place in Burkina Faso (Ouoba et al., 2004). Among the 49 extracted study level outcomes, 39% of outcomes for which there was baseline similarity showed significant differences between the groups. Most of these (74%) were for the secondary outcomes attitudes/beliefs and knowledge regarding FGM/C in the community-based interventions. While we could perform four meta-analyses – belief that FGM/C compromise the human rights of women, prevalence of FGM/C among girls 0-10 years, knowledge of harmful consequences of FGM/C among women and men

– there was either high heterogeneity among the studies included or one study contributed disproportionately to the pooled result, which affected the interpretation of effect sizes and raised doubt about the validity of all results. In sum, the effectiveness of the included interventions is limited but the results point to potential advantageous developments as a result of the abandonment efforts.

3.3 CONTEXT STUDIES

3.3.1 Quality assessment of context studies

To judge the methodological quality of the 27 context studies, we used the NOKC checklist for cross-sectional studies and the CASP tool for qualitative studies, arriving at a final decision of high, moderate or low methodological quality for each study.

Among the 26 quantitative context studies, we judged that nine of them had high methodological quality, 11 had moderate and six had low methodological quality (appendix 4). All reports adequately described the population from which the sample was drawn and used appropriate statistical methods in their analyses. In most studies, the sample was representative of the population, the response rate was adequate, and standardized data collection methods were used. On the other hand, in all but two studies (Aigbodion et al., 2004; El-Gibaly et al., 2002), there was a failure to explain whether (and how) the participants who agreed to participate were different from those who refused to participate and to show that the measures were reliable and valid.

There was one qualitative study (Ragab et al., 2000), which we judged to have moderate methodological quality (appendix 4). While the study was generally methodologically sound, we assessed that the relationship between researcher and participants had not been adequately considered, data analysis was not sufficiently rigorous and it was unclear whether the recruitment strategy was appropriate to the aims of the research.

3.3.2 Syntheses 2-4: Descriptions and results in the context studies

All in all, we included 27 context studies. We identified studies which described the situation of FGM/C in each of the seven areas in which a controlled study of an intervention to reduce the prevalence of FGM/C has been carried out, except for Ethiopia (table 11). To be context relevant, one inclusion criterion was that studies had to have collected data within five years of the start of the included intervention. The scarcity of identified context studies from Ethiopia and Somalia impelled us to extend the inclusion further five years, but only one study, from Somalia, was identified. We briefly describe all context studies in table 12.

Table 12. Description of included context studies (N=27), organized by country

Author	Study design	Method. quality[1]	N	Population country	region	gender
Mali DHS 1996	Representative[2]	High	9,704	Mali	country	women
Afifi 2009	Representative[2]	High	14,393	Egypt	country	women
Al-Hussaini 2003	Cross-sectional	Low	254	Egypt	north	women
Allam 1999	Cross-sectional	High	1,020	Egypt	north	both
Dandash 2001a	Cross-sectional	High	352	Egypt	northeast	women
Dandash 2001b	Cross-sectional	Moderate	282	Egypt	northeast	women
Egypt DHS 2000	Representative[2]	High	15,573	Egypt	country	women
El-Gibaly 2002	Cross-sectional	Moderate	1,236	Egypt	country	both
Ragab 2004	Qualitative	Moderate	not stated	Egypt	north	both
Yount 2004	Cross-sectional	Moderate	3,331	Egypt	north	women
Bayoudh 1995	Cross-sectional	Low	370	Somalia	southeast	both
Abubakar 2004	Cross-sectional	Moderate	210	Nigeria	northwest	women
Aigbodion 2004	Cross-sectional	Moderate	180	Nigeria	southwest	both
Briggs 2002	Cross-sectional	Moderate	195	Nigeria	southeast	both
Dare 2004	Cross-sectional	Moderate	522	Nigeria	southwest	women
Freyermeyer 2007	Representative[2,3]	High	8,199	Nigeria	country	women
Kandala 2009	Representative[2,3]	High	7,620	Nigeria	country	women
Nigeria DHS 2000	Representative[2]	High	8,206	Nigeria	country	women
Odimegwu 1998	Cross-sectional	Moderate	1,445	Nigeria	southeast	men
Odimegwu 2001	Cross-sectional	Moderate	1,488	Nigeria	southeast	both
Okemgbo 2002	Cross-sectional	Moderate	308	Nigeria	southeast	women
Osifo 2009	Cross-sectional	Low	67	Nigeria	southwest	both
Snow 2002	Cross-sectional	Moderate	1,709	Nigeria	southwest	women
Ugboma 2004	Cross-sectional	Low	600	Nigeria	southeast	both
Bop 2001	Cross-sectional	Low	900	Senegal	multiple	both
Burkina Faso DHS 2000	Representative[2]	High	7,005	Burkina Faso	country	both
Dehne 1997	Cross-sectional	Low	70	Burkina Faso	northeast	women

Note: [1] Methodological quality. [2] Nationally representative household sample. Demographic and Health Surveys have been carried out since 1984 in multiple countries (every five years) to collect information on key population and health indicators, although not all contain a module on FGM/C. [3] The study is based on the Nigeria DHS 2000.

Only one qualitative study was included (Ragab et al., 2000) in addition to the 26 quantitative studies, four of which were nationally representative household surveys. As mentioned above, nine studies (33%) had high methodological quality and 12 (44%) had moderate methodological quality. The number of participants in the included studies ranged from 67 to 15,573, with a median number of 1,020

participants (mean=3,278). We included 13 studies from Nigeria, nine from Egypt, two from Burkina Faso and one each from Mali, Senegal, and Somalia.

Among the 27 context studies, 23 (85%) were journal articles while the remainder of the publications were organisational reports, not published in the peer-reviewed literature.

3.3.2.1 Mali

According to the Mali Demographic and Health Survey (DHS, 1996), Mali is a patriarchal country where wife beating is common and the prevalence of FGM/C is 94%. Diop and colleagues (1998) found that of the clients observed in health centres, 92% were cut. There is no law explicitly prohibiting FGM/C, but existing general provisions of criminal codes in Mali can be applied to the practice (PRB, 2008). In the study describing the effectiveness of training health personnel in Segou region, south-central Mali (Diop et al., 1998), it was explained that activities aimed at reducing the practice of FGM/C have been ongoing since the late 1970s, consisting mainly of raising awareness among population subgroups, such as opinion leaders. Similarly, Easton and colleagues (2002) write that while there are few infrastructures in the Kati area where the Tostan educational program took place, most communities had contact with, and assistance from, various nongovernmental organizations, some of which had addressed FGM/C. According to the authors (2002), villages in the Kati area are predominantly Muslim with a small Christian minority. The residents are predominantly Bambara and their main economic activity is subsistence and commercial agriculture.

Diop and colleagues (1998) state that FGM/C is increasingly performed by health personnel, but that training curriculums of health workers do not cover the health problems related to FGM/C. The Mali DHS (1996) showed that across the country, 2% of all respondents (3.3% in Bamako) who were cut had been cut by medical personnel. The rate increased with lower age, from 0.2% of 45-49 year olds being cut by health personnel to 3.6% of 15-19 year olds, indicating a tendency toward medicalisation.

In addition to searching the effectiveness studies (Diop et al., 1998; Easton et al., 2002) for information on context, we included one study from Mali, the Demographic and Health Survey from 1995/1996 (DHS Mali, 1996), which included a module on FGM/C (table 13).

Table 13. Description of included context study from Mali (N=1)

Author	Population	Outcomes
Mali DHS, 1996 (High)	9,704 women. Nationally representative household sample. Age 15-49. 29% Bambara, 15% Peulh, 13% Sarakolé. 81% none, 12% primary, 7% secondary or higher education. 91% Muslim, 4% animist, 3% Christian.	Attitude towards FGM/C; Reasons to continue to practice FGM/C; Reasons to stop FGM/C

Note: High, stated in parentheses under author, refer to methodological quality.

According to Mali DHS data, more than nine of ten women were cut, the most common types being FGM/C type I (52%) and type II (47%). Three quarters (73%) had been cut between the ages of 0-10 (median age 6.3), most (82%) by a traditional circumciser. The practice was found in most parts of the country, but was particularly prevalent among the ethnic groups in the south of Mali (Bambara, Malinke, Senufo). Rates of FGM/C did not vary much among the three religious groups: Muslims (94%), Christians (85%), and animists (88%).

In line with 71% of the participants intending to cut their daughter, three quarters (75%) of the participants believed that FGM/C should continue (12% were undecided). The proportion of respondents who voiced specific reasons for continuing the practice of FGM/C were, in order of proportion who endorsed the reason: custom tradition (61%), good tradition (28%), religious necessity (13%), hygiene (6%), preserve virginity and prevent immorality (5%), and get good marriage (3%). On the other hand, 13% stated that they thought FGC should be stopped, voicing the following reasons: medical complications (45%), bad tradition (30%), it prevented sexual satisfaction (13%), and painful experience (13%).

3.3.2.2 Egypt

Above we presented one controlled intervention that had taken place among female university students in Alexandria, Lower Egypt. According to the Egypt DHS (2000), recent economic growth has been accompanied by improvements such as school enrolment (95%) and literacy (53%), but gaps remain for women, who have lower literacy rates and participation in the labour force, and the rural population, who lag behind with respect to for example access to sanitation facilities and literacy. Both the 1995 and 2000 Egypt DHS confirmed that prevalence of FGM/C in Egypt is near universal, with 97.3% of ever-married women age 15-49, and 99.1% of girls age 15-19, reporting that they had been cut. This is despite a national law from 1996 (extended in 2007) specifically prohibiting the practice, also when performed by health personnel (UNICEF, 2010). In 61.4% of cases of FGM/C among daughters, the DHS (2000) found that trained medical personnel had performed the practice, up from 54.8% in the 1995 Egypt DHS, suggesting a trend toward medicalisation of FGM/C in Egypt.

Since the one included effectiveness study took place among female students in the Lower Egypt area, we emphasize context data most pertinent for this group, but present all data extracted from the nine included context studies. We included one qualitative study (Ragab et al., 2000) and eight cross-sectional studies presented in ten publications (Afifi, 2009; Al-Hussaini, 2003; Allam, 2001; Allam et al., 1999; Dandash et al., 2001a; Dandash et al., 2001b; Egypt DHS, 2000; El-Gibaly et al., 2002; Yount et al., 2004; Yount et al., 2002). In total, 36,755 participants were included, mainly from the Lower Egypt area (table 14). Only about 800 of these participants were men (the qualitative study did not state the number of participants and Allam et al. (2001) did not state the proportion of men in the sample).

Table 14. Description of included context studies from Egypt (N=9)

Author	Population	Outcomes
Afifi, 2009[1] (High)	14,393 women. Age 15-49. 32% secondary or higher educ. Religion not stated.	Predictors of not intending to cut daughter
Al-Hussaini, 2003 (Low)	254 women. From Assiut (upper Egypt). Age 16-37. 30% none, 11% primary, 55% secondary, 5% higher educ. 96% Muslim, 4% Christian.	Reasons for FGM/C
Allam, 2001 (High)	1,020 women and men. From Cairo (lower Egypt). Mean age 20.8. All were university students. Religion not stated.	Predictors of supporting FGM/C; Attitude towards FGM/C
Dandash, 2001a (High)	352 women. From Shakira governorate (lower Egypt). 64% over age 40. 25% none, 18% literate, 58% high school educ. Religion not stated.	Predictors of having cut daughter; Reasons for FGC; Reasons against FGM/C
Dandash, 2001b (Moderate)	282 women. From Ismailia governorate (lower Egypt). All were university students. Religion not stated.	Predictors of positive attitude to FGM/C; Reasons for FGM/C; Reasons against FGM/C
Egypt DHS, 2000[1] (High)	15,573 women. Age 15-49. 43% none, 13% some primary, 13% some secondary, 31% secondary or higher educ. Religion not stated.	Predictors of intending to/have had daughter cut, believing FGM/C should continue, believing men want FGM/C to continue; Reasons for FGM/C; Reasons against FGM/C
El-Gibaly, 2002[1] (Moderate)	1,236 girls, 314 boys. Age 10-19. Education and religion not stated.	Predictors of being circumcised, believing that circumcision is necessary, believing that women should be circumcised before marriage; Reasons for FGM/C; Reasons against FGM/C
Ragab, 2000 (Moderate)	Women and men, health workers, religious leaders and their wives. From El-Fayom, El-Minia, Bani-Suif (governorates in upper Egypt). Age, education and religion not stated.	Beliefs about FGM/C; Reasons for FGM/C; Role of husbands in FGM/C
Yount, 2004 (Moderate)	3,331 women. From Minia (lower Egypt). Age15-54. 71% none, 18% primary, 2% preparatory, 9% secondary or higher educ. 56% Muslim, 44% Coptic/Christian.	Predictors of having cut daughter, intention to cut daughter, reporting positive effects of FGM/C, reporting negative effects of not doing FGM/C

Note: Note: High, moderate and low, stated in parentheses under author, refer to methodological quality. We are not including information for ethnic groups because this information was missing in all included studies for Egypt. Educ.= education. [1] = Nationally representative household survey, the Demographic and Health Survey (DHS).

The Egypt DHS (2000) identified rural residence and low education as factors associated with women believing that FGM/C should continue and believing that men want the practice to continue. In addition, factors associated with having a daughter already cut/intending to cut their daughter were: older age, rural Upper Egypt residency, and no education. Four studies (Allam et al., 2001; Dandash et al., 2001b; El-Gibaly et al., 2002; Yount et al., 2004) presented results of multivariate logistic regression models to identify predictors perpetuating FGM/C, such as

predictors of condoning FGM/C and predictors of believing FGM/C is necessary. Afifi (2009) presented results of a regression model with predictors of women not intending to perform genital cutting on their daughter. The results of these regression models are presented in appendix 7, table 8. A comparison of factors across these regression models reveals that four recurrent factors seem to perpetuate FGM/C in Egypt:

- Rural (Upper Egypt) residence significantly increased the chance of respondents being in favour of FGM/C.
- Girls and women with no or low education were more likely than women with some or higher education to favour FGM/C.
- Women who had parents with at least preparatory education were less likely to either have had their daughter cut or intended to have their daughter cut.
- Women from Muslim families were at least four times more likely than women from families with other religions to either have had their daughter cut or intended to have their daughter cut.

With respect to female students in Lower Egypt, we note that factors predicting a positive attitude to FGM/C was rural residency and being a first year students (vs being a later year student) (Dandash et al., 2001b).

Five studies (Al-Hussaini, 2003; Dandash et al., 2001a; 2001b; Egypt DHS, 2000; El-Gibaly et al., 2002) listed the proportion of respondents who voiced various reasons in support of FGM/C (table 15). The majority (63%) of the female students in one study (Dandash et al., 2001b) stated that their reason for supporting FGM/C was that it was a good habit.

Table 15. Results of context studies from Egypt, reasons for FGM/C

Reasons	Al-Hussaini 2003	Dandash 2001a	Dandash 2001b	Egypt DHS 2000	El-Gibaly 2002
Custom/tradition	47%	30%	63%	58%	26%
Religion	1%	18%	7%	12%	7%
Reduce sexual desire	10%			32%	5%
(sexual) morality		49%	12%		
Cleanliness/hygiene	2%	4%	19%	29%	5%
Marriage prospects				4%	4%
Parental ignorance	4%				
Preservation of virginity				9%	
Social acceptance				4%	
Good health					4%

Note: The percentages refer to proportion of respondents who endorsed the reason in the study. We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

Four studies (Dandash et al., 2001a; 2001b; Egypt DHS, 2000; El-Gibaly et al., 2002) listed the proportion of respondents who voiced various reasons in opposition to FGM/C (table 16). Among the female students (Dandash et al., 2001b), no reasons dominated as far as stated reason for having a negative attitude towards FGM/C, but reasons included that they didn't believe in it (36%), they feared complications (29%), it affected marriages (23%), and they thought it was against religion (13%).

Table 16. Results of context studies from Egypt, reasons against FGM/C

Reasons	Dandash 2001a	Dandash 2001b	Egypt DHS 2000	El-Gibaly 2002
(Fear of) complications		29%	17%	24%
Not religious requirement		13%	2%	18%
Sexual problems	22%		11%	
No value/benefits	16%			14%
Old custom	62%			
Form of violence				3%
Can cause marital problems				3%

Note: The percentages refer to proportion of respondents who endorsed the reason in the study. We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

We used a generic inverse variance approach to synthesise reasons voiced for and against FGM/C across studies. The results of the meta-analyses are presented in appendix 7, figures 1-10. Due to high heterogeneity across the studies and/or unequal weight, which raise doubt about the validity of the results, we do not detail the pooled estimates here, but we note the main results. On average, across the studies, we found that:

- 45% of the respondents (5 studies, range 26-63%, 95% CI=0.28-0.61) stated that the reason in support of FGM/C was custom and tradition
- 30% of the respondents (2 studies, range 12-45%, 95% CI=-0.06-0.67) said the reason was sexual morals or morals
- 16% of the respondents (3 studies, 5-32%, 95% CI=-0.05-0.36) said the reason was to reduce sexual desires or preserve virginity
- 12% of the respondents (5 studies, range 2-29%, 95% CI=-0.03-0.26) said the reason was to ensure hygiene and cleanliness
- 9% of the respondents (5 studies, range 1-18%, 95% CI=0.03-0.15) voiced religion as a reason for FGM/C.

With respect to reasons stated in opposition to FGM/C, we found that:

- 22% of the respondents (3 studies, range 17-29%, 95% CI=0.15-0.30) said it was fear of complications or harmful health outcomes

- 16 % of the respondents (2 studies, range 11-22%, 95% CI=0.06-0.27) said sexual problems
- 14% of the respondents (2 studies, range 14-16%, 95% CI=0.10-0.19) said it had no value or benefit
- 11% of the respondents (3 studies, range 2-18%, 95% CI=-0.00-0.22) said they opposed FGM/C because it was against religion and it was not a religious requirement.

Additionally, two studies described participant's beliefs about FGM/C. First, Allam and colleagues (2001) found that among female and male college students from universities in Cairo, 95% opinioned that FGM/C was not important for marriage, 72% believed that abolishing FGM/C would be an improvement for society, and 69% thought that discussions in the media (including television) were important to ban FGM/C. Second, in the Egypt DHS (2000), the proportion of female respondents who agreed with various statements regarding FGM/C were: religious tradition (73%), husband prefers (67%), prevents adultery (51%), lessens sexual satisfaction (37%), can lead to death (29%), causes infertility (8%), can lead to childbirth difficulties (8%).

The qualitative study (Ragab et al., 2000) described that the participants held various 'misbeliefs' regarding FGM/C, including that it is a religious necessity and beautifying procedure, it protects and controls the sexuality of women, it has health benefits, it keeps fertility and is beneficial for childbirth, as well as facilitating sexual intercourse. The researchers found that the participants showed strong support for the procedure, stating that it was a good tradition and women believed that men prefer women who are cut. Lastly, the researchers concluded that the role of husbands in decision-making related to FGM/C was important, both directly and indirectly.

3.3.2.3 Nigeria

Snow and colleagues (2002) describe Nigeria as culturally diverse, with several hundred ethnic groups, about half of the population describing themselves as Muslim, and women's status and educational attainment varying greatly across the country. According to other researchers (Odimegwu & Okemgbo, 2003; Okemgbo et al., 2002), Nigeria is a patriarchal society in which there is an absence of government policy on gender equality, predisposing women to various acts of gender-based violence. Acts include battery, marital rape, sexual abuse of female children, and FGM/C. While there is no federal law against FGM/C in Nigeria (PRB, 2008), researchers state that there have been decades of ongoing efforts to stop FGM/C (Babalola et al., 2006; Briggs, 2002; Freymeyer & Johnson, 2007; Snow et al., 2002).

According to the Nigeria DHS (2000), one in every four woman in the country has gone through FGM/C (25.1%), with 82.2% having FGM/C type I, 6.7% type II, and 3.8% type III. However, rates of FGM/C were found to be highest in the southwest

(48.4%) and southeast (36.5%) regions, which include Enugu State and Ebonyi State, the areas in which the included effectiveness study from Nigeria took place. The practice is commonly performed in infancy (42% were cut before the age of 1), primarily by a circumcision practitioner or traditional birth attendant (73%). Rates of FGM/C did not vary much among rural and urban women, but a slightly larger proportion of urban women were cut. Similarly, FGM/C was more common among better-educated women.

We included 13 context studies from Nigeria (table 17), including the Nigeria DHS (2000), referred to above. Since the one included effectiveness study from Nigeria took place among female and male community members in the southeast, we emphasize context data most pertinent for this group, but present all data extracted from the 13 included context studies (presented in 14 publications). All included context studies were cross-sectional studies (Abubakar et al., 2004; Aigbodion et al., 2004; Briggs, 2002; Dare et al., 2004; Freymeyer & Johnson, 2007; Kandala et al., 2009; Nigeria DHS, 2000; Odimegwu et al., 1998; Odimegwu et al., 2001; Odimegwu & Okemgbo, 2003; Okemgbo et al., 2002; Osifo & Evbuomwan, 2009; Snow et al., 2002; Ugboma et al., 2004). In total, 30,749 participants were included, mainly from the south of the country (table 17). Only 2,607 of these participants were men and we include them with women, unless otherwise stated.

Table 17. Description of included context studies from Nigeria (N=13)

Author	Population	Outcomes
Abubakar, 2004 (Moderate)	210 women (pregnant). Age 16-45. 50% secondary or higher educ. 58% Hausa, 17% Igbo. 76% Muslim, 24% Christian. 23% cut.	Predictors of favouring continuation of FGM/C; Intention to cut daughter; Reasons for FGM/C; Reasons against FGM/C
Aigbodion, 2004 (Moderate)	180 (100 female and 80 male) undergraduate students at University of Ibadan. 66% from low educ. background. 61% Christian, 39% Muslim. Ethnicity not stated.	Differences among groups with regards to having favourable attitude to FGM/C[3]
Briggs, 2002 (Moderate)	195 (95 women and 100 men). Age 15-49. 15% no educ., 74% secondary or higher educ. 75% cut. Ethnicity, religion not stated.	Beliefs about FGM/C; Intention to cut daughter
Dare, 2004 (Moderate)	522 women (pregnant). Mean age 26. 100% cut. Educ., ethnicity, religion not stated.	Know FGM/C is illegal; Intention to cut daughter; Reasons for FGM/C
Freymeyer, 2007 [1,2] (High)	8,199 women. Age 15-49. 25% cut. Educ., ethnicity, religion not stated.	Predictors of favouring continuation of FGM/C
Kandala, 2009 [1,2] (High)	7,620 women. Age 15-49. 22% cut. Educ., ethnicity, religion not stated.	Predictors of being cut, having cut daughter

Author	Population	Outcomes
Nigeria DHS, 2000 [1] (High)	8,206 women. Age 15-49. 37% secondary or higher educ. 25% cut. Ethnicity, religion not stated.	Intention to cut daughter; Reasons for FGM/C; Reasons against FGM/C
Odimegwu, 1998 (Moderate)	1,445 boys/men (683 adolescent boys). Mean age for adult men 37. Mean age for boys 15. 47% secondary or higher educ. 74% mainline churches, 17% traditional religion. Ethnicity not stated.	Reasons for FGM/C; Reasons against FGM/C
Odimegwu, 2001 (Moderate)	1,488 (726 women and 762 men). Mean age 37. Igbo. 47% secondary or higher educ. 81% Christian, 19% traditional.	Reasons for FGM/C; Reasons against FGM/C
Okemgbo, 2002 (Moderate)	308 women. Age 15-49. 63% secondary or higher educ. Igbo. 94% Christian, 4% traditional.	Reasons for FGM/C; Reasons against FGM/C
Osifo, 2009 (Low)	67 (47 women and 20 men). Parents. Mean age for women 37 (20-59). 59% of low socioeconomic group (reside rural). Edo. 100% of women cut.	Reasons for cutting daughter
Snow, 2002 (Moderate)	1,709 women. Age 15-49. 76% secondary or higher educ. 30% Esan, 29% Bini, 13% Igbo. 86% Christian, 6% Muslim, 5% other/none. 46% cut.	Predictors of being cut
Ugboma, 2004 (Low)	600 (400 women and 200 men). Age 10 and older (40% age 20-29). 52% of women cut. Educ., ethnicity, religion not stated.	Reasons for FGM/C

Note: High, moderate and low, stated in parentheses under author, refer to methodological quality. DHS= Demographic and Health Survey. Educ.= education. [1] Nationally representative household sample. [2]The study is based on the Nigeria DHS 2000. [3]The publication did not provide the data, only the summary of the analysis.

Aigbodion and colleagues (2007), who examined the attitudes of students towards FGM/C, found neither a difference between women and men, nor between Muslim and Christian students, regarding favourable attitude towards the practice. However, those students from a higher educational background and with higher awareness level held less favourable attitudes towards FGM/C. Four studies (Abubakar et al., 2004; Freymeyer & Johnson, 2007; Kandala et al., 2009; Snow et al., 2002) presented results of multivariate logistic regression models to identify predictors perpetuating FGM/C, such as predictors of favouring continuation of FGM/C and predictors of having a cut daughter. The results of these regression models are presented in appendix 7, table 9. A comparison of factors across these regression models reveals that three recurrent factors seem to perpetuate FGM/C in Nigeria:

- Women who were themselves cut (and living in a community where most women were cut) were more likely than women who were not cut to favour FGM/C.
- Women with no or low education were more likely than women with some or higher education to favour FGM/C.
- The older a women was the more likely she was to favour FGM/C.

Seven studies (Abubakar et al., 2004; Dare et al., 2004; Nigeria DHS, 2000; Odimegwu et al., 1998; Odimegwu et al., 2001; Okemgbo et al., 2002; Ugboma et al., 2004) listed the proportion of respondents who voiced various reasons in support of FGM/C (table 18). The two most frequently stated reasons for favouring FGM/C among study participants from the southeast area of Nigeria were tradition/culture (57%) and that the practice reduces female sexual desire (37%).

Table 18. Results of context studies from Nigeria, reasons for FGM/C

Reasons	Abubakar 2004	Dare 2004	DHS 2000	Odimegwu 1998	Odimegwu 2001	Okemgbo 2002	Ugboma 2004
Tradition/culture	65%	62%	75%	60%	48%	79%	40%
Religion	38%	24%	2%				14%
Prevention of promiscuity/ preserve virginity		51%	14%	23%		2%	7%
Cleanliness/hygiene			5%	3%			
Better marriage prospects			7%				8%
Prevent pregnancy/delivery problems				7%	8%	9%	16%
Reduce/control female sexual desire				17%	36%		58%
Beautify female organ				5%	2%		11%
Maturity attainment				5%	4%		
Good for health						39%	7%
Enhancement of fertility	18%						
Greater pleasure of husband			3%				
Initiation to womanhood							30%
Prevent overgrowth of clitoris							20%
Improve female sexual satisfaction							13%
Be faithful to husband							13%
Cure infertility							3%

Note: The percentages refer to proportion of respondents who endorsed the reason in the study. We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

Five studies (Abubakar et al., 2004; Nigeria DHS, 2000; Odimegwu et al., 1998; Odimegwu et al., 2001; Okemgbo et al., 2002) listed the proportion of respondents who voiced various reasons against FGM/C (table 19). The most frequently stated reason opposing FGM/C among study participants from the southeast area was that there may be health complications (36%).

Table 19. Results of context studies from Nigeria, reasons against FGM/C

Reasons	Abubakar 2004	DHS 2000	Odimegwu 1998	Odimegwu 2001	Okemgbo 2002
Bad tradition	37%	61%			
Medical complications/ health problems	60%	22%	28%	52%	29%
It's not necessary			16%	22%	
Against religion		22%			
Painful personal experience		10%			
Against dignity of women		10%			
Prevents sexual satisfaction		7%			
Pregnancy/delivery problems			14%		
Tradition is not static				21%	

Note: The percentages refer to proportion of respondents who endorsed the reason in the study. We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

We used a generic inverse variance approach to synthesise reasons voiced for and against FGM/C across studies from Nigeria. The results of the meta-analyses are presented in appendix 7, figures 11-22. Due to high heterogeneity across the studies and/or unequal weight, which raise doubt about the validity of the results, we do not detail the pooled estimates here, but we note the main results. On average, across the studies, we found that:

- 61% of the respondents (7 studies, range 40-79%, 95% CI=0.50-0.72) stated that the reason in support of FGM/C was tradition and culture
- 37% of the respondents (3 studies, range 17-58%, 95% CI=0.12-0.62) stated that the reason in support of FGM/C was to reduce or control female sexual desire
- 19% of the respondents (5 studies, range 2-51%, 95% CI=0.08-0.31) stated that the reason in support of FGM/C was religion
- 18% of the respondents (4 studies, range 2-38%, 95% CI=0.06-0.30) stated that the reason in support of FGM/C was prevention of promiscuity or preserve virginity
- 10% of the respondents (4 studies, range 7-16%, 95% CI=0.06-0.14) stated that the reason in support of FGM/C was to avoid problems with pregnancy and delivery.

With respect to reasons stated in opposition to FGM/C, we found that:

- 38% of the respondents (5 studies, range 22-60%, 95% CI=0.25-0.51) said it was they though there were medical complications or harmful health outcomes

- 49% of the respondents (2 studies, range 37-61%, 95% CI=0.26-0.73) simply said they thought FGM/C was a bad tradition and 19% (2 studies, range 16-22%, 95% CI=0.13-0.25) believed that it was unnecessary.

Additionally, four studies (Abubakar et al., 2004; Briggs, 2002; Dare et al., 2004; Nigeria DHS, 2000) reported that 14%, 50%, 19% and 4%, respectively, of the respondents in these studies stated they intended to cut their daughter. Respondents in one study (Osifo & Evbuomwan, 2009) explained that the reasons for having performed FGM/C on their daughter were: "uncircumcised female was taboo" (100%), mark of cultural identity (63%), to prevent promiscuity (36%), for hygiene (31%), for safe delivery during reproductive age (15%), babies will die on making contact with clitoris during delivery (12%), and initiation to womanhood (6%). A study with female and male community members from the southeast of Nigeria explained that 80% of them believed that FGM/C was compulsory in their community. Yet, half of the respondents stated that they would not like their daughter to be cut and that most men in the community would marry a woman who was not cut. Regarding their beliefs about FGM/C, they said that it was a custom (38%), it reduced sexual desire (32%), it reduced promiscuity (25%), and it enhanced a woman's dignity (5%).

3.3.2.4 Somalis

As described earlier, we included one intervention which targeted Somali refugees in Kenya. It is believed that as many as a quarter of the total Somali population (about two million) may be living in exile, many as refugees in neighbouring countries (Talle, 2003). According to baseline assessment by the authors of the effectiveness study (Chege et al., 2004), the Somali participants were predominantly of Islamic faith and all women of reproductive age had undergone FGM/C, mainly type III (infibulation). The authors of the effectiveness study (Chege et al., 2004) write that views contributing to the continuation of FGM/C among the Somali community were: maintaining culture and tradition, ensuring fertility, conferring respect upon the girl and her family, being a devout Moslem, and ensuring the marriageability of girls and their possibilities of fulfilling lives as mothers and responsible wives. The authors state that pre-intervention research was undertaken to understand the context of the practice. Results were subsequently applied towards designing a relevant intervention in which context-specific activities were included. There is no law against FGM/C in Somalia, but Kenya passed a national law in 2001 (PRB, 2008).

We identified no relevant studies about Somalis residing in Kenya, only one study set in Mogadishu, Somalia (Bayoudh et al., 1995). This cross-sectional survey was of low methodological quality and included 300 men and 70 women. All the women reported being cut; 80% of them were infibulated. The survey asked about respondents' attitudes towards FGM/C (table 20).

Table 20. Description of included context study from Somalia (N=1)

Author	Population	Outcomes
Bayoudh, 1995 (Low)	370 women and men, age 20-60. Ethnicity, education and religion not stated.	Attitude towards FGM/C; Reasons for FGM/C

Note: Low, stated in parentheses under author, refers to methodological quality.

Almost all (97%) of the respondents were in favor of FGM/C. The authors write "only women who are infibulated and circumcised are well regarded and can marry. It is a passport to marriage. An uncircumcised woman is considered a marginal woman, hypersexual, infidel, non-Muslim and non-pure" (p240). The study participants' reasons for FGM/C were, in order of proportion of respondents offering the reason: Somali custom license for marriage (84%), religious obligation (70%), protection of virginity (27%), control sexual desire (3%).

3.3.2.5 Ethiopia

We included one intervention which targeted the Afar people in villages in the northeast of Ethiopia. According to baseline assessment by the authors of the effectiveness study (Chege et al., 2004), the prevalence of FGM/C among the Afar people is 91%, with the predominant type being infibulation. The Afar people are pastoralist and semi-nomadic. The authors of the effectiveness study further described the Afar as conservative, traditional, and practicing FGM/C mainly for religious reasons (Islam). Regrettably, the only information about the context of FGM/C in Ethiopia is from the effectiveness report as we were unable to identify any additional context studies. We mention, however, that according to the PRB (2008) and GTZ (2007) there is a national law against FGM/C since 2004.

3.3.2.6 Senegal

The Kolda region, the area where the included effectiveness study from Senegal took place, is, according to these study authors (Diop et al., 2004), economically and socially under-developed. The challenges include poorly developed health infrastructure, malnutrition among children, high maternity mortality rate, and about 90% illiteracy. Diop and Askew (2009) write that the Kolda region is characterised by gender disparities and patriarchal structures that discriminate against women. According to the Senegal DHS from 2005, 28.2% of women in Senegal have undergone FGM/C (82.7% type I, 0.2% type II, 11.9% type III), but there is great variation in rates of FGM/C across ethnic groups (1.8-93.8%). In Kolda, where the majority of the population is Pulaar and Mandingo, 93.8% of the women are cut. Tostan's FGM/C abandonment activities in the region of Kolda started in 1988, with return in 1996 and in 1998 (Diop et al., 2004). The practice of FGM/C was outlawed in 1999 (PRB, 2008).

We included one context study from Senegal (Bop, 2001) (table 21). The cross-sectional survey was conducted in 1999 in three regions (Saint Louis in the north, Kolda in the southeast, and Tambacounda in the south). It included 900

participants, 300 from each region. Almost all of the 450 women participants (98%) had undergone FGM/C.

Table 21. Description of included context study from Senegal (N=1)

Author	Population	Outcomes
Bop, 2001 (Low)	900 women and men, mean age 40. 97% Muslim, 2% Christian. 98% of women cut. Education, ethnicity not stated.	Attitude towards FGM/C; Reasons for FGM/C

Note: Low, stated in parentheses under author, refers to methodological quality.

Almost half (48%) of the respondents approved of FGM/C. They were asked to state the reasons or justifications for FGM/C, which, in order of proportion agreeing with the statement, were: respect tradition (94%), obey religious demand (39%), guarantee women's cleanliness (52%), initiate girls (53%), for women to get married (22%), and men prefer cut women (21%).

3.3.2.7 Burkina Faso

Above we presented one controlled intervention that had taken place among females and males in 23 villages in south-central Burkina Faso in 2001-2002 (Ouoba et al., 2004). According to the United Nations Development Program (2010), Burkina Faso is among the 20 least economically developed countries in the world, with a mean length of adults' schooling of only 1.3 years, high maternal mortality rate, and a large gender inequality index. According to the 2000 DHS, 71.6% of the women were genitally cut (31.3% type I, 67.7% type II, 0.7% type III, 0.3% other/don't know). A national law against FGM/C was adopted in 1996 (PRB, 2008), as one of several signs of Burkina Faso having "succeeded in achieving and sustaining a remarkable level of activity against the practice" (Diop et al., 2008b p2). Activities against the practice in Burkina Faso are varied and include political leaders, religious leaders, the security forces, medical personnel, youth associations, and the media (Diop et al., 2008b).

The search and screening process identified two context studies set in Burkina Faso (Burkina Faso DHS, 2000; Dehne et al., 1997) (table 22). The largest study, the Burkina Faso DHS from 1998/1999 included a nationally representative household sample of 7,005 women and men from across the country. The second was a cross-sectional study from the northeast area of the country (department of Sebba, province of Seno) with 70 female participants, all of whom were genitally cut (the most frequent type appears to have been type II).

Table 22. Description of included context studies from Burkina Faso (N=2)

Author	Population	Outcomes
Burkina Faso DHS, 2000 (High)	7,005 (4,615 women and 2,390 men). Age 15-49. 61% Mossi, 6% Fulfuldé/Peul, 6% Bobo, 6% Gurmances. 85% none, 9% primary, 6% secondary or higher education. 61% Muslim, 20% Catholic, 14% traditional/other, 5% Protestant. 71.6% of women cut.	Reasons for and against FGM/C
Dehne, 1997 (Low)	70 women. Mean age 37 (19-58). 33% Rimaibe, 33% Gurmances, 20% Fulani, 14% Mossi. 77% Muslim, 23% Animist.100% cut. Education not stated.	Reasons for FGM/C

Note: High and low, stated in parentheses under author, refer to methodological quality.

The Burkina Faso DHS report was descriptive. It showed that the prevalence of FGM/C was highest among respondents residing in rural areas (82% vs 70% in urban areas); and among those who were Muslim (78% vs 67% Catholic, 56% Protestant, 61% traditional/other religion). The highest proportions of women who thought FGM/C should continue were found among those who were themselves genitally cut (74% vs 7% who were not cut), those living in rural areas (23% vs 11% living in urban areas), those with no formal education (23% vs 17% with primary education vs 5% with secondary education). Women were generally more supportive of FGM/C compared to men: 21% of women and 17% of men thought FGM/C should continue while 65% of women and 69% of men thought the practice should stop. The reason most frequently endorsed for its continuation was custom/tradition (table 23). Medical complications and the fact that the practice is prohibited by law were by far the two most frequently offered reasons for why FGM/C should stop, among both female and male participants (table 24).

Table 23. Results of context studies from Burkina Faso, reasons for FGM/C

Reasons	Women	Men
Custom/tradition	70.7%	65.3%
Hygiene	10.4%	19.6%
Avoid immoral behavior/preserve virginity	9.7%	13.7%
Religion	9.1%	19.0%
Chance for marriage	4.3%	13.1%
Pleasure	1.5%	0.3%
Other	24.3%	17.5%

Note: The percentages refer to proportion of respondents who endorsed the reason in the study.
We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

Table 24. Results of context studies from Burkina Faso, reasons against FGM/C

Reasons	Women	Men
Bad tradition	1.7%	1.3%
Against religion	0.7%	0.7%
Medical complications	51.1%	67.0%
Painful experience	7.7%	0.8%
Against women's rights/dignity	1.6%	6.9%
Prevents male sexual pleasure	0.9%	6.0%
Prevents female sexual pleasure	4.4%	0%
Prohibited by law	43.6%	29.2%
Other	22.0%	25.3%

Note: The percentages refer to proportion of respondents who endorsed the reason in the study.
We do not present reasons given by less than 2% of the respondents in a study and other/no reason.

The survey by Dehne and colleagues (1997) asked the female study participants to provide reasons for FGM/C. Their answers and the proportion of respondents who endorsed the reasons included: to avoid *bourowel* (alleged disease of the genital organs) (57%), religion demands that all women are excised (17%), to reduce sexual desire ("women/girls who are not excised like men too much") (17%), because this is our tradition (15%), those who are not excised are not allowed to attend the prayers (13%), to facilitate childbirth (4%), excision was already practiced when/before we were born (26%). Two percent voiced the following reasons: our grandparents already practiced excision, it is forbidden for men to marry uncircumcised women, my friends will reject me, unexcised women are not even allowed to fetch water. Lastly, 6% did not know what the reason for FGM/C was.

3.4 REALIST SYNTHESIS – CMO CONFIGURATIONS

Realist synthesis is concerned with explaining the relationship between the context in which the intervention is applied, the mechanisms supposed to underpin the program, that is, the catalyst behind behaviour, and the outcomes which are produced (Pawson 2006; Pawson et al., 2004; 2005). In this section we describe the results of our realist synthesis of the five broad categories of interventions included. They are summarised in tables 25-32. Research question two is further addressed here.

3.4.1 Training of health personnel in Mali

In the first intervention presented, training of health personnel in the south of Mali, the program was embedded in local public health services. It was aimed at improving health providers' involvement with FGM/C, given their routine contact with people affected by FGM/C in this area of the country, where FGM/C is

particularly prevalent. In such a context where FGM/C may be increasingly performed by health personnel, the intervention involved an appropriate stakeholder group with a potential role in reducing the prevalence of the practice. Such personnel are furthermore figures of authority in many communities; that they possess extensive, accurate knowledge about the practice is essential. In the same context, when the intervention involved training health clinic staff in FGM/C and its health complications and their treatment, the aim was for the training to improve their knowledge, attitudes, and skills. However, the results show that such improvements were not triggered by the intervention: there was no significant difference between the intervention and comparison groups regarding any of the outcomes. The study authors (Diop et al., 1998) suggest that this may be due to the time span devoted to training, 3 group sessions, being too short (table 25). When implementing a training program, allowing adequate time for adoption of the intended ideation- or behavior changes is essential.

While the training may have been too short, the intervention content – which consisted of "recall of female anatomy and FGC, its context and local rationale, its prevalence in Mali and elsewhere, and the different types of cutting. The health complications and their treatment was [sic] emphasized" (Diop et al., 1998, p16) – seems to be a fitting response both to strengthen health workers' clinical involvement in FGM/C and preparing them for supporting and informing patients and communities about the benefits of discontinuing the tradition of FGM/C. Among Malians thinking FGM/C should be stopped, the most frequently voiced reason for opposing the practice was medical complications (Mali DSH, 1996). It appears the intervention build upon an important factor already accepted in the community discourse regarding FGM/C. It is unknown, however, the degree to which the intervention content's "context and local rationale for FGM/C" were grounded in empirical data about these issues. The study publication did not refer to any pre-studies although it was explained that a workgroup of seven experts was responsible for preparing the training for the participants.

Table 25. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding intervention for health personnel in Mali

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Mali, districts of Bamako and Bla (Segou region, south-central Mali), health personnel (Bambara, 96% Muslim). - 94% prevalence (mostly type I&II) - Especially ethnic groups in south - No national law, codes can be applied - Increasingly performed by health personnel, but no formal FGM/C training - Most support and intend to continue - Reasons for: tradition, religion, hygiene - Reasons against: complications, bad tradition, prevents sexual satisfaction.	Delivered to health personnel in Mali. Group training sessions + supervision. Duration: 2 months (likely in 1997). Techniques: - Provide information on consequences of behaviour in general - Provide information on consequences of behaviour to the individual - Goal setting (behaviour) - Provide instruction on how to perform the behaviour - Prompt practice - Prompt identification as role model/ position advocate.	Training leads to improved knowledge, attitudes, and skills. - Insufficient time span for training.	There was no significant difference between the intervention and comparison groups regarding any of the outcomes at endline (knowledge, beliefs).

3.4.2 Education of female students in Egypt

According to Mounir and colleagues (2003), the authors of the second controlled intervention study aimed at preventing the continuation of FGM/C, adolescents in Egypt lack sufficient and correct knowledge about reproductive health. Mounir and colleagues' intervention was a short-term health education program on reproductive health provided to female second year students in Alexandria, Lower Egypt. The intervention's "educational contents were based on the results of the pretest" (Mounir et al., 2003, p439), suggesting that the program incorporated target population specific factors regarding the issue. Further, the health talks included dangers of FGM/C and the role play "experience and precautions against" FGM/C and social pressures on the practice (p440). The authors stated that there was an increase in the proportion of students who agreed that traditions perpetuate the continuation of FGM/C, suggesting this point was addressed in the health talks. The behaviour change technique used, provide information on consequences of FGM/C to the individual, appears sound given that context studies showed that the most common reason for opposing the practice was fear of complications (table 26).

In the context of an urban university setting, the promotion of sexual health embedded in a reproductive health curriculum though health education was an appropriate strategy. Health education interventions have been touted as one of the most appropriate strategies for promoting young people's sexual health (Oakley et al., 1995), and the provision of formal education did increase students' knowledge of likely complications following FGM/C. While the strategy's impact on attitudes and behaviors is uncertain, the intervention proved its feasibility in increasing knowledge and went beyond proof of concept. The benefits of placing FGM/C in a reproductive health context was also supported by the Egypt DHS data showing that few women 15-49 years old recognized the potential adverse physical consequences of the practice for women. The women were more likely to recognise that FGM/C may have an effect on sexual relations: 37% believed that it lessens sexual satisfaction for a couple. Thus, in circumstances where FGM/C is near universal and a sizeable proportion believe the practice has adverse health consequences, such as in Egypt, an understanding of health implications following FGM/C through education seems promising in deterring people from supporting it and helping bring about change. Furthermore, FGM/C in Egypt reflects and reinforces ideologies about gender and women's sexuality. Likely, dissemination of factual information about FGM/C to communities in general, and educational attainment of women in particular, could increase exposure to and knowledge of FGM/C issues.

The formal educational intervention was provided to second grade, urban Lower Egypt, female students. In a country where FGM/C is practiced on more than nine of ten daughters, to prevent the continuation of FGM/C there are many possible target groups. The context data showed that rural, Upper Egypt residence and low education were the most influential factors in the continuation of FGM/C. The social pressure to follow traditions and habits is strong in such areas. These groups would

perhaps have been more appropriate recipients of an intervention than urban students, but the context studies showed that also university students favoured performing FGM/C on their daughter (over 60%) and considered it a good habit. Studies involving students suggested that interventions targeting this group should be provided in first grade, given these students' positive attitude towards FGM/C, compared to later year students (Dandash et al., 2001b). These studies also suggested that students would be receptive to discussions in media. Student respondents opinioned that information about the practice in the media, especially television, would be important to combat FGM/C in the country (Allam et al., 2001).

Table 26. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding education of students in Egypt

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Egypt (lower), urban setting in city of Alexandria, female students (47% from low social class families). - 99% prevalence among 15-19 year olds - National law prohibiting FGM/C - Role of men important but most think FGM/C is not important for marriage - Reasons for: custom and good tradition, preserve sexual morals/reduce women's sexual desires, hygiene, religion - Reasons against: physical complications/ harm, sexual problems, no value/benefit, religion.	Delivered to female students at Alexandria university in Egypt. 2 x 1-hr sessions on reproductive health. (likely in 2001). Techniques: - Provide information on consequences of behaviour to the individual.	Education leads to improved knowledge and attitudes.	Reproductive health sessions for female students achieved a significant increase in mean knowledge score about dangers of FGM/C compared to no intervention (MD= 0.75 points, 95%CI= 0.63, 0.87).

3.4.3 Communication program in Nigeria

In Babalola and colleagues' study (2006), the intervention was implemented in one Nigerian state among beneficiary communities who had low or no education. While the Nigerian community's 'need' regarding FGM/C may not be as great as in some other countries, with about one in four women being cut, FGM/C was considered an essential traditional practice in the intervention state, Enugu. Estimates indicated that 37% of women in the state were cut (table 27). In this area, the authors of the effectiveness study stated that the factors underpinning FGM/C were cultural: uncut women were considered unmarriageable, unclean, and potentially promiscuous (Babalola et al., 2006). These results appear to be derived from a pre-intervention exploration performed by one of the program agencies, the Health Communication Partnership. They largely mirror the results of our included context studies, except in the latter, cleanliness failed to emerge as an important factor for FGM/C. It was not clear the extent to which identified cultural factors were embedded in the communication intervention, but the effectiveness report suggest they were taken into account.

Specifically, Babalola and colleagues (2006) explained that the factors underpinning the practice "support the convention hypothesis" (p1595), which was the theory of behaviour change applied. In this context, with an intervention that involved multimedia communication, action plans, and community meetings with e.g. traditional and religious leaders, it appears that a strategy of facilitating group interactions centred around FGM/C among the community members and promoting advocacy with peers improved not only ideational factors but advocacy efforts. There was evidence of a shift in the social convention of FGM/C through the provision of knowledge and actions of some spread to others through social networks: value orientation, advocacy and perceived social support improved. Rather than outright condemnation, it seems the program promoted careful reflection about what FGM/C meant in the target culture and why it was perpetuated. There was a sound fit between the program theory of change and program core components. In this context, with convention theory as a driver of change, dosage of program messages seems important as outcome data documented a clear advantage of exposure to a combination of activities and mass media. Presumably, progress could have been greater had more community members been exposed to the communication program. If an intervention based on convention theory is implemented, it is essential that both program fidelity and program exposure are high. Consistent with convention theory, in a Nigerian context, education about FGM/C, public discussions and declarations of opposition to FGM/C all contributed to developing a critical mass of individuals who changed their beliefs about FGM/C.

Table 27. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding communication program in Nigeria

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Nigeria, southeast (Enugu state), community members (most low education, 58% Protestant). - 25% prevalence (37% in southeast) - No national law against FGM/C - About 20% intended to cut daughter - Considered 'compulsory' in southeast - Support for FGM/C strongest among older, cut, lower educated women - Reasons for: tradition, control female sexual desire/prevent promiscuity, religion, avoid pregnancy/delivery issues - Reasons against: physical complications/ harm, unnecessary/bad tradition.	Delivered to community members in Enugu state, southeast Nigeria. Mutimedia communication, action plans, community meetings. Duration: ~12 months (2003-2004). Techniques: - Provide information on consequences of behaviour in general - Provide information on consequences of behaviour to the individual - Goal setting (behaviour) - Goal setting (outcome) - Plan social support/ social change.	Convention theory: - program leads to increased awareness, which leads to self-examination of beliefs and values, which triggers ways of thinking and value orientations - program leads to dialogue and group/social interactions and advocacy, which in turn improves self-efficacy and perceived social support - High degree of program exposure improved FGM/C-related ideation, but 36.6% no exposure to program - Program exposure through both mass media and community activities affected change more so than exposure through either one alone.	Communication program for community members achieved a significant increase in the proportion of - women who encouraged someone not to perform FGM/C on their daughter (RR=2.68) - women who had no intention of performing FGM/C on their daughter (RR=1.13) - men who did not believe there were benefits from FGM/C (RR=1.17) - men who believed most community members favoured discontinuation of FGM/C (RR=1.76).

3.4.4 Outreach and advocacy in Somali refugee camps and in Ethiopia

In two of the eight included effectiveness studies, an outreach and advocacy intervention was implemented amongst Muslim communities who had no or low education and unmet needs. In the absence of national laws against the practice, FGM/C was near universal in both communities, who mainly practiced infibulation (tables 28-29). The intervention was embedded in existing reproductive health projects in the two sites of Kenya and Ethiopia and was implemented by locally trusted agencies that appeared 'neutral'. The fact that the issue of FGM/C was addressed as part of a larger set of health issues likely facilitated acceptance of the program, although opposition towards the project and its staff did occur in both sites. Presumably, in contexts where FGM/C is widely embraced as an integral part of a community's traditions, it is advantageous that well-regarded agencies incorporate FGM/C abandonment efforts within a larger set of health and well-being issues.

In addition to providing the intervention to appropriate communities and linking FGM/C to a wider health agenda, the intervention agencies properly conducted pre-intervention research to understand the context of the practice, the results of which were subsequently applied towards the design and content of the intervention. Specifically, the intervention included context-specific activities, took into account the health, psychosexual, and social attitudes and beliefs that influenced the communities' continuance of the practice, and provided factual information on negative health and social consequences relating to FGM/C. Further, acknowledging the strong link between FGM/C and Islam among Somalis and Ethiopians, concerted efforts were made in both sites to engage religious leaders.

In the context of the Somali refugee camp, this context-based outreach and advocacy intervention neither trigger individuals' improved knowledge, advocacy, intention, nor groups' collective understanding, norms, action. Rather, in the comparison group, which should only receive education, there was a significantly greater increase in the proportion of individual who believed that FGM/C compromised the human rights of women. One critical factor that likely accounted for both the change in the comparison group and impeded the program's success in the intervention group is the engagement of religious leaders. In the comparison camp, several religious leaders maintained previously established advocacy efforts towards ending FGM/C, while in the intervention camp, religious leaders declined to advocate against the practice. In both camps, religious leaders continued to provide mixed messages about FGM/C. It is also likely, as the authors of the effectiveness report allude to, that efforts by a Christian group to end a practice – closely linked with Islam – in Muslim communities antagonised the beneficiary community. The authors claim the organization National Council of Churches of Kenya was a neutral actor with no religious or other agenda; whether the program beneficiaries perceived it as such is questionable. Thus, considerations for future interventions include, first, in circumstances where FGM/C is strongly linked with religious beliefs to ensure religious leaders' engagement and unified position, and two, for

organisations to operate under a politically acceptable, ideally nonsectarian, umbrella of a community.

Briefly, the outreach and advocacy intervention likely did not work in the Somali context also because of insufficient program exposure and because recipients did not recall certain messages. For example, only 1% of the Somali participants recalled the message that FGM/C compromises the human rights of women. This is a reminder that intervention fidelity, and intervention receipt, are of critical importance in outreach and advocacy programs that rely on individuals' improved knowledge and ideation to affect collective action.

In a similar context of need involving Ethiopians, when the intervention succeeded in exposing participants to anti-FGM/C messages and mobilizing religious leaders, it triggered an improvement in knowledge of harmful consequences of FGM/C, belief that it compromised the human rights of women, and intentions not to perform FGM/C in the future. For outreach and advocacy program success, it seems critical to engage religious leaders in creating an enabling environment for individual behaviour change.

Somalis and Ethiopians, especially religious leaders, are likely to interpret FGM/C abandonment efforts differently. Organisations should proceed carefully and respectfully when working with communities who consider FGM/C a necessary tradition, especially Somali populations. In the outreach and advocacy intervention, as a reaction to publicly addressing the negative consequences of FGM/C, threats against staff were greater in the Somali context than the Ethiopian context. Similarly, groups working to change FGM/C should be prepared to protect community members who are marginalized and threatened when they come out against the continuation of the practice. This became necessary in the outreach and advocacy intervention.

Table 28. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding outreach and advocacy program in Somali refugee camps in Kenya

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Kenya, refugee camp for Somalis (located in northeast), community members (adult, Muslim, low education). - 100% prevalence (mainly type III) - Kenyan law against FGM/C in 2001 - Reasons for: ensuring marriageability, religion, protection of virginity, tradition.	Delivered to Somalis residing in refugee camp in northeast Kenya. Community-level information and education outreach plus advocacy. Duration: 18 months (2001-2002). Techniques: - Provide information on consequences of behaviour to the individual - Provide instruction on how to perform the behaviour - Plan social support/ social change - Prompt identification as role model/ position advocate.	- Education leads to increased awareness - Training and education trigger advocacy - IEC activities affect intentions - IEC activities lead to individuals' improved knowledge and attitudes, which lead to groups' increased mutual understanding and agreement, which translates into collective action, which in turn shapes social norms. - Program changes were possible - FGM/C was addressed as part of a larger set of reproductive health issues - Intervention organization appeared 'neutral' by not imposing its values - Project staff were trusted. - Community objected to intervention - Planned work with religious leaders did not occur, they gave mixed messages - Insufficient exposure to the intervention - Some program messages not recalled - Somalis objected to law against FGM/C - Likely distrust by Muslim program recipients towards Christian program implementers.	Outreach and advocacy for community members achieved a significant increase in the proportion of people in *comparison* group who believed that FGM/C compromised the human rights of women (RR=0.77).

Table 29. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding outreach and advocacy program in Ethiopia

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Ethiopia, near Awash town (northeast), community members (Afar, Muslim, low education). - 91% prevalence (mainly type III) - National law against FGM/C in 2004 - Strong link between FGM/C and Islam.	Delivered to community members in villages in northeast Ethiopia. Community-level information and education outreach plus advocacy. Duration: 15 months (2001-2002) Techniques: - Provide information on consequences of behaviour to the individual - Provide instruction on how to perform the behaviour - Plan social support/ social change - Prompt identification as role model/ position advocate.	- Education leads to increased awareness - Training and education trigger advocacy - IEC activities affect intentions - IEC activities lead to individuals' improved knowledge and attitudes, which lead to groups' increased mutual understanding and agreement, which translates into collective action, which in turn shapes social norms. - Program changes were possible - FGM/C was addressed as part of a larger set of reproductive health issues - Intervention organization appeared 'neutral' by not imposing its values - Project staff were trusted - Intervention succeeded in mobilizing religious leaders. - Community objected to intervention - Insufficient exposure to intervention.	Outreach and advocacy for community members achieved a significant increase in the proportion of community members who - had no intention to perform FGM/C (RR=2.62) - believed that FGM/C compromised the human rights of women (RR=2.21) - knew of harmful consequences of FGM/C (RR=1.37).

3.4.5 Tostan education program in Mali, Senegal, and Burkina Faso

One intervention was implemented in three rural village areas where the beneficiaries held no or low education. In all three countries where the Tostan education program was implemented there were national laws or codes that could be applied towards FGM/C, but there were considerable unmet needs, with over 70% prevalence of FGM/C in all three intervention areas. In two of the areas, Mali and Senegal, the program participants were mainly Muslim while in Burkina Faso 45% were Muslim (tables 30-32). The issue of FGM/C was integrated within a larger project curriculum. Regarding FGM/C, the human rights module intended for participants to better understand and argue for rights to bodily integrity, the information in the hygiene module was imparted to deepen their knowledge about the harmful effects of the practice, and the women's health module discouraged negative traditions such as FGM/C. That Tostan embedded FGM/C within a larger curriculum likely facilitated acceptance of the program, similar to the outreach and advocacy intervention among Somalis and Ethiopians, although resistance towards the project did occur. Again, in contexts where FGM/C is widely practiced and a deep-seated tradition in people's lives, it seems advantageous to incorporate FGM/C abandonment efforts within a larger set of related issues, such as health and well-being.

While the intervention was provided to appropriate beneficiary communities and integrated in a wider health and human rights agenda, it is unclear the degree to which Tostan conducted pre-implementation research to understand the contexts of the practice. However, it seems at least in Senegal the original curriculum was created in a participatory manner with the communities; "incorporating language and cultural forms familiar to participants" (Easton et al., 2002, p6), and involving facilitators who were from the target communities. In Mali, the information about FGM/C was "grounded in the local context" as the "different arguments for and against FGC were presented in the very words used by Malian women and men who had testified during previous village workshops" (Easton et al., 2002, p6). Both in Senegal and Mali, the Tostan educational program was somewhat tailored to the target group, although not necessarily from the start, as in Senegal classes were only offered to men after recognizing that not including them created problems. Yet, it remains unclear the degree to which the Tostan curriculum took into account the specific attitudes and beliefs that influenced the communities' continuance of the practice, which the context studies revealed were chiefly tradition and religion in all three areas. None of the texts implied that religion was addressed in the Tostan program modules, although one of the outcome measures for men in Senegal was believing that FGM/C was recommended by religion (the proportion who believed this *increased* in both the intervention and comparison groups). The general implication is for future programs intending to reduce the prevalence of FGM/C to recognise that gathering appropriate and sufficient data to inform the development of a program that address a group's needs and wants is the harbinger of effective practice. Specifically, as mentioned earlier, in contexts where there is a strong link

between FGM/C and religion, program planners must attract religious leaders' engagement and commitment to halting its continuation. In Senegal, one Pulaar religious leader strongly expressed his support for FGM/C, which may have influenced the faithful. In such a context, it is essential that program planners take heed of the influence of religious and other community leaders.

In contexts where FGM/C was strongly linked with tradition and religion, and program beneficiaries received an education program on hygiene, problem solving, human rights and women's health that may not have addressed their needs and wants, it triggered low attendance and drop-out, especially among male participants. Some beneficiaries believed "the program was coming to fight against the traditional culture" (Diop et al., 2004 p8) and many neither took on a role as position advocate nor spread the information learned, as expected by the program responsible.

In patriarchal contexts, when women failed to participate because "their husbands forbade it" (Diop et al., 2004 p8) and men dominated in committees, programs dealing with FGM/C should strive to include both genders and sessions should be partially co-ed and partially separate, as supported by the success of separate women's groups in Mali. Those programs that do not adequately address men are likely to experience conflict, as occurred in Senegal.

In all three contexts, there were major implementation problems, including insufficient training of facilitators, difficulties with retaining facilitators, delayed execution of modules, alteration of the program, and failure to achieve initial and ongoing consensus among those responsible for the program. These issues may have affected the study outcomes. The implications seem for complex and extended initiatives like Tostan to carefully plan all aspect of the program, including program scheduling, as well as regularly monitor implementation to safeguard against deviations from the established protocol. When deviations and drifts are detected, they must not only be recorded, but accounted for. For intervention planners who are considering a facilitator-based program like Tostan, make sure facilitators are sufficiently trained and comfortable with the topics to be raised during sessions (this was especially problematic in Mali); take also care of recruiting facilitators from the target communities (this did not occur in Mali and Burkina Faso). Skilled, community-based peer-facilitators who have background characteristics similar to those of the target population will help to ensure that the language and messages used are relevant and appropriate and make the target population better relate to a sensitive, context-bound issue such as FGM/C.

In such contexts when the same program was replicated with the aim of education increasing empowerment, the four-module Tostan education program achieved inconsistent results. In Mali, the only outcome was that a greater proportion of intervention participants than comparison participants were opposed to FGM/C. More importantly, in Senegal and Burkina Faso, the program led to fewer women stating that they had cut their daughter and a higher proportion of participants having knowledge of harmful consequences of FGM/C. Thus, depending on the context, the Tostan education program can result in both negligible- and small

positive effects. What remains unclear, as stressed by Easton and colleagues (2002), is the flexibility of the Tostan curriculum to be adjusted to differences in context when being replicated. The suggestion is for the non-governmental organization Tostan to clarify expectations and acceptability at the beginning concerning adjustments, preferably avoiding slavish fidelity that may result in the education program being true to the form, but not the essence (theory), of the program. A prerequisite is for Tostan to take care to explicate the program theory, which was differentially alluded to in the various reports about the program.

Table 30. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding communication program in Mali

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Mali, Kati area (south), community members (Bambara, Muslim, no or low education). - 94% prevalence (mostly type I&II) - Especially ethnic groups in south - No national law, codes can be applied - Most support and intend to continue - Reasons for: tradition, religion, hygiene - Reasons against: complications, bad tradition, prevents sexual satisfaction.	Delivered to community members in villages in south Mali. Tostan four-module education program (hygiene, problem solving, human rights, women's health). Duration: 6 months (2000). Techniques: - Provide information on consequences of behavior to the individual - Goal setting (outcome) - Prompt practice - Prompt identification as role model/ position advocate.	- Education leads to increased knowledge, which foster confidence + empowerment, which affect a sense of activism - Education affects intentions, attitudes, and skills - Education leads to public discussions - Education increases empowerment, which affects attitudes and behaviors. - Program 'grounded' in local context - Program included both genders - Separate women's circles were important reinforcement mechanisms - Human rights framework was meaningful to the participants - Participants received FGM/C information from other sources than the intervention. - Drop-out from sessions, especially men - Insufficient pre-service facilitator training - Disagreements, lack of mutual expectations among implementers - Implementation problems - Lack of clarity how to fit into local context.	Tostan education program for community members achieved an increase in the proportion of intervention participants who was opposed to FGM/C, but there was a baseline difference between the groups.

Table 31. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding communication program in Senegal

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Senegal, Kolda region (south), community members (Pulaar and Mandingo, Muslim, no or low education). - 28% prevalence (94% in Kolda) (mostly type I&III) - National law against FGM/C in 1999 - About half approved of FGM/C - Reasons for: respect tradition, obey religious demand, for cleanliness, initiation of girls, for women to get married, men prefer cut women.	Delivered to community members in villages in south Senegal. Tostan four-module education program (hygiene, problem solving, human rights, women's health). Duration: 6 months (2001). Techniques: - Provide information on consequences of behaviour to the individual - Prompt identification as role model/ position advocate.	- Education leads to increased knowledge - Education leads to improved attitudes and skills - Education leads to public discussions/ social interactions, which leads to public commitment - Education empowers people. - Participants satisfied with the program - Only the most motivated women participated in the complete program - Many received information about FGM/C before the intervention started - Participants received FGM/C information from other sources than the intervention - Intervention villages selected met certain criteria, many willing to abandon FGM/C. - Religious leader openly favoured FGM/C - Many who signed-up did not attend - Inconsistent participation and drop-out - Some objected to the intervention - Low participation in FGM/C events - Implementation problems.	Tostan education program for community members achieved a significant - decrease in the proportion of 0-10 year old girls who were cut (RR=0.77) - increase in the proportion of women who knew at least two consequences of FGM/C (RR=2.92) - increase in the proportion of men who knew at least two consequences of FGM/C (RR=3.10).

Table 32. Summary of the contextual factors identified, the intervention, the mechanisms triggered, and the outcomes regarding communication program in Burkina Faso

Context	Intervention characteristics	Mechanisms	Outcomes/Conclusions
Burkina Faso (Zoundwéogo, south-central), community members (Mossi, 45% Muslim, no or low education). - 72% prevalence (mostly type I&II) - National law against FGM/C in 1996 - FGM/C most common in rural areas, among Muslim women - Support for FGM/C strongest among cut, rural, non-educated women - Reasons for: tradition and custom, religion, reduce sexual desire, avoid an alleged disease of genital organs - Reasons against: medical complications, prohibited by law, painful experience, prevents sexual pleasure.	Delivered to community members in villages in south-central Burkina Faso. Tostan four-module education program (hygiene, problem solving, human rights, women's health). Duration: 8 months (2001-2002). Techniques: - Provide information on consequences of behaviour in general - Prompt identification as role model/ position advocate.	- Education leads to increased knowledge - Education leads to improved motivation and skills - Education leads to increased knowledge, engagement and triggers confidence - Education empowers participants. - Delayed implementation of part of program - Difficult retaining facilitators, which disrupted program progress - Inconsistent participation and drop-out - Participants didn't distribute information - Lack of tangible incentives to motivate participants.	Tostan education program for community members achieved a significant increase in the proportion of - women who regretted having had daughter cut (RR=1.26) - women who disapproved of FGM/C (RR=1.04) -women who knew at least two consequences of FGM/C (RR=2.92) - men who had no intention to perform FGM/C on daughter (RR=1.05) - men who disapproved of FGM/C (RR=1.10) - men who believed FGM/C was unnecessary (RR=1.06) - men who knew at least two consequences of FGM/C (RR=1.47).

4 Discussion

The present systematic review summarised the effectiveness of interventions designed to reduce the prevalence of FGM/C and contextual factors that may help explain the extent of effectiveness of such interventions. The final realist synthesis drew these two sets of data together in an attempt to explain theoretically how results from the effectiveness evaluations came about.

4.1 SUMMARY OF THE EVIDENCE

4.1.1 Effectiveness of FGM/C abandonment efforts

The present systematic review shows that there is a paucity of high quality evaluations of the effectiveness of interventions to reduce the prevalence of FGM/C. We included eight controlled studies assessing the effectiveness of five broad categories of interventions, set in seven different countries in Africa. We identified no controlled interventions that had taken place in other parts of the world. All of the evaluation studies were characterized by low methodological quality. Thus, while our calculated effect sizes for prevalence of FGM/C, knowledge, beliefs, and intentions about FGM/C suggested that there appear to be positive developments as a result of interventions, the low quality of the body of evidence affects the interpretation of results and draws the validity of the findings into doubt. We note that no more than two studies could be synthesised in a meta-analysis for any given outcome given the heterogeneity in the types of interventions and outcomes.

Reservations regarding the validity of the study findings concern in particular the observation that none of the studies were randomised, most contained prognostically dissimilar intervention and comparison groups, and contamination of the intervention seems to have occurred in four sites. Application of adequate methods of randomisation and allocation concealment typically achieve similarity of prognostic factors at baseline, which is a basic prerequisite in studies examining the effectiveness of an intervention (Higgins & Green, 2009). Our assessment of similarity of prognostic factors at baseline suggested that the endline results may be biased in favour of the intervention groups in four of the studies (Babalola et al., 2006; Chege et al., 2004; Ouoba et al., 2004) due to dissimilarities in education, ethnicity, prevalence of FGM/C, and religion. Additionally, baseline values of outcomes were in some cases 'worse' in the intervention group and in other cases in the comparison group. In Senegal, the differential selection of intervention villages and attrition among participants meant that intervention participants who remained

at endline were likely predisposed to exhibit positive changes regarding FGM/C. In both Mali and Senegal, intervention participants were likely exposed to another FGM/C program, not just the intervention examined. While acknowledging the difficulties inherent in carrying out randomised studies for complex interventions in community settings, researchers are urged to attempt randomisation and take the similarity of prognostic factors into account when planning future intervention studies. An encouragement to conduct process evaluations, which should measure implementation aspects such as program exposure and fidelity, is also called for, again, acknowledging that implementation and adherence are often difficult to achieve and to measure in multi-component health promotion interventions (Armstrong et al., 2009). Only one of the studies included a comparison group that received some form of services (Chege et al., 2004), but there was limited information provided as to the content of this service. This is regrettable since such data can illuminate effectiveness results. Relatedly, there was some information on programmatic aspects in the included studies, which allowed us to formulate hypotheses for mechanisms of differential effects, but future evaluations should strive to provide more complete process evaluations as well as report on other context relevant data. Such information ultimately helps improve future programs' effects.

As identified by our literature search as well as PRB's survey from 2007 (Feldman-Jacobs & Ryniak, 2007), a range of FGM/C-abandonment projects have taken place in African countries. Unfortunately, the great majority of interventions have not been evaluated and those that have, have used observational designs that make it difficult to draw causal inferences. For example, the Tostan homepage (http://www.tostan.org) announces that the program has been carried out in ten countries (Burkina Faso, Djibouti, The Gambia, Guinea, Guinea Bissau, Mali, Mauritania, Senegal, Somalia, and Sudan) but in addition to the three evaluations we included from Burkina Faso, Mali, and Senegal, we only identified a report of the Tostan education program from Sudan (Easton et al., 2004). This report was not included in the systematic review because it did not include any relevant outcomes regarding FGM/C, which is perplexing given that "FGC, [is] the issue driving the funding of the Tostan replications" (Monkman et al., 2007 p452).

Our results of eight effectiveness evaluations of FGM/C abandonment efforts do not provide convincing evidence of their effect, the validity of the results is uncertain, although promising trends were observed from some of the interventions. It seems training health personnel does not change their knowledge and beliefs/attitudes regarding FGM/C, while educational sessions about reproductive health, including dangers of FGM/C, seem to improve female students' awareness about FGM/C. Similarly, the extended communication program among Nigerian villagers affected both women and men's cognitions about FGM/C and our effect size estimates suggested that knowledge about harmful consequences of FGM/C among Ethiopian, but not Somali study participants may have increased as a result of the outreach and advocacy intervention. Lastly, the Tostan education program changed Mali, Senegal

and Burkina Faso villagers' cognitions – attitudes and/or knowledge – regarding FGM/C. Several of these results are encouraging and point to the promising role of especially education in the work to prevent further continuation of the practice. We return to important program elements below. In two of the eight included effectiveness evaluations, the Tostan evaluation in Senegal and Burkina Faso, the cutting status of girls was measured (reports by mothers), with endline results suggesting a drop in cutting among daughters in Senegal. The decline is significant and must be interpreted with caution given the study's methodological limitations including self-reporting by mothers. In Burkina Faso, 4-6% of girls were reportedly cut at baseline, hence limited improvement in this outcome was possible over the program period. Rate of FGM/C is the most clinically meaningful outcome for FGM/C abandonment initiatives and most reliable as an outcome when medical examinations, rather than self-reporting or reporting by parents, are used. In the UNICEF initiated international agreement on standard indicators for situation analysis towards ending FGM/C, prevalence by age cohorts 15-49 is highlighted as the most important indicator (UNICEF, 2004).

Outcome values in the comparison groups showed improvement in several cases. Among Somalis (Chege et al., 2004) a significantly greater proportion of comparison group participants than intervention group participants believed that FGM/C compromised the human rights of women. Possible reasons for comparison group improvement in the latter case may be the presence of educational activities and in the general case contamination of the intervention or that baseline surveying sensitised the respondents. In the study by Diop and colleagues (1998), data showed that from before to after the training fewer Malian health personnel, in both the intervention and comparison groups, wished to play a role in educating others about FGM/C. Several explanations are possible, including that they may have realised the additional burden such a task would constitute to their workload.

Overall, the included studies indicated that anti-FGM/C programs could improve knowledge, attitudes and beliefs related to FGM/C, but the tradition is not easily forsaken. Examination of the most recent cutting estimates in included countries from which DHS data are available offer a mixed picture of the success of combined efforts to end the practice. DHS estimates show that in Mali, rates of FGM/C are decreasing, down from 93.7% in 1996 to 85.2% in 2006 (Mali DHS, 1996; 2007). A similar decline is found in Egypt where FGM/C is down from 97.3% in 2000 to 91.1% in 2008 (Egypt DHS, 2000; 2008). Conversely, DHS data from Nigeria and Burkina Faso show that rates of FGM/C are increasing in these two countries, up from 25.1% in 1999 to 29.6% in 2008 (Nigeria DHS 2000; 2009) and up from 71.6% in 1999 to 76.6% in 2003 (Burkina Faso DHS, 2000; 2004), respectively. The country-wide fluctuations in rates are observed also in the intervention areas presented in this systematic review.

4.1.2 Factors related to the continuance and discontinuance of FGM/C

Motivated by a belief that effectiveness of interventions depends on marshalling the most appropriate program elements with an understanding of the target

populations' beliefs, attitudes, values, knowledge, and past behaviours, we synthesised results from empirical studies on the target populations' views of FGM/C. Once we had identified that controlled effectiveness studies had taken place in seven contexts, we search for and included a total of 27 eligible studies from these contexts. Regrettably, we did not identify any studies from Ethiopia. It is possible that studies from Ethiopia have been published since the effectiveness study took place, but the gap in research regarding factors sustaining the practice in this country is highlighted. The extent to which we could conclude regarding how factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions was limited. The main reason was the difficulty in judging the level of match between the interventions' content components and empirical data about factors related to FGM/C's continuation due to the effectiveness reports' lack of descriptions on intervention content. Still, the included context studies allowed us to form an understanding of the interventions' target audiences, their legislative-cultural milieu, social environment, and perceptions regarding FGM/C. We used the information to offer provisional hypotheses concerning mechanisms of effect or lack thereof, which we return to below.

Across the seven intervention areas, FGM/C ranged from 25%-100%, although all countries except Nigeria have passed legislation or codes prohibiting the practice. As stated by UNICEF (2005b), laws signal a country's disapproval of FGM/C, send a clear message of support to those who abandon the practice, and can act as a deterrent to FGM/C. Although their function of signalling official condemnation at an early stage is likely considerable, their deterrent effect have not been confirmed. Only study participants from Burkina Faso mentioned legal prohibition as a reason for not performing FGM/C. Rather, participants from all included countries, except Somalis, consistently argued that FGM/C was simply a cultural tradition and so must continue. This result is largely congruent with the WHO's "mental map" of why the practice continues. The organisation found that members of practicing communities held culturally entrenched beliefs about FGM/C, which formed a "mental map", and largely included psycho-sexual and social reasons, religion, society, and hygiene and aesthetics (WHO, 1999). By their nature, traditions are so woven into the web of thoughts, opinions, general life, and daily habits that they are difficult to abandon. Following traditions means being culture congruent. In fact, as Gali (1998) explained, in areas where FGM/C is prevalent the practice is embedded in many cultural systems through multiple ties to historical tradition, religion, and marriageability.

While tradition was the most common explanation for FGM/C given by study participants in almost all the countries included, the other reasons were more prominent in certain countries than others. That is, the 'determinants' of FGM/C or its contributing factors were not everywhere the same. After tradition, however, religion –specifically, Islam – was a common explanation in favour of FGM/C. It was mentioned by some also as a reason against FGM/C, especially among study participants in Egypt, reflecting the fact that a 'true' Islamic position on FGM/C is

impossible to identify, given those involved argue from their own interpretation of the written sources. There are four Islamic law schools, of which three regard FGM/C as recommended and the last, the Shafi'i law school, regards FGM/C as required, but each law school manifests differently in different countries according to socio-cultural practices in those countries (Roald, 2001). What is important is to draw attention to the fact that some Muslims, including Imams who promote the practice (Gali, 1998), consider FGM/C as an authentic way to be religious and a sign of religious devotion.

The results showed that another common argument for FGM/C, which is linked to religion, was reduction of women's sexual desire. In many of the included reports, FGM/C was believed to attenuate sexual desire, thereby preventing promiscuity and protecting both virginity and morality. There was a belief among respondents that a woman's sexuality is irresponsible and wanton and therefore must be controlled through FGM/C. Implications of FGM/C on sexual desire and satisfaction have been confirmed (Berg & Denison, 2011), but the procedure does not ensure virginity: deinfibulation and reinfibulation can be and are performed (e.g. Lannon, 2000; Thierfelder et al., 2005). Interestingly, Bullough (1995), a specialist in historical attitudes to sexuality in different religions, explains that Islam is essentially a sex positive religion. However, while some salafi Islamists consider FGM/C as a means to heighten female sexual desire, others regard it as a tool to reduce sexual desire (Roald, 2001). Concerning morality, the perception among many respondents was that through FGM/C, girls bear witness of unquestionable moral status, which in turn was a fundamental assurance of marriageability. The links to convention theory are clear: the theory posits that FGM/C emerged as a strategy to secure marriage by signaling fidelity and later spread to become a prerequisite for marriage (Mackie, 1996). According to Johansen (2006), the concept of virginity is of utmost importance in many practicing communities in that a family's and indeed the whole wider group's honour depends on girls' chastity. In fact, Kassamali (1998) writes that in patrilineal societies in which women have limited influence, family honour is customarily closely associated with women's sexual behaviour. It is important to note that for most women in these societies being a wife and mother is the only 'career' available, as few have the opportunity to seek education, employment or a life economically independent of a husband.

While most of the participants in the included context studies seemed to accept FGM/C as a social organising principle, among those who expressed reasons against the practice many simply explained it as a 'bad tradition', but many also mentioned harm and medical complications, as well as the prevention of sexual satisfaction. These results suggest possible factors, already accepted in community discourses, upon which to build abandonment messages. Research exists showing the detrimental sexual consequences from FGM/C for women (Berg & Denison, 2011) as well as for their male partners (Abdal Magied & Musa, 2004; Almroth et al., 2001). In fact, three of the interventions included in this systematic review – the education and training of health personnel, education of students, outreach and advocacy –

seem to have incorporated the medical harms following FGM/C into their contents. This is in agreement with UNICEF's (2005b) suggestion of key elements that can encourage abandonment of FGM/C, one of which is "an awareness on the part of a community of the harm caused by the practice" (p13).

There were less data regarding socio-demographic differentials affecting FGM/C which could be used to strategically inform abandonment efforts. A UNICEF report (2005a) describes that countries in which the prevalence of FGM/C is 80% or higher – such as in Egypt, Ethiopia, and Mali – display negligible variation by socio-demographic factors, and that abandonment programs will be most effective if they target women from all regional and socio-demographic groups in the country. In light of the results discussed above, we concur with this assessment. At the same time, across the studies from Egypt, Nigeria, and Burkina Faso, we found that rural residence, having no or little formal education, and being cut predicted higher levels of support for FGM/C, which indicates possible stakeholder groups and capacity for change within these groups.

4.1.3 Realist approach results

Strengths and limitations of our systematic review are addressed below, but it is important to start this section with the caveat that applying a realist approach proved difficult because a lot of information was missing in the effectiveness publications. In all but one case, no explicit theory was described underpinning the intervention and information on mechanisms was scarce. Nonetheless, the CMO configurations helped us better understand why and how an intervention had an impact in a specific context. On the basis of these results, we were able to formulate suggestions for future programs.

Contextual influence on the success of the intervention was reported in five studies (Chege et al., 2004; Diop et al, 2004; Easton et al., 2002; Ouoba et al., 2004) and four studies appeared to have undertaken pre-intervention research to understand the context of the practice, the results of which were subsequently used to tailor the intervention content (Babalola et al., 2006; Chege et al., 2004; Diop et al., 2004). An elementary health promotion notion is that successful interventions are not created by chance – they are the product of much effort and based on systematic planning involving a community analysis or assessment that identifies the factors linked to the health issue, which in turn are used to engage the target community (Mckenzie & Smeltzer, 2001). Programs must conform to the norms and customs of the communities they target, that is, be tailored to the local context. Applicability help ensure not only relevancy and a sense of community ownership but also sustainability of health promotion programmes, and as reports regarding FGM/C stress (WHO, 2008), community-led, sustained action is vital to ensure long lasting, positive results.

All of the programs in our systematic review were based on a theory that provision of information improves cognitions about FGM/C. There was an assumption that information corrects faulty or lacking knowledge about FGM/C and triggers changes

in beliefs. Several social change theories, such as the health belief model, theory of reasoned action, and the precaution process adoption model (Glanz et al., 2002), work from the assumption that behavioural change can be arrived at by inducing change in beliefs. Appropriately, all eight of the effectiveness studies included in this systematic review measured change in knowledge or beliefs related to FGM/C, and the results from six interventions showing changes in beliefs are encouraging. Previous reports reveal that high levels of support for the continuation of FGM/C are strongly correlated with high prevalence among daughters (UNICEF, 2005a).

We found that success of FGM/C-related information dissemination was contingent upon a range of contextual factors. One, integrating the issue of FGM/C in a larger set of community-relevant issues appears to have facilitated acceptance of the programs. In seven of the studies, the issue of FGM/C was addressed as part of local health services, reproductive health education, or a wider curriculum. Chege and collegues (2004) and Easton and colleagues (2002) found that program success depended upon their embedment in holistic projects that took into account the community context. Two, in areas where FGM/C is closely associated with Islam, it seems critical to establish an alliance with religious leaders, who often function as norm authorities, to create an enabling environment for individual behaviour change. This is in agreement with an interagency statement from 2008 (WHO), which puts forth that the perpetuation of FGM/C is partially due to local structures of power and authority, such as religious leaders.

Important process factors that seemed to reduce the effectiveness of programs included that research participants were not aware of or signed up to take account of the research dimension of the study, and the information was not recalled/retained or the information dose was not large enough. One study reported that only 63% of community members who participated in the programme completed all modules (Diop et al., 2004) and another stated that virtually no respondents recalled any messages that FGM/C violates women's rights (Chege et al., 2005) drawing treatment exposure into doubt. These are factors that may have reduced the impact of the interventions, making a potentially real difference non-significant. Further, in some instances the study design involved role conflict or uncertainties for staff (e.g. Easton and colleagues (2002) write that "staff of Tostan and of IEP ran into some major disagreements during the course of implementation" p7), there were insufficient measures in place to reduce confounding (e.g. controlling for other program influences) and adverse prevailing program and evaluation climate. In at least one case, the intervention was not implemented as planned (Easton et al., 2002). The results underscore that important implementation aspects of the interventions contributing to change included the selection, training, and supervision of facilitators and support personnel. A similar important factor for program performance was that organisations implementing the program were trusted and well-regarded, operating under a culturally acceptable, nonsectarian, umbrella of a community. Thus, our results support a development phase in which the program planners work in partnership with the local communities, especially

traditional and religious leaders, to optimise the intervention before the program is rolled out and tested in a controlled design.

4.2 IMPLICATIONS

In accordance with the realist approach, which concludes with hypotheses regarding likely caveats and considerations concerning the programs examined and some of the most applicable measures to deal with these issues, we offered recommendations in the realist approach section above. Thus, here we simply draw attention to important implications, which include:

- To stem the continuation of FGM/C there is a need to form community-approved, nonsectarian organisations – non-governmental and voluntary – that can work towards FGM/C abandonment while operating under a politically and culturally acceptable umbrella of community development.
- Program planners should attempt to gain the support of community decision makers.
- It is imperative that program planners understand the determinants of the problem and contributing factors to FGM/C, and ensure programs are culturally responsive. A precondition is the completion of a needs assessment, the harbinger of effective practice.
- Future evaluations of anti-FGM/C programs should be designed in light of the present systematic review results. Attention should be paid not only to the quality of the program but also to the way it is implemented. Specifically:
- There is a need to conduct methodologically rigorous intervention evaluations, preferably those based on randomisation of groups.
- Evaluations must document the causal chains that affect the outcomes, including extensive reporting of context related information.
- Measurement of outcomes in community-based programs should be long-term (multi-year follow-up) and include measures of prevalence, preferably ascertained through medical examination by trained personnel, rather than self-report.
- Interventions should be based on (well-defined) theories and researchers should report extensively on the intervention, mechanisms by which the intervention should bring about change, as well as the process of implementation.

4.3 STRENGTH AND LIMITATIONS

The systematic review is subject to strengths as well as limitations. We used a comprehensive and systematic literature search and systematic process for identifying relevant publications. Reports published in a language not mastered by the review team were translated. We attempted to identify the behaviour change theories underpinning the interventions as well as characterised interventions in terms of a taxonomy of techniques (Abraham & Michie, 2008), which help facilitate communication between program designers, adopter, and reviewers. For the effectiveness question, we included only controlled studies, which is encouraged by researchers in the field (Askew, 2005). In addition to performing a 'standard' systematic review of the effectiveness of FGM/C abandonment interventions, we incorporated a realist approach which allowed an assessment of what facilitates the success of these interventions.

Unfortunately, despite repeated requests to the study researchers, we failed to obtain a full text description of one potentially relevant effectiveness study (Akweongo et al., forthcoming) which was reported in a PRB publication (2006). It is possible that the study would have met our study eligibility criteria: It is a field experiment which investigates two community-based FGM/C abandonment strategies, education activities and livelihood and development activities, in northern Ghana. The main outcome is prevalence of FGM/C over a five year period. Inclusion of this study might have affected the results regarding prevalence.

In the realist approach discussion, we pointed out that it was difficult to identify strong patterns in how contexts influenced outcome of interventions due to their limited descriptions and the fact that there were few similar interventions implemented in different contexts. In addition to limited reporting on the process of implementation, the studies did not explicitly report on the relevant effective components of the mechanisms that were assumed to bring about FGM/C related behavior change. Thus, in an attempt to delineate the CMO configurations we examined the publications for implicit and explicit statements about mechanisms and context as well as relied on logic and inference to try to articulate what was often not explicated. This strategy was inspired by van der Knaap and colleagues (2008). Most importantly, we examined the outcome measures used by the researchers of the effectiveness studies: if they included attitudes and knowledge, we proposed that the intervention was intended to achieve changes in attitudes and knowledge, which in turn were assumed to lead to behavioural change. Trying to disentangle the mechanisms through which various components of the interventions affected the outcome was even more challenging. In effect, nearly all of the mechanisms we identified could mostly be described in general terms only, a challenge identified also by other reviewers of realist syntheses (van der Knaap et al., 2008). We call attention to these challenges because they can be informative for future researchers undertaking a realist approach and encourage more complete reporting of mechanisms in behavior change interventions.

In terms of undertaking, our experience with this systematic review corroborates Pawson's (2006) caution that conducting a realist approach is a complex and time consuming task which requires a high level of experience and training in multiple areas. Our understanding of the realist logic of inquiry was based on descriptive publications (Pawson, 2006; Pawson et al., 2004; 2005) and four worked examples of the approach (Dieleman et al., 2009; Greenhalg et al., 2007; Kane et al., 2010; van der Knaap et al., 2008), but we recognise that our approach is different from these four, which in turn are different from each other. This may reflect Pawson and colleagues' statement that the realist approach "cannot be used as a formulaic, protocol driven approach" (2005 p32); that it is not a review technique, but rather a review logic.

On the one hand, compared to the before mentioned examples, our approach appears to be closer to the Campbell- and Cochrane style of systematic review methodology, including assessment of methodological study quality and quality of the evidence, which Pawson (2006) recommends should be theory-driven and made not on the basis of checklists but in relation to the usage of each fragment. We also had a set of inclusion criteria for context studies while Pawson and colleagues (2005) put forth that purposive sampling and selection is necessary because there is not a finite set of relevant papers that can be defined and then found. Indeed, we acknowledge that in a realist approach, including ours, there are many more potentially relevant sources of information than any systematic review could practically cover. We prioritised the investigation of particular processes, notably factors driving and hindering the continuation of FGM/C because knowledge of why FGM/C continues may lead to the identification of more ways to work towards the practice's abandonment. On the other hand, contrary to the recommendation by Pawson and colleagues (2005), we neither entered into dialogue with the people who developed and delivered the interventions, nor did we enter into linkage, by which we involved the commissioners of the review in the production of the research synthesis.

In the end, we believe our systematic review both adhered to the hallmarks of systematic reviewing, transparency and reproducibility, while incorporating the "quintessential realist strategy" (Pawson, 2006 p95) in that it focused on contextual constrains on the workings of intervention mechanisms. As argued also by van der Knaap and colleagues (2008), unpacking the 'black box' of potentially effective interventions helps to identify what facilitates the success of interventions and the conditions that are necessary for mechanisms to work. Yet, in line with Kane and colleagues (2010), our realist synthesis results should be seen as a generic set of hypotheses.

4.4 CONCLUSIONS

We conducted a systematic review of the effectiveness of FGM/C abandonment efforts and factors related to FGM/C in accordance with the criteria that the Campbell Collaboration applies to systematic reviews. Additionally, using the realist approach, we delineated the context-mechanism-outcome configurations of the interventions in an attempt to explain how the outcomes of the interventions varied depending on the mechanisms supposed to underpin the interventions and the contexts in which they were used.

Our findings show that much work remains to be conducted regarding the evaluation of FGM/C abandonment efforts. With only eight controlled intervention evaluations meeting our inclusion criteria, all characterised by low methodological quality and no more than two studies synthesised in a meta-analysis for any given outcome, few firm conclusions can be drawn and we encourage intensified research efforts in this area. There is considerable doubt about the validity of the findings, but our results point to possible advantageous developments as a result of FGM/C abandonment interventions, including lower prevalence of FGM/C, and changes in cognitions about FGM/C, such as beliefs about the practice. Similarly, our ability to conclude regarding how factors related to the continuance and discontinuance of FGM/C help explain the effectiveness of interventions was limited. We argue for interventions to address the local communities' core values, beliefs, and general enforcement systems that support FGM/C, which according to our findings vary across contexts. However, our results from 27 included context studies showed that the main factors that supported the practice were tradition, religion, and reduction of women's sexual desire. The main factors identified in this systematic review that hindered FGM/C were harm and medical complications, as well as the prevention of sexual satisfaction.

Based on the context-mechanism-outcome configurations we could formulate tentative recommendations for future programs, some of which were to include a community analysis and development phase before the intervention is launched, to consider integrating the issue of FGM/C in a larger set of community-relevant issues, to establish an alliance with traditional and religious leaders, and to base the intervention on an articulated behaviour change theory.

5 References

Included studies

Abubakar, I. S., Illyasu, Z., Kabir, M., Uzoho, C. C., & Abulkadir, M. B. (2004). Knowledge, attitude and practice of female genital cutting among antenatal patients in Aminu Kano Teaching Hospital, Kano. *Nigerian Journal of Medicine, 13*(3), 254-258.

Afifi, M. (2009). Women's empowerment and the intention to continue the practice of female genital cutting in Egypt. *Archives of Iranian Medicine, 12*(2), 154-160.

Aigbodion, A. I., Imhonde, H. O., & Aluede, O. (2004). A pilot study of the attitude of Nigerian University students towards female genital mutilation. *Anthropologist, 6*(4), 247-252.

Allam, M. F., de Irala-Estévez, J., Navajas, R. F., del Castillo, A., S., Hoashi, J. S., Pankovich, M. B., & Liceaga, J. R. (1999). Students' knowledge of and attitudes about female circumcision in Egypt. *The New England Journal of Medicine, 341*(20), 1552-1553.

Allam, M. F., de Irala-Estévez, J., Navajas, R. F., del Castillo, A., S., Hoashi, J. S., Pankovich, M. B., & Liceaga, J. R. (2001). Factors associated with the condoning of female genital mutilation among university students. *Public Health, 115*, 350-355.

Al-Hussaini, T. (2003). Female genital cutting: types, motives and perineal damage in laboring Egyptian women. *Medical Principles and Practice, 12*, 123-128.

Babalola, S., Brasington, A., Agbasimalo, A., Helland, A., Nwanguma, E., & Onah, N. (2006). Impact of a communication programme on female genital cutting in eastern Nigeria. *Tropical Medicine and International Health, 11*, 1594-1603.

Bayoudh, F., Barrak, S., Ben Fredj, N., Allani, R., & Hamdi, M. (1995). Etude d'une coutume en Somalie: la circoncision des filles/Study of a common practice in Somalia: female circumcision. *Médicine Tropicale, 55*,238-242.

Bop, C. (2001). Female genital mutilation data from Senegal. *Echo, 6*, 5-10.

Briggs, L. A. (2002). Male and female viewpoints on female circumcision in Ekpeye, Rivers State, Nigeria. *African Journal of Reproductive Health, 6*(3), 44-52.

Burkina Faso Demographic and Health Survey (Institut National de la Statistique et de la Demographie Ministere de l'Economie et des Finances Ouagadougou, Burkina

Faso & Macro International Inc. Calverton, Maryland USA). (2000). *Burkina Faso: DHS, 1998-99 - Final report*. Calverton, MD: Measure DHS.

Chege, J., Askew, I., Igras, S., & Mutesh, J. K. (2004). Testing the effectiveness of integrating community-based approaches for encouraging abandonment of female genital cutting into CARE's reproductive health programs in Ethiopia and Kenya. Washington, D.C.: Population Council.

Dandash, K. F., Refaat, A. H., & Eyada, M. (2001a). Female genital mutilation: a descriptive study. *Journal of Sex & Marital Therapy, 27*, 453-458.

Dandash, K. F., Refaat, A. H., & Eyada, M. (2001b).Female genital mutilation: a prospective view. *Journal of Sex & Marital Therapy, 27*, 459-464.

Dare, F. O., Oboro, V. O., Fadiora, S. O., Orji, E. O., Sule-Odu, A. O., & Olabode, T. O. (2004). Female genital mutilation: an analysis of 522 cases in south-western Nigeria. *Journal of Obstetrics and Gynaecology, 24*(3), 281-283.

Dehne, K. L., Wacker, J., Nadembega, J., & Ira, R. (1997). Female genital mutilation in the North of Burkina Faso. *curare, 20*, 221-242.

Diop, N. J., & Askew, I. (2009). The effectiveness of a community-based education program on abandoning female genital mutilation/cutting in Senegal. *Studies in Family Planning, 40*(4), 307-318.

Diop, N. J., Badge, E., Ouoba, D., & Melching, M. (2003). Replication of the TOSTAN programme in Burkina Faso. How 23 villages participated in a human right-based education programme and abandoned the practice of female genital cutting in Burkina Faso. Mwangaza Action Association.

Diop, N. J., Faye, M. M., Moreau, A., Cabral, J., & Benga, H. (2004). The TOSTAN Program. Evaluation of a community based education program in Senegal. New York: Population Council.

Diop, N. J., Moreau, A., & Benga, H. (2008a). Evaluation of the long-term impact of the TOSTAN programme on the abandonment of FGM/C and early marriage: results from a qualitative study in Senegal. Dakar, Senegal: Population Council.

Diop, N.J., Traoré, F., Diallo, H., Traoré, O., Touré, A. H., Diallo, Y., et al. (1998). Study of the effectiveness of training Malian social and health agents in female genital cutting issues and in educating their clients/ Etude de l'efficacité de la formation du personnel socio-sanitaire dans l'educacion des client(e)s sur l'excision et dans le traitemaent de ses complications au Mali. Bamako, Mali: Population Council.

Easton, P., Miles, R., & Monkman, K. (2002). Final report on the evaluation of the Tostan/IEP village empowerment program pilot project in the Republic of Mali. Florida State University, Tallahassee, Florida.

Egypt Demographic and Health Survey (Fatma El-Zanaty, Ann A. Way, Ministry of Health and Population, National Population Council, Cairo, Egypt & ORC Macro,

Calverton, Maryland, USA). (2000). *Egypt: DHS, 2000 – Final report*. Calverton, MD: Measure DHS.

El-Gibaly, O., Ibrahim, B., Mensch, B. S., & Clark, W. H. (2002). The decline of female circumcision in Egypt: evidence and interpretation. *Social Science & Medicine, 54*, 205-220.

Freymeyer, R. H., & Johnson, B. E. (2007). An exploration of attitudes toward female genital cutting in Nigeria. *Population Research and Policy Review, 26*, 69-83.

Kandala, N., Nwakeze, N., & Kandala, S. N. I. I. (2009). Spatial distribution of female genital mutilation in Nigeria. *American Journal of Tropical Medicine and Hygiene, 81*(5), 784-792.

Mali Demographic and Health Survey. (1996). *Mali: DHS, 1995-96 - Final report*. Calverton, MD: Measure DHS.

Monkman, K., Miles, R., & Easton, P. (2007). The transformatory potential of a village empowerment program: The Tostan replication in Mali. *Women's Studies International Forum, 30*, 451-464.

Mounir, G. M., Mahdy, N. H., & Fatohy, I. M. (2003). Impact of health education program about reproductive health on knowledge and attitude of female Alexandria University students. *Journal of the Egypt Public Health Association, 78*, 433-466.

Nigeria Demographic and Health Survey. (2000). *Nigeria: DHS, 1999 - Final report*. Calverton, MD: Measure DHS.

Odimegwu, C., Okemgbo, C., & Luqman, B. (1998). Female genital mutilation in Nigeria: What do adult and adolescent men know? *Journal of the Nigerian Anthropological and Sociological Association, 1-2*, 97-108.

Odimegwu, C., Modupeola, O., & Okemgbo, C. N. (2001).Tradition and health: the predicament of female adolescents among the Igbo. *Journal of Cultural Studies, 3*(1), 284-300.

Odimegwu, C., & Okemgbo, C. N. (2003). Gender role ideologies and prevalence of violence against women in Imo State, Nigeria. *Anthropologist, 5*(4), 225-236.

Okemgbo, C. N., Omideyi, A. K., & Odimegwu, C. O. (2002). Prevalence, patterns and correlates of domestic violence in selected Igbo communities of Imo State, Nigeria. *African Journal of Reproductive Health, 6*(2), 101-114.

Osifo, D. O., & Evbuomwan, I. (2009). Female genital mutilation among Edo people: the complications and pattern of presentation at a pediatric surgery unit, Benin City. *African Journal of Reproductive Health, 13*(1), 17-25

Ouoba, D., Congo, Z., Diop, N. J., Melching, M., Banza, B., Guiella, G., & Baumbarten, I. (2004). Experience from a community based education program in Burkina Faso, the Tostan Program. Washington, D.C.: Population Council.

Ragab, A. R. A., Mahmoud, M., & Serour, G. (2000). Decisions regarding harmful practices (female genital cutting (FGC), and virginity testing) and the role of husbands. *Africa Regional Sexuality Resource Centre*. Available online http://www.arsrc.org/downloads/features/ragab009.pdf

Snow, R. C., Slanger. T. E., Okonofua, F. E., Oronsaye, F., & Wacker, J. (2002). Female genital cutting in southern urban and peri-urban Nigeria: self-report validity, social determinants and secular decline. *Tropical Medicine and International Health, 7*(1), 91-100.

Ugboma, H. A. A., Akani, C. I., & Babatunde, S. (2004). Prevalence and medicalization of female genital mutilation. *Nigerian Journal of Medicine, 13*(3), 250-253.

United Nations Children's Fund (UNICEF). (2008). Long-term evaluation of the Tostan programme in Senegal: Kolda, Thiès and Fatick regions. Working paper. New York: UNICEF.

Yount, K. M. (2002). Like mother, like daughter? Female genital cutting in Minia, Egypt. *Journal of Health and Social Behavior, 42*(3), 336-358.

Yount, K. M. (2004). Symbolic gender politics, religious group identity, and the decline in female genital cutting in Minya, Egypt. *Social Forces, 82*(3), 1063-1090.

Excluded studies considered in full text

Abdel-Tawab, N., & Hegazi, S. (2000). Critical analysis of interventions against FGC in Egypt. Population Council.

Abdi, M. S., & Askew, I. (2008). A religious oriented approach to address female genital mutilation/cutting among the Somali community of Wajir, in Kenya. Frontiers in Reproductive Health.

Abdi, M. S., Jaldesa, G., & Askew, I. (2008). Managing and preventing female genital mutilation/cutting (FGM/C) among the Somali community in Kenya. Frontiers in Reproductive Health.

Adeokun, L. A., Oduwole, M., Oronsaye, F., Gbogboade, A. O., Aliyu, N., Wumi, A., et al. (2006). Trends in female circumcision between 1933 and 2003 in Osun and Ogun States, Nigeria (a cohort analysis). *African Journal of Reproductive Health, 10*, 48-56.

Adinma, J. I., & Agbai, A. O. (1999). Practice and perceptions of female genital mutilation among Nigerian Igbo women. *Journal of Obstetrics and Gynaecology, 19*, 44-48.

Afifi, M. (2009). Wealth index association with gender issues and the reproductive health of Egyptian women. *Nursing and Health Sciences, 11*, 29-36.

Afifi, M., & von Bothmer, M. (2007). Egyptian women's attitudes and beliefs about female genital cutting and its association with childhood maltreatment. *Nursing and Health Sciences, 9*, 270-276.

Ahlberg, B. M., Njau, W., Kiiru, K., & Krantz, I. (2000). Gender masked or self-inflicted pain: Female circumcision, eradication and persistence in central Kenya. *African Sociological Review, 4*, 35-54.

Akinsete, K. (1998). Female circumcision in Yorubaland: Its nature, practice and persistence among the people of Iwo. *West African Journal of Archaeology, 28*, 53-68.

Amusan, O. A., & Asekun-Olarinmoye, E. O. (2006). Knowledge, beliefs, and attitudes to female genital mutilation (FGM) in Shao community of Kwara State, Nigeria. *International Quarterly of Community Health Education, 27*, 337-349.

Asekun-Olarinmoye, E. O., & Amusan, O. A. (2008). The impact of health education on attitudes towards female genital mutilation (FGM) in a rural Nigerian community. *The European Journal of Contraception and Reproductive Health Care, 13*(3), 289-297.

Berhane, Y., Gossaye, Y., Emmelin, M., & Hogberg, U. (2001). Women's health in a rural setting in societal transition in Ethiopia. *Social Science and Medicine, 53*, 1525-1539.

Bitong, L. (2005). Fighting genital mutilation in Sierra Leone. Bulletin of the World Health Organisation, 83, 806-807.

Boyle, E. H., McMorris, B. J., & Gomez, M. (2002). Local conformity to international norms: The case of female genital cutting. *International Sociology, 17*, 5-33.

Caldwell, J. C., Orubuloye, I. O., & Caldwell, P. (2000). Female genital mutilation: Conditions of decline. *Population Research and Policy Review, 19*, 233-254.

Carr, D. (1997). Female genital cutting: Findings from the Demographic and Health Surveys program. Calverton, MD: Macro International.

Carr, D. (2001). Female genital cutting: Findings from the Demographic and Health Surveys program. Calverton, MD: Macro International.

Chege, J. N., Askew, I., & Liku, J. (2001). An assessment of the alternative rites approach for encouraging abandonment of female genital mutilation in Kenya. New York: Population Council.

Creel, L. (2001). Abandoning female genital cutting: Prevalence, attitude and efforts to end the practice. Washington, D.C.: Population Reference Bureau.

Diagne, A. F. (2008). Lutte contre la pratique de l'excision au Mali. De l'approche santé à l'approche basée sur les droits de l'enfant. Rapport d'Evaluation du programme du Centre Djoliba. Senegal: Population Council.

Diop, N., Congo, Z., Ouédraogo, A., Sawadogo, A., Saloucou, L., & Tamini, I. (2008b). Analysis of the evolution of the practice of female genital mutilation/cutting in Burkina Faso. New York: Population Council.

Easton, P., Miles, R., & Monkman, K. (2004). Final report on the evaluation of the Tostan/Entishar Society Village Empowerment program pilot project in the republic of Sudan. Tallahassee: Florida State University.

Ehigiegba, A. E., Selo-Ojeme, D. O., & Omorogbe, F. I. (1998). Female circumcision and determinants in southern Nigeria. *East African Medical Journal, 75*, 374-376.

Ekenze, S. O., Ezegwui, H. U., & Adiri, C. O. (2007). Genital lesions complicating female genital cutting in infancy: a hospital-based study in south-east Nigeria. *Annals of Tropical Paediatrics, 27*, 285-290.

Ericksen, K. P. (1989). Female genital mutilations in Africa. *Behavior Science Research, 23*, 182-184.

Ericksen, K. P., & Hodge, R. W. (1991). Occupations and ritual in contemporary Egypt. *Journal of Quantitative Anthropology, 3*, 207-227.

Feyi-Waboso, P., & Akinbiyi, A. (2006). Knowledge of, attitudes about, and practice of female genital cutting in antenatal patients among Igbos in Nigeria. *Journal of Gynecologic Surgery, 22*, 89-95.

Ghadially, R. (1992). All for "Izzat". *Newsletter: Women's Global Network on Reproductive Rights*, 7-8.

Gordon, D. (1991). Female circumcision and genital operations in Egypt and the Sudan: A dilemma for medical anthropology. *Medical Anthropology Quarterly (New Series), 5*, 3-14.

Gosselin, C. (2000). Feminism, anthropology and the politics of excision in Mali: Global and local debates in a postcolonial world. *Anthropologica, XLII*, 43-60.

Hassanin, I. M. A., Saleh, R., Bedaiwy, A. A., Peterson, R. S., & Bedaiwy, M. A. (2008). Prevalence of female genital cutting in Upper Egypt: 6 years after enforcement of prohibition law. *Reproductive Biomedicine Online, 16*, 27-31.

Hayford, S. R. (2005). Conformity and change: community effects on female genital cutting in Kenya. *Journal of Health and Social Behavior, 46*, 121-140.

Hernlund, Y., & Shell-Duncan, B. (2007). Contingency, context, and change: Negotiating female genital cutting in the Gambia and Senegal. *Africa Today, 53*, 42-57.

Igwegbe, A. O., & Egbuonu, I. (2000). The prevalence and practice of female genital mutilation in Nnewi, Nigeria: the impact of female education. *Journal of Obstetrics and Gynaecology, 20*, 520-522.

Jackson, E. F., Akweongo, P., Sakeah, E., Hodgson, A., Asuru, R., & Phillips, J. F. (2003). Inconsistent reporting of female genital cutting status in northern Ghana:

explanatory factors and analytical consequences. *Studies in Family Planning, 34*, 200-210.

Kamau, A., Bornemann, R., & Laaser, U. (2006). Psychosocial influences on adolescent sexuality and identity in rural Kenya. *Health Sociology Review, 15*(3), 305-316.

Kipuri, N. (2004). Female genital mutilation. *Indigenous Affairs, 1-2*, 22-27.

Mandara, M. U. (2004). Female genital mutilation in Nigeria. *International Journal of Gynaecology and Obstetrics, 84*, 291-298.

Masho, S. W., & Matthews, L. (2009). Factors determining whether Ethiopian women support continuation of female genital mutilation. *International Journal of Gynecology and Obstetrics*, *107*, 232-235.

Meskal, F. H., Kefene, H., Sellassie, A. H., & Khodakevich, L. (1994). A survey of harmful traditional practices in Ethiopia. *Proceedings of the eleventh International conference of Ethiopian studies, Addis Ababa*, April 1-6 1991: volume II, p. 494-506.

Missailidis, K., & Gebre-Medhin, M. (2000). Female genital mutilation in eastern Ethiopia. *Lancet*, *356*, 137-138.

Mitike, G., & Deressa, W. (2009). Prevalence and associated factors of female genital mutilation among Somali refugees in eastern Ethiopia: a cross-sectional study. *BMC Public Health, 9*, 264.

Mostafa, S. R. A., El Zeiny, N. A. M., Tayel, S. E. S., & Moubarak, E. I. (2006). What do medical students in Alexandria know about female genital mutilation? *Eastern Mediterranean Health Journal*, *12* Suppl 2, S78-S92.

Nelson, D., & Newman, C. (2002). Mali: female genital cutting. Advocacy, counselling, support. Provider training and community outreach to eliminate harmful practice. Results review/ Mali: Mutilation génitale féminine. Plaidoyer, counseling, appui. Formation des fournisseurs et initiatives communautaires en vue de l'élimination d'une pratique nocive. Examen des resultants. Chapel Hill, North Carolina, University of North Carolina at Chapel Hill, Program for International Training in Health. PRIME Project.

Nkwo, P. O., & Onah, H. E. (2001). Decrease in female genital mutilation among Nigerian Ibo girls. *International Journal of Gynaecology and Obstetrics, 75*, 321-322.

Obi, S. N. (2004). Female genital mutilation in south-east Nigeria. *International Journal of Gynaecology and Obstetrics, 84*, 183-184.

Odu, B. K. (2008). The attitude of undergraduate female students toward genital mutilation in a Nigerian university. *Research Journal of Medical Sciences, 2*(6), 295-299.

Ogunlola, I. O., Orji, E. O., & Owolabi, A. T. (2003). Female genital mutilation and the unborn female child in southwest Nigeria. *Journal of Obstetrics and Gynaecology, 23*, 143-145.

Olenick, I. (1998). Female circumcision is nearly universal in Egypt, Eritrea, Mali and Sudan. *International Family Planning Perspectives, 24*(1), 47-49.

Orji, E. O., & Babalola, A. (2006). Correlates of female genital mutilation and its impact on safe motherhood. *Journal of the Turkish German Gynecology Association Artemis, 7*, 319-324.

Ouedraogo, S. (2008). Social effect and female genital mutilation (FGM). Unpublished.

Popoola, B. I. (2007). Prevalence of female genital mutilation and parents' attitude among the Yorubas in western Nigeria. *African Journal for Physical, Health Education, Recreation and Dance, 13*(1), 17-31.

Rahlenbeck, S. I., & Mekonnen, W. (2009). Growing rejection of female genital cutting among women of reproductive age in Amhara, Ethiopia. *Culture, Health & Sexuality, 11*(4), 443-452.

Refaat, A., Dandash, K. F., el Defrawi, M. H., & Eyada, M. (2001). Female genital mutilation and domestic violence among Egyptian women. *Journal of Sex and Marital Therapy, 27*, 593-598.

Refaat, A. H., Dandash, K. F., Lotfy, G., & Eyada, M. (2001). Attitudes of medical students towards female genital mutilation. *Journal of Sex and Marital Therapy, 27*, 589-591.

Shabana, K. A., El-Shiek, M., El-Nazer, M., & Samir, N. (2003). Women's perceptions and practices regarding their rights to reproductive health. *Eastern Mediterranean Health Journal, 9*(3), 296-308.

Shell-Duncan, B., & Hernlund, Y. (2006). Are there "Stages of Change" in the practice of female genital cutting?: Qualitative research findings from Senegal and the Gambia. *African Journal of Reproductive Health, 10*(2), 57-71.

Spadacini, B., & Nichols, P. (1998). Campaigning against female genital mutilation in Ethiopia using popular education. *Gender and Development, 6*, 44-52.

Ssali, S. N. (2009). Law, circumcision and gendered sexuality in eastern Uganda and western Kenya. *Eastern African Journal of Peace & Human Rights, 15*(1), 128-157.

Suzuki, C., & Meekers, D. (2008). Determinants of support for female genital cutting among ever-married women in Egypt. *Global Public Health, 3*, 383-398.

Tag-Eldin, M. A., Gadallah, M. A., Al Tayeb, M. N., Abdel-Aty, M., Mansour, E., & Sallem, M. (2008). Prevalence of female genital cutting among Egyptian girls. *Bulletin of the World Health Organization, 86*, 269-274.

Other references

Abdal Magied, A., & Musa, S. (2004). Psycho-sexual effects of female genital mutilation on Sudanese men. *Ahfad Journal: Women and Change, 21*(1), 18-28.

Abraham, C., & Michie, S. (2008). A taxonomy of behaviour change techniques used in interventions. *Health Psychology, 27,* 379-387.

Akweongo, P., Jackson, E., Phillips, J., Adongo, R., Naik, R., Sakeah, E., et al. (forthcoming). Reducing female genital cutting through community mobilization: The impact of a field experiment in northern Ghana.

Almroth, L., Almroth-Berggren, V., Hassanein, O. M., Al-Said, S. S. E., Hasan, S. S. A., Lithell, U., et al. (2001). Male complications of female genital mutilation. *Social Science & Medicine, 53*, 1455-1460.

Armstrong, R., Waters, E., & Doyle, J. *Reviews in health promotion and public health*. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions . The Cochrane Collaboration, 2009.

Askew, I. (2005). Methodological issues in measuring the impact of interventions against female genital cutting. *Culture, Health & Sexuality, 7*(5), 463-477.

Berg, R. C., & Denison, E. (2011). Does female genital mutilation/cutting affect women's sexual functioning? A systematic review of the sexual consequences of FGM/C. Sexuality Research and Social Policy, DOI: 10.1007/s13178-011-0048-z

Berg, R., Denison, E., & Fretheim, A. (2010a). Psychological, social and sexual consequences of female genital mutilation/cutting (FGM/C): a systematic review of quantitative studies. Report from Kunnskapssenteret nr 13–2010. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.

Berg, R., Denison, E., & Fretheim, A. (2010b). Factors promoting and hindering the practice of female genital mutilation/cutting (FGM/C). Report from Kunnskapssenteret nr 23-2010. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.

Boyle, M. (1998). Guidelines for evaluating prevalence studies. *Evidence Based Mental Health, 1,* 37-39.

Briggs, M., & Flemming, K. (2007). Living with leg ulceration: a synthesis of qualitative research. *Journal of Advanced Nursing, 59,* 319-28.

Bullough, V. (1995). *Sexuality and religion*. In: Diamant L, McNulty RD, editors. Psychology of sexual orientation, behavior and identity (444-456). Westport: Greenwood.

Burkina Faso Demographic and Health Survey (Institut National de la Statistique et de la Demographie, Ministere de l'economie et du Developpement, Ouagadougou, Burkina Faso, & ORC Macro, Calverton, Maryland, USA). (2004). *Burkina Faso: DHS, 2003 - Final report*. Calverton, MD: Measure DHS.

Critical Appraisal Skills Programme (CASP). (2006). 10 questions to help you make sense of qualitative research. National Health Services, Public Health Resource Unit. http://www.phru.nhs.uk/Doc_Links/Qualitative%20Appraisal%20Tool.pdf

Denison, E., Berg, R. C., Jensen, J. O., & Lewin, S. (2011). Interventions to reduce the prevalence of female genital mutilation/cutting in African countries. Protocol. Campbell Collaboration.

Denison, E., Berg, R. C., Lewin, S., & Fretheim, A. (2009). Effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting. Report from Kunnskapssenteret nr 25–2009. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.

Dieleman, M., Gerretsen, B., & van der Wilt, G. J. (2009). Human resource management interventions to improve health workers' performance in low and middle income countries: a realist review. *Health Research Policy and Systems, 7*(7).

Dixon-Woods, M., Bonas, S., Booth, A., Jones, D., Miller, T., Sutton, A., Shaw, R., Smith, J., & Young, B. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research, 6*(1), 27-44.

Egypt Demographic and Health Survey (Fatma El-Zanaty, Ann Way & Macro International Inc. Calverton, Maryland, USA). (2008). *Egypt: DHS, 2008 - Final report*. Calverton, MD: Measure DHS.

Feldman-Jacobs, C., & Ryniak, S. (2007). Abandoning female genital mutilation/cutting: an in-depth look at promising practices. Washington D.C.: Population Reference Bureau.

Gali, M. A. (1998). *Female circumcision: A transcultural study of attitudes, identity and reproductive health of East African immigrants*. Dissertation Abstracts International: Section B: The Sciences and Engineering 1998;58(11-B):May-B.

Glanz, K., Rimer B. K., & Lewis, F. M. (2002). Health behaviour and health education. Theory, research and practice. (3rd ed). San Francisco: Jossey-Bass.

Greenhalg, T., Kristjansson, E., & Robinson, V. (2007). Realist review to understand the efficacy of school feeding programmes. *BMJ, 335,* 858-861.

GTZ. (2007). Available from http://www.gtz.de/en/dokumente/en-fgm-countries-ethiopia.pdf

Harden, A., Garcia, J., Oliver, S., Rees, R., Shepherd, J., Brunton, G., & Oakley, A. (2004). Applying systematic review methods to studies of people's views: an example from public health research. *Journal of Epidemiology and Community Health, 58,* 794-800.

Higgins, J. P. T., & Green, S. (2009). *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.0.2*. The Cochrane Collaboration.

HRP. (2006). The UNDP/UNFPA/WHO/World bank special programme of research, development and research training in human reproduction. Progress in sexual and reproductive health, Newsletter 72. Geneva: World Health Organization.

Johansen, R. E. B. (2006). *Experiences and perceptions of pain, sexuality, and childbirth: A study of female genital cutting among Somalis in Norwegian exile, and their health providers.* Institute of general practice and social medicine, University of Oslo.

Kane, S., Gerretsen, B., Scherpbier, R., Dal Poz, M., & Dieleman, M. (2010). A realist synthesis of randomised controlled studies involving use of community health workers for delivering child health interventions in low and middle income countries. *BMC Health Services Research, 10*(286).

Kassamali, N. J. (1998). *When modernity confronts traditional practices: female genital cutting in northeast Africa* (39-62). In: Bodin HL, Tohidi N, editors. Women in Muslim societies. Diversity within unity. London: Lynne Rienner Publishers.

Lannon, C. M. (2000). Circumcision -The debates goes on. *Pediatrics, 105*(3), 681-4.

Mackie, G. (1996). Ending footbinding and infibulation: A convention account. *American Sociological Review, 61*(6), 999-1017.

Mali Demographic and Health Survey. (2007). *Mali: DHS, 2006 - Final report*. Calverton, MD: Measure DHS.

Mckenzie, J. F., & Smeltzer, J. L (2001). *Planning, implementing, and evaluating health promotion programs. A primer* (3rd ed). Boston: Allyn and Bacon.

Muteshi, J., & Sass, J. (2005). Female genital mutilation in Africa: an analysis of current abandonment approaches. Nairobi: PATH.

Nigeria Demographic and Health Survey. (2009). *Nigeria: DHS, 2008 - Final report*. Calverton, MD: Measure DHS.

Oakey, A., Fullerton, D., Holland, J. (1995). Sexual health education interventions for young people: a methodological review. *BMJ 310*, 158-162.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: an introduction*. ESRC Research Methods Programme. Available online http://www.ccsr.ac.uk/methods/publications/documents/RMPmethods2.pdf

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services and Research Policy, 10*(1) Suppl1, 21-34.

Pawson, R. (2006). *Evidence-based policy. A realist perspective*. London: Sage Publications.

Petrosino, A., Boruch, R. F., Soydan, H., Duggan, L., & Sanchez-Meca, J. (2001). Meeting the challenges of evidence-based policy: the Campbell Collaboration. *The Annals of the American Academy, 578,* 14-34.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences*. A practical guide. Oxford: Blackwell Publishing.

Population Reference Bureau (PRB). (2006). Abandoning female genital mutilation/cutting: An in-depth look at promising practices. Available online http://www.prb.org/pdf07/FGM-C_Report.pdf

PRB. (2008). Abandoning female genital mutilation/cutting: data and trends. Washington DC.

Roald, A. S. (2001). *Women in Islam: the Western experience*. London: Routledge.

Shepherd, J., Harden, A., Rees, R., Brunton, G., Garcia, J., Oliver, S., & Oakley, A. (2006). Young people and healthy eating: a systematic review of research on barriers and facilitators. *Health Education Research, 21*(2), 239-257.

Smith, L. K., Pope, C., & Botha, J. L. (2005). Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis. *Lancet, 366*(9488), 825-31.

Talle, A. (2003). *Om kvinneleg omskjering*. Trondheim: Samlaget.

Thierfelder, C., Tanner, M., & Bodiang, C. M. K. (2005). Female genital mutilation in the context of migration: Experience of African women with the Swiss health care system. *European Journal of Public Health, 15*(1), 86-90.

Thomas, H. (No date). Quality assessment tool for quantitative studies. Effective Public Health Practice Project. Toronto: McMaster University.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*, 45.

United Nations Population Fund (UNFPA). (2007). *A holistic approach to the abandonment of female genital mutilation/cutting*. New York: United Nations.

United Nations Children's Fund (UNICEF). (2004). UNICEF global consultation on indicators. Child protection indicators framework. Female genital mutilation and cutting. New York: UNICEF.

UNICEF. (2005a). Female genital mutilation/female genital cutting: A statistical report. New York: UNICEF.

UNICEF. (2005b). Changing a harmful social convention: Female genital mutilation/cutting. Florence: UNICEF.

UNICEF. (2010). Media backgrounder: Female genital mutilation/cutting (FGM/C) in Egypt. Retrieved from http://www.unicef.org/egypt/media_4115.html

United Nations Development Program (UNDP). (2010). Human development index. Available online http://www.beta.undp.org/undp/en/home.html

van der Knaap, L. M., Leeuw, F. L., Bogaerts, S., & Nilssen, L. T. J. (2008). Combining Campbell standards and the realist evaluation approach. The best of two worlds? *American Journal of Evaluation, 29*(1), 48-57.

Wheeler, P. (2003). The role of the law. *International Journal of Children's Rights, 11,* 257-271.

World Health Organization (WHO). (1997). A Joint WHO/UNICEF/UNFPA Statement. Female genital mutilation. Geneva: World Health Organization.

WHO. (1999). Female genital mutilation. Programmes to date: what works and what doesn't? A review. Geneva: World Health Organization.

WHO. (2000). A systematic review of the health complications of female genital mutilation including sequelae in childbirth. Geneva: World Health Organization.

WHO. (2008). Eliminating female genital mutilation: An interagency statement. Geneva: World Health Organization.

WHO study group. (2006). Female genital mutilation and obstetric outcome: WHO collaborative prospective study in six African countries. *Lancet,* 367:1835-1841.

Yoder, S., Abderrahim, N., & Zhuzhuni, A. (2004). Female genital cutting in the Demographic and Health Surveys: A critical and comparative analysis. DHS Comparative Reports No. 7. USA, Maryland: ORC Macro, Calverton.

Yoder, S., & Kahn, S. (2008). Numbers of women circumcised in Africa: the production of a total. DHS working papers. USA: United States Agency for International Development.

Appendix 1: Explanation of terms

CASP	Critical Appraisal Skills Programme
DHS	Demographic and Health Survey
EPPI	Evidence for Policy and Practice Information and Co-ordinating Centre
FGM/C	Female genital mutilation/cutting
HRP	Human Reproduction Programme (part of WHO)
IEC	Information, Education, Communication
PRB	Population Reference Bureau
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
WHO	World Health Organisation

Appendix 2: Search strategies

African Index Medicus

Search: Sari Ormstad

Date: 05.02.2009; updated 19.02.2010 and 03.03.2011

Retrieval: 12+12+13

Strategy:

"CIRCUMCISION" [Descriptor] or "CIRCUMCISION, FEMALE" [Descriptor] or "INFIBULATION" [Descriptor]

Anthropology Plus

Search: Hege Oswald

Date: 18.02.2009

Retrieval: 200

Strategy:

((kw: female* or kw: wom#n) or kw: girl*) and ((kw: mutilation* or kw: infibulate*) or kw: cutting*) or kw: infibulate*

British Nursing Index and Archive

Search: Sari Ormstad

Date: 04.02.2009; updated 20.01.2010 and 02.03.2011

Retrieval: 135+153+167

Strategy:

1. Circumcision/
2. ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.
3. "fgm/c".tw.
4. ((removal$ or alteration$ or excision$) adj6 female genital$).tw.
5. pharaonic circumcision$.tw.

6. sunna.tw.
7. (clitoridectom$ or clitorectom$).tw.
8. (infibulate$ or reinfibulat$ or deinfibulat$).tw.
9. or/1-8

The Cochrane Library: CENTRAL, CDR, DARE

Search: Sari Ormstad
Date: 04.02.2009; updated 21.01.2010 and 02.03.2011
Retrieval: 3+6+10
Strategy:

#1 MeSH descriptor Circumcision, Female, this term only

#2 ((female* or woman or women or girl or girls) near/3 (mutilation* or circumcis* or cutting*)) or "fgm/c" or ((removal* or alteration* or excision*) near/6 (female next genital*)) or (pharaonic next circumcision*) or sunna or clitoridectom* or clitorectom* or infibulat* or reinfibulat* or deinfibulat*:ti or ((female* or woman or women or girl or girls) near/3 (mutilation* or circumcis* or cutting*)) or "fgm/c" or ((removal* or alteration* or excision*) near/6 (female next genital*)) or (pharaonic next circumcision*) or sunna or clitoridectom* or clitorectom* or infibulat* or reinfibulat* or deinfibulat*:ab

#3 (#1 OR #2)

EMBASE

Search: Sari Ormstad
Date: 04.02.2009; updated 20.01.2010 and 02.03.2011
Retrieval: 570+60+125
Strategy:

1. female circumcision/ or female genital mutilation/ or female genital cutting/ or infibulate105ns/
2. ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.
3. "fgm/c".tw.
4. ((removal$ or alteration$ or excision$) adj6 female genital$).tw.
5. pharaonic circumcision$.tw.
6. sunna.tw.
7. (clitoridectom$ or clitorectom$).tw.
8. (infibulate$ or reinfibulat$ or deinfibulat$).tw.
9. or/1-8

EPOC

Search: Sari Ormstad
Date: 09.02.09
Retrieval: 0
Strategy:
Title or Abstract or Keyword: circumcise% or mutilation% or FGM or clitoridectom% or clitorectom% or infibulate% or deinfibulat% or de-infibulat% or reinfibulat% or re-infibulat% or female genital%

MEDLINE
Search: Sari Ormstad
Date: 04.02.2009; updated 20.01.2010 and 02.03.2011
Retrieval: 1100+79+103
Strategy:
1. Circumcision, Female/
2. ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.
3. "fgm/c".tw.
4. ((removal$ or alteration$ or excision$) adj6 female genital$).tw.
5. pharaonic circumcision$.tw.
6. sunna.tw.
7. (clitoridectom$ or clitorectom$).tw.
8. (infibulate$ or reinfibulat$ or deinfibulat$).tw.
9. or/1-8

PILOTS
Search: Sari Ormstad
Date: 04.02.2009; updated 25.01.2010 and 02.03.2011
Retrieval: 14+15+17
Strategy:
((DE=("genital mutilation")) or (TI=(((female* or woman or women or girl or girls) within 3 (mutilation* or infibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic circumcision* or sunna or clitoridectom* or clitorectom* or 106nfibulate* or reinfibulat* or deinfibulat*)) or (AB=(((female* or woman or women or girl or girls) within 3 (mutilation* or infibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic circumcision* or sunna or clitoridectom* or clitorectom* or infibulate* or reinfibulat* or deinfibulat*)))

POPLINE
Search: Sari Ormstad
Date: 05.02.2009; updated 03.03.2011
Retrieval: 1566+105

Strategy:
POPLINE Advanced Search KEYWORDS:
FEMALE GENITAL CUTTING

PsychINFO
Search: Sari Ormstad
Date: 04.02.2009; updated 21.01.2010 and 02.03.2011
Retrieval: 401+60+63
Strategy:
1. Circumcision/
2. ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.
3. "fgm/c".tw.
4. ((removal$ or alteration$ or excision$) adj6 female genital$).tw.
5. pharaonic circumcision$.tw.
6. sunna.tw.
7. (clitoridectom$ or clitorectom$).tw.
8. (infibulate$ or reinfibulat$ or deinfibulat$).tw.
9. or/1-8

Social Services Abstracts
Search: Sari Ormstad
Date: 04.02.2009; updated 25.01.2010 and 02.03.2011
Retrieval: 40+39+49
Strategy:
((DE=("circumcision" or "genital mutilation")) or (TI=(((female* or woman or women or girl or girls) within 3 (mutilation* or 107nfibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic circumcision* or sunna or clitoridectom* or clitorectom* or infibulate* or reinfibulat* or deinfibulat*)) or (AB=(((female* or woman or women or girl or girls) within 3 (mutilation* or infibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic circumcision* or sunna or clitoridectom* or clitorectom* or infibulate* or reinfibulat* or deinfibulat*)))

Sociological Abstracts
Search: Sari Ormstad
Date: 04.02.2009; updated 25.01.2010 and 02.02.2011
Retrieval: 325+334+346
Strategy:
((DE=("circumcision" or "genital mutilation")) or (TI=(((female* or woman or women or girl or girls) within 3 (mutilation* or infibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic

circumcision* or sunna or clitoridectom* or clitorectom* or infibulate* or reinfibulat* or deinfibulat*)) or (AB=(((female* or woman or women or girl or girls) within 3 (mutilation* or infibulate* or cutting*)) or fgm or ((removal* or alteration* or excision*) within 6 female genital*) or pharaonic circumcision* or sunna or clitoridectom* or clitorectom* or infibulate* or reinfibulat* or deinfibulat*)))

WHOLIS

Search: Sari Ormstad
Date: 05.02.2009; updated 19.02.2010 and 03.03.2011
Retrieval: 65+67+72
Strategy:
words or phrase "((female$ or wom?n or girl or girls) near3 (mutilation$ or infibulate$ or cutting$))"
OR
words or phrase ""fgm/c""
OR
words or phrase "((removal$ or alteration$ or excision$) near6 (female adj genital$))"
OR
words or phrase "(pharaonic adj circumcision$)"
OR
words or phrase "sunna"
OR
words or phrase "(clitoridectom$ or clitorectom$)"
OR
words or phrase "(infibulate$ or reinfibulat$ or deinfibulat$)"

Appendix 3: Excluded studies

5.1 EXCLUDED EFFECTIVENESS STUDIES

We excluded 7 effectiveness studies read in full text. These are listed in table 1 with reasons for exclusion. For the complete citation we refer to the reference list of excluded studies.

Table 1. Excluded effectiveness studies and cause for exclusion

Study	Cause for exclusion of study
Asekun-Olarimoye 2008	There was no comparison group
Bitong 2005	Not a primary effectiveness study
Chege 2001	There was no comparison group
Diop 2008b	Only qualitative data reported, review
Easton 2004	No relevant outcome data
Nelson 2002	There was no comparison group
Spadacini 1998	Not a primary effectiveness study

5.2 EXCLUDED CONTEXT STUDIES

We excluded 56 context studies read in full text. These are listed in table 2 with reasons for exclusion. For the complete citation we refer to the reference list of excluded studies.

Table 2. Excluded context studies and cause for exclusion

Study	Cause for exclusion of study
Abdel-Tawab 2000	Focus not factors related to FGM/C continuance or discontinuance
Abdi 2008	Not primary study with empirical data
Abdi 2009	Qualitative data not analyzed in qualitative way
Adeokun 2006	Focus not factors related to FGM/C continuance or discontinuance
Adinma 1999	Data collected >5 years before intervention occurred
Afifi 2007	Data collected after intervention occurred

Study	Cause for exclusion of study
Afifi 2009	Data collected after intervention occurred
Ahlberg 2000	Participants not from one of included intervention countries/groups
Akinsete 1998	Not primary study with empirical data
Amusan 2006	Data collected after intervention occurred
Asekun-O 2008	Data collected after intervention occurred
Berhane 2001	Focus not factors related to FGM/C continuance or discontinuance
Boyle 2002	Participants from multiple countries
Caldwell 2000	Not primary study with empirical data
Carr 1997	Participants from multiple countries
Carr 2001	Participants from multiple countries
Creel 2001	Participants from multiple countries
Diagne 2008	Not primary study with empirical data
Ehigiegba 1998	Focus not factors related to FGM/C continuance or discontinuance
Ekenze 2007	Focus not factors related to FGM/C continuance or discontinuance
Ericksen 1989	Participants from multiple countries
Ericksen 1991	Focus not factors related to FGM/C continuance or discontinuance
Feyi-Waboso 2006	Data collected after intervention occurred
Ghadialli 1991	Not primary study with empirical data
Gordon 1991	Not primary study with empirical data
Gosselin 1991	Not primary study with empirical data
Hassanin 2008	Data collected after intervention occurred
Hayford 2005	Participants not from one of included intervention countries/groups
Hernlund 2007	Focus not factors related to FGM/C continuance or discontinuance
Igwegbe 2000	Data collected >5 years before intervention occurred
Jackson 2003	Focus not factors related to FGM/C continuance or discontinuance
Kamau 2006	Participants not from one of included intervention countries/groups
Kipuri 2004	Not primary study with empirical data
Mandara 2004	Data collected >5 years before intervention occurred
Masho 2009	Data collected after intervention occurred
Meskal 1994	Focus not factors related to FGM/C continuance or discontinuance
Missailidis 2000	Not primary study with empirical data
Mitike 2009	Data collected after intervention occurred
Mostafa 2006	Data collected after intervention occurred
Nkwo 2001	Focus not factors related to FGM/C continuance or discontinuance
Obi 2004	Focus not factors related to FGM/C continuance or discontinuance

Study	Cause for exclusion of study
Odu 2008	Data collected after intervention occurred
Ogunlola 2003	Focus not factors related to FGM/C continuance or discontinuance
Olenick 1998	Participants from multiple countries
Orji 2006	Data collected after intervention occurred
Ouedraogo 2008	Data collected after intervention occurred
Popoola 2007	Data collected after intervention occurred
Rahlenbeck 2009	Data collected after intervention occurred
Refaat 2001a	Focus not factors related to FGM/C continuance or discontinuance
Refaat 2001b	Data collected >5 years before intervention occurred
Shabana 2003	Focus not factors related to FGM/C continuance or discontinuance
Shell-Duncan 2005	Focus not factors related to FGM/C continuance or discontinuance
Spandacini 1998	Not primary study with empirical data
Ssali 2009	Participants not from one of included intervention countries/groups
Suzuki 2008	Data collected after intervention occurred
Tag-Eldin 2008	Data collected after intervention occurred

Appendix 4: Quality assessment

5.3 QUALITY ASSESSMENT OF THE EFFECTIVENESS STUDIES

The outcome of the quality assessment using the Quality Assessment Tool for Quantitative Studies, developed by McMaster University Effective Public Health Practice Project, is presented in table 3.

Table 3. Quality assessment of effectiveness studies

Study	Selection bias	Study design	Confoun-ders	Blinding	Data collection	Withdraw-als, drop-outs	Final decision
Babalola 2006	Strong	Weak	Weak	Moderate	Weak	Moderate	Weak
Chege 2004a	Strong	Weak	Weak	Moderate	Weak	Strong	Weak
Chege 2004b	Strong	Weak	Weak	Moderate	Weak	Strong	Weak
Diop 2004	Strong	Weak	Moderate	Moderate	Weak	Weak	Weak
Diop 1998	Moderate	Weak	Weak	Moderate	Weak	Strong	Weak
Easton 2002	Moderate	Weak	Weak	Moderate	Weak	Weak	Weak
Mounir 2003	Strong	Weak	Strong	Moderate	Weak	Weak	Weak
Ouoba 2004	Moderate	Weak	Weak	Moderate	Weak	Moderate	Weak

Note: Chege 2004a refers to the study set in Kenya; Chege 2004b refers to the study set in Ethiopia

The McMaster Quality Assessment tool for quantitative studies is reproduced below. For the dictionary which accompanies the tool we refer to Thomas (no date).

A) Selection bias

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population? 1. Very likely 2. Somewhat likely 3. Not likely 4. Can't tell

(Q2) What percentage of selected individuals agreed to participate? 1. 80-100% agreement 2. 60-79% agreement 3. Less than 60% agreement 4. Not applicable 5. Can't tell

Rate this section: 1. Strong 2. Moderate 3. Weak

B) Study design – Indicate the study design

Randomized controlled trial

Controlled clinical trial

Cohort analytic (two groups pre + post)

Case-control

Cohort (one group pre + post (before and after))

Interrupted time series

Other specify _____

Can't tell

Was the study described as randomized? If No, go to Component C. No Yes

If Yes, was the method of randomization described? No Yes

If Yes, was the method appropriate? No Yes

Rate this section: 1. Strong 2. Moderate 3. Weak

C) Confounders

(Q1) Were there important differences between groups prior to the intervention?

1 Yes 2 No 3 Can't tell

The following are examples of confounders:

1 Race, 2 Sex, 3 Marital status/family, 4 Age, 5 Socio-economic status (income or class), 6 Education, 7 Health status, 8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)? 1 80 – 100% 2 60 – 79% 3 Less than 60% 4 Can't Tell

Rate this section: 1. Strong 2. Moderate 3. Weak

D) Blinding

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants? 1 Yes 2 No 3 Can't tell

(Q2) Were the study participants aware of the research question? 1 Yes 2 No 3 Can't tell

Rate this section: 1. Strong 2. Moderate 3. Weak

E) Data collection methods

(Q1) Were data collection tools shown to be valid? 1 Yes 2 No 3 Can't tell

(Q2) Were data collection tools shown to be reliable? 1 Yes 2 No 3 Can't tell

Rate this section: 1. Strong 2. Moderate 3. Weak

F) Withdrawals and drop-outs

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group? 1 Yes 2 No 3 Can't tell

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest). 1 80 -100% 2 60 – 79% 3 less than 60% 4 Can't tell

Rate this section: 1. Strong 2. Moderate 3. Weak

G) Intervention integrity

(Q1) What percentage of participants received the allocated intervention or exposure of interest? 1 80 -100% 2 60 – 79% 3 less than 60% 4 Can't tell

(Q2) Was the consistency of the intervention measured? 1 Yes 2 No 3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results? 1 Yes 2 No 3 Can't tell

H) Analyses

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design? 1 Yes 2 No 3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received? 1 Yes 2 No 3 Can't tell

GLOBAL RATING

Component ratings

Please transcribe the information from the gray boxes on pages 1-4 onto this page.

Rate this section (see dictionary)

A Selection bias	1. Strong 2. Moderate 3. Weak
B Study design	1. Strong 2. Moderate 3. Weak
C Confounders	1. Strong 2. Moderate 3. Weak
D Blinding	1. Strong 2. Moderate 3. Weak
E Data collection methods	1. Strong 2. Moderate 3. Weak
F Withdrawals and drop-outs	1. Strong 2. Moderate 3. Weak

Global rating for this paper (circle one):

1 Strong (four strong ratings with no weak ratings)

2 Moderate (less than four strong ratings and one weak rating)

3 Weak (two or more weak ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings? No Yes If yes, indicate the reason for the discrepancy

1 Oversight 2 Differences in interpretation of criteria 3 Differences in interpretation of study

Final decision of both reviewers (circle one): 1 Strong 2 Moderate 3 Weak

5.4 QUALITY ASSESSMENT OF THE CONTEXT STUDIES

The outcome of the quality assessment of quantitative context studies is presented in table 4.

Table 4. Quality assessment of quantitative context studies

Study	1	2	3	4	5	6	7	Assess-ment
Abubakar 2004	Yes	Yes	No	Unclear	Yes	No	Yes	Moderate
Afifi 2009	Yes	Yes	No	Yes	Yes	No	Yes	High
Aigbodion 2004	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Moderate
Al-Hussaini 2003	Yes	Unclear	No	Unclear	Yes	No	Yes	Low
Allam 1999	Yes	Yes	No	Yes	Yes	No	Yes	High
Bayoudh 1995	Yes	Unclear	No	Unclear	Unclear	No	Yes	Low
Bop 2001	Yes	Unclear	No	Unclear	Unclear	No	Yes	Low
Briggs 2002	Yes	Yes	No	Unclear	Yes	No	Yes	Moderate
Dandash 2001a	Yes	Yes	No	Yes	Yes	No	Yes	High
Dandash 2001b	Yes	Unclear	No	Unclear	Yes	No	Yes	Moderate
Dare 2004	Yes	Yes	No	Unclear	Yes	No	Yes	Moderate
Dehne 1997	Yes	Unclear	No	Unclear	Unclear	No	Yes	Low
DHS BF 1999	Yes	Yes	No	Yes	Yes	No	Yes	High
DHS Egypt 2000	Yes	Yes	No	Yes	Yes	No	Yes	High
DHS Mali 1996	Yes	Yes	No	Yes	Yes	No	Yes	High
DHS Nigeria 1999	Yes	Yes	No	Yes	Yes	No	Yes	High
El-Gibaly 2002	Yes	Yes	Unclear	Unclear	Yes	Unclear	Yes	Moderate
Freymeyer 2007	Yes	Yes	No	Yes	Yes	No	Yes	High
Kandala 2009	Yes	Yes	No	Yes	Yes	No	Yes	High
Odimegwu 1998	Yes	Yes	No	Yes	Unclear	No	Yes	Moderate
Odimegwu 2001	Yes	Yes	No	Yes	Yes	No	Yes	Moderate

Study	1	2	3	4	5	6	7	Assess-ment
Okemgbo 2002	Yes	Yes	No	Unclear	Yes	No	Yes	Moderate
Osifo 2009	Yes	Unclear	No	Unclear	Unclear	No	Yes	Low
Snow 2002	Yes	Unclear	No	Yes	Yes	No	Yes	Moderate
Ugboma 2004	Yes	Unclear	No	Yes	Unclear	No	Yes	Low
Yount 2002	Yes	Yes	No	Unclear	Yes	No	Yes	Moderate

Note: The numbers in the top row refer to the checklist question – the questions are stated below

The seven quality assessment questions for quantitative (cross-sectional) studies:

1. Was the population from which the sample was drawn clearly defined?
2. Was the sample representative of the population?
3. Is it explained whether (and how) the participants who agreed to participate are different from those who refused to participate?
4. Is the response rate adequate?
5. Were standardized data collection methods used?
6. Were measures shown to reliable and valid?
7. Were the statistical methods appropriate?

Each question was answered 'yes', 'unclear/somewhat', or 'no'. The checklist is based on EBM Notebook, Guidelines for evaluating prevalence studies. May 1998, No 2 p 37-9). Description of assessment of study quality:

High quality (few limitations): All or almost all of the criteria from the checklist are met. If some of the criteria are not met, it must be unlikely that the study conclusions will change.

Moderate quality (some limitations): Some of the criteria are not met and/or the study does not adequately address the criteria. It is unlikely that the study conclusions will change.

Low quality (serious limitations): Few or no criteria are met and/or the study does not adequately address the criteria. It is likely that the study conclusions will change.

The outcome of the quality assessment of the qualitative context study using the CASP checklist for qualitative studies is presented in table 5.

Table 5. Quality assessment of the qualitative context study

Study	1	2	3	4	5	6	7	8	9	10	Assess-ment
Ragab 2004	Yes	Yes	Yes	Unclear	Yes	No	Yes	No	Yes	Somewhat	Moderate

We do not have permission to reproduce the CASP checklist questions, but they may be accessed via the following link:
http://www.sph.nhs.uk/what-we-do/public-health-workforce/resources/critical-appraisals-skills-programme
The first two questions are answered 'yes' or 'no' while all other questions are answered 'yes', 'unclear/somewhat', or 'no'. Assessment of study quality as high, moderate or low is the same as for quantitative views studies described above.

Appendix 5: Behaviour change techniques

Across the eight included effectiveness studies, eight different behavior change techniques were applied. The number of behavior change techniques used in one study ranged from one to six (table 6).

Table 6. Summary of behavior change techniques used in effectiveness studies

Study	Country	Techniques							
		1	2	5	6	21	26	29	30
Diop 1998	Mali	X	X	X		X	X		X
Mounir 2003	Egypt		X						
Chege 2004a	Kenya		X			X		X	X
Chege 2004b	Ethiopia		X			X		X	X
Babalola 2006	Nigeria	X	X	X	X			X	
Easton 2002	Mali		X		X		X		X
Diop 2004	Senegal		X						X
Ouoba 2004	Burkina Faso	X							X

Note: Chege 2004a refers to the study set in Kenya; Chege 2004b refers to the study set in Ethiopia.

Numbers along the top row in the table above (table 6) refer to behavior change techniques, as listed and defined by Mackie:

1= Provide information on consequences of behaviour in general: information about the relationship between the behaviour and its possible or likely consequences in the general case, usually based on epidemiological data, and not personalised for the individual (contrast with technique2)

2= Provide information on consequences of behavior to the individual: Information about the benefits and costs...

5= Goal setting (behaviour)

6= Goal setting (outcome)

21= Provide instruction on how to perform the behaviour

26= Prompt practice

29= Plan social support/ social change

30= Prompt identification as role model/ position advocate

5.5 TRAINING IN MALI

Six behaviour change techniques were used: Techniques 1, 2, 5, 21, 26, 30.

Provide information on consequences of behaviour in general (technique 1). The intervention provided information about the relationship between FGM/C and its possible consequences in the general case. It is explained that the intervention "consisted of the two following activities... Training sessions with social/health personnel ... recall of female anatomy and FGC ... and the different types of cutting. The health complications and their treatment was emphasized." (p15).

Provide information on consequences of behavior to the individual (technique 2). As evident from the following quote, the training of health staff included information tailored to the Mali client population "The training consisted of ... FGC, its context and local rational, its prevalence in Mali and elsewhere." (p15).

Goal setting (behaviour) (technique 5). Health staff was encouraged and trained to introduce "FGC-related activities within health talks at the clinics and during individual consultations with clients whenever feasible." (p7).

Provide instruction on how to perform the behaviour (technique 21). Health staff were told and instructed in how to communicate about FGM/C at the clinic and with clients. This was evident from for example "An entire day was devoted to the use of visual aids. Role playing was used to stimulate counseling sessions. All were trained in filling out health cards" (p15) and "All experimental sites were supplied with a flip chart to facilitate the IEC activities [IEC=information, education, communication]" (p9).

Prompt practice (technique 26). Health staff was prompted to rehears counselling sessions about FGM/C. It was stated that "communication technique used during the training session was practical role-playing." (p10).

Prompt identification as role model/ position advocate (technique 30). The training provided opportunities for health staff to persuade others of the importance of not performing FGM/C. It was stated that "They were expected to conduct anti-FGC talks with their clients." (p9).

5.6 EDUCATION IN EGYPT

One behaviour change technique was used: Technique 2.

Provide information on consequences of behavior to the individual (technique 2). The health education intervention provided information about FGM/C tailored to University of Alexandria students. This is evident through the explanation "The educational contents were based on the results of the pretest. For proper provision of information about RH [reproductive health] aspects and changing attitudes, the researcher used integration of communication skills with appropriate use of teaching methods. The educational methods included a health talk, group discussion, role play and use of educational aids such as posters about RH aspects with colored drawings to grab attention, colored leaflet about RH aspects. Handouts including all items mentioned in the session in the form of points were distributed to the participants in the program at the end of the session." (p439). And "Teaching methods and contents of the implemented RH program. 1. Health talks: importance of premarital counseling, dangers of FGM, antenatal care, family planning, breast feeding, sexually transmitted diseases." (p440).

5.7 COMMUNICATION PROGRAM IN NIGERIA

Techniques included 1, 2, 5, 6, 29:

Provide information on consequences of behavior in general (technique 1). The intervention, a multimedia communication programme, appears to have provided information about FGM/C and its possible consequences in the general case. It is stated that: "the program was designed to raise awareness of the possible negative effects of FGC ... address the cultural and socio-economic props that reinforce the practice" (p1595); "NAWOJ supported regular newspaper columns, radio call-in shows and public forums on FGC" (p1595).

Provide information on consequences of behaviour to the individual (technique 2). It seems the programme provided information about the benefits and costs of stopping FGM/C, tailored to the Nigerian groups involved based on their characteristics. Data in support include: "The programme ...aimed to contribute to the elimination of FGC in the project communities by challenging individual and communities to examine their beliefs and values around FGC and encouraging action towards the elimination of the practice" (p1595); "assisting community core groups to examine their knowledge, attitudes and practices surrounding women's reproductive health" (p1595).

Goal setting (behaviour) (technique 5). The groups were encouraged to decide to change. The reports states for example: "A key WARO input was the facilitation of community meetings where FGC and other maternal health issues were discussed and analysed. WARO also trained the core community groups and supported them in making action plans for the elimination of FGC in their communities" (p1595).

Goal setting (outcome) (technique 6). In addition to being encouraged to set behavioral goals regarding FGM/C, the programme encouraged the community members to set general goals regarding women's health. It was stated that the

programme “provide[ed] the groups with guidance and support as they developed action plans to improve women’s health situation.” (p1595).

Plan social support/ social change (technique 29). The communication programme appears to have prompted the participants to consider how to elicit social support from other people to help them achieve the target behavior of stopping FGM/C. It was stated that: “Programme methodology of facilitating group interactions among women and between men and women as well as promoting advocacy with peers is especially effective for improving both self-efficacy and perceived social support.” (p1600).

5.8 OUTREACH AND ADVOCACY IN KENYA

Four behavior change techniques were used: Techniques 2, 21, 29, 30.

Provide information on consequences of behavior to the individual (technique 2). The intervention’s focus was on “increasing knowledge of social and psychosexual effects” (p ix), it “systematically developed and implemented information and education activities using BCC* approaches” (p4). Information and education about the practice of FGM/C was tailored to the various community groups, as evident in the following quote: “specific training and messages were constructed for each group, e.g., religious leaders would speak to the erroneous association of Islam and FGC” (p8).

Provide instruction on how to perform the behavior (technique 21). Selected community members were trained in how to present FGM/C issues to others: “Training figured prominently in the early phases of the project to ensure a common understanding, approach, and educational messages related to FGC abandonment” (p8). This was also evident in a table describing key activities (p 43), where it is shown that training of traditional health providers took place as well as basic and refresher trainings for groups such as religious leaders, elders, traditional police, and youth peer educators.

Plan social support/social change (technique 29). It is stated that the intervention was “Supporting collective action/social changes: As community level changes began to be noted, the project helped support those who wanted to publicly commit to ending the practice” by setting up protection, working with community leaders, and helping create support groups (p9).

Prompt identification as role model/ position advocate (technique 30). The intervention had a clear advocacy focus; it was “built on education and advocacy interventions (p4). It was further stated that “Project staff systematically targeted influential people who could serve as advocates –such as religious leaders, teachers, elders, and the community leaders—with activities designed to provide information, raise issues and inform such influential leaders in order for them to advocate on issues of FGM” (p 8).

* Behavior change communication (BCC) is the strategic use of communication to promote positive health outcomes, based on proven theories and models of behavior change. BCC employs a systematic process beginning with formative research and behavior analysis, followed by communication planning, implementation, and monitoring and evaluation. Audiences are carefully segmented, messages and materials are pre-tested, and both mass media and interpersonal channels are used to achieve defined behavioral objectives (http://www.globalhealthcommunication.org/strategies/behavior_change_communication)

5.9 OUTREACH AND ADVOCACY IN ETHIOPIA

Four behavior change techniques were used: Techniques 2, 21, 29, 30.

Provide information on consequences of behavior to the individual (technique 2). According to the intervention descriptions, the intervention's focus was on "increasing knowledge of social and psychosexual effects" (p ix) and was comprised of "community-level educational outreach activities using Behaviour Communication Change (BCC)* approaches" (p4) as well as trainings which included "FGC consequences (health, psychosexual, social)) (p8). Information and education about the practice of FGM/C was tailored to the various community groups, as evident in the following quote: "specific training and messages were constructed for each group, e.g., religious leaders would speak to the erroneous association of Islam and FGC" (p8).

Provide instruction on how to perform the behavior (technique 21). Selected community members were trained in how to present FGM/C issues to others: "Training figured prominently in the early phases of the project to ensure a common understanding, approach, and educational messages related to FGC abandonment" (p8). This was also evident in a table describing key activities (p 43), where it is shown that training, including themes such as "health effects of FGC", took place. Women leaders, religious leaders, male leaders "were trained in advocacy."

Plan social support/social change (technique 29). It is stated that the intervention was "Supporting collective action/social changes: As community level changes began to be noted, the project helped support those who wanted to publicly commit to ending the practice" by setting up protection, working with community leaders, and helping create support groups (p9). The intervention included events such as "evening video sessions that showed recoded discussions by religious leaders speaking out on FGC issues" (p8).

Prompt identification as role model/ position advocate (technique 30). The intervention had a clear advocacy focus; it was comprised of educational outreach activities and "community-level advocacy" (p 4). It was further stated that "Project staff systematically targeted influential people who could serve as advocates –such as religious leaders, teachers, elders, and the community leaders—with activities

designed to provide information, raise issues and inform such influential leaders in order for them to advocate on issues of FGM" (p 8).

* Behavior change communication (BCC) is the strategic use of communication to promote positive health outcomes, based on proven theories and models of behavior change. BCC employs a systematic process beginning with formative research and behavior analysis, followed by communication planning, implementation, and monitoring and evaluation. Audiences are carefully segmented, messages and materials are pre-tested, and both mass media and interpersonal channels are used to achieve defined behavioral objectives (http://www.globalhealthcommunication.org/strategies/behavior_change_communication)

5.10 TOSTAN EDUCATION PROGRAM IN MALI

Four behavior change techniques were used: Techniques 2, 6, 26, 30.

Provide information on consequences of behavior to the individual (technique 2). The "Tostan Village Empowerment Program... is a trimmed-down version of Tostan's nonformal education and literacy program ... focuses on four modules that have the most relevance for the issue of FGC: human rights, problem solving skills, community hygiene and reproductive health" (p3). Information about FGC was "grounded in the local context" as the "different arguments for and against FGC were presented in the very words used by Malian women and men who had testified during previous village workshops" (p6). It was further stated that "The program actually implemented varied somewhat from the Tostan model, due both to conscious adaptation of the strategy to local circumstances" (p5).

Goal setting (outcome) (technique 6). The participants were encouraged to set goals as evident by intervention staff visiting "each village to follow up with committees and help them develop action plans" (p6).

Prompt practice (technique 26). The participants were prompted to rehearse the behaviour or preparatory behaviours numerous times. It was stated that "The program also gives participants the chance to practice active learning and problem-solving methodologies and to apply them to their own lives with the assistance of the facilitators assigned to the villages" (p 4).

Prompt identification as role model/ position advocate (technique 30). The intervention focused on how the community members could be an example to others and affect their behavior: "Participants are encouraged to take responsibility for sharing the information they have received and behavioral changes they have experiences with non-participants: other women and men from their own and neighboring villages" (p4). The report also included the following descriptions: "Committees were formed in each experimental village at the initiative of the participants, as part of the application of the concepts learned" (p 6) and

"...participants undertook a 'village exchange'. Five delegates from each of the participating communities met...to share their experience with actions taken... and to discuss how the committees that had been established might continue their activities now that the sessions were over" (p 6).

5.11 TOSTAN EDUCATION PROGRAM IN SENEGAL

Two behavior change techniques were used: Techniques 2, 30.

Provide information on consequences of behavior to the individual (technique 2). It was stated that "The overall objective of the evaluation was to provide information to support a strategy for improving women's health and for abandonment of the practice of FGC" (p2). The objective was sought through "a basic education program, developed by TOSTAN ... The basic education program consists of four modules: hygiene, problem solving, women's health, and human rights" (pi) and was implemented with contributions from the communities themselves. The educational program was tailored to the group also in that after first providing classes only to women – "TOSTAN recognized not including men created problems among them, who felt their exclusion was not fair" – the organizers invited also men to participate (p8).

Prompt identification as role model/ position advocate (technique 30). The intervention focused on how the community members could be an example to others and affect their behavior: "it was an expectation of the program that participants would share information on FGC with others in the community" (p33). It was also stated that "TOSTAN facilitators encouraged the leaders in each village, who wanted to bring about changes, to open discussions with the leaders of two other villages selected by themselves as 'parental linkage villages'."

5.12 TOSTAN EDUCATION PROGRAM IN BURKINA FASO

Two behavior change techniques were used: Techniques 1, 30.

Provide information on consequences of behavior in general (technique 1): Although it is quite possible that the information was tailored to the target population, as it was in one of the other Tostan programs, we were unable to identify such information. It seems information about FGM/C and its possible likely consequences in the general case, was provided, through "a broad-based functional educational program. TOSTAN's educational model includes four modules on human rights, problem solving, environmental hygiene and women's health" (pi). Further "the group sessions were held every two days, lasting 2 hours per session" and "the 63 group sessions were given to classes of men and women" (p3).

Prompt identification as role model/ position advocate (technique 30). The intervention focused on how the community members could be an example to others

and affect their behavior. The organizers “encouraged a system of formally expecting participants to share the information learned with other individuals in the community” (pii). Specifically, Tostan set up a mentoring system, also referred to as a sponsorship system, in which each participant “choose a friend in the village (this person is not a participant) with whom he / she share all information received in the classroom” (p17). Additionally, after having completed a study visit to the Tostan program in Senegal,”village meetings and inter-village meetings allowed members of the delegation to share the experience of Senegal” (p41).

Appendix 6: Mechanisms

Across the eight included effectiveness studies, a number of different behavior change theories were identified. These are listed in table 7. They were identified- and originate from the effectiveness reports. For transparency, we include relevant texts from these reports below, organized by country.

Table 7. Summary of mechanisms in the included effectiveness studies

Author	Population and intervention	Behaviour change theory
Diop 1998	Mali. Training, supervision	Training leads to improved knowledge, attitudes, skills
Mounir 2003	Egypt. Education	Education leads to improved knowledge, attitudes
Babalola 2006	Nigeria. Multimedia communication	Convention theory: Communication program will lead to increased awareness, which will lead to self-examination of beliefs, values, which will trigger ways of thinking, value orientations. Program will lead to dialogue, group/social interactions, advocacy, which will improve improves self-efficacy, perceived social support
Chege 2004a	Kenya (Somalis). Outreach, advocacy	Education leads to increased awareness; Training and education trigger advocacy; IEC activities affect intentions; IEC activities lead to individuals' improved knowledge, attitudes, which lead to groups' increased mutual understanding, agreement, which translates into collective action, which in turn shapes social norms
Chege 2004b	Ethiopia. Outreach, advocacy	Same as Chege 2004a (above)
Easton 2002	Mali. Tostan education program	Education leads to increased knowledge, understanding (critical awareness), which foster confidence, empowerment, which affect a sense of activism/motivation to act; Education affects intentions, attitudes, skills; Education leads to public discussions; Education increases empowerment, which affects attitudes, behaviors
Diop 2004	Senegal. Tostan education program	Education leads to increased knowledge, understanding; Education leads to improved attitudes, skills; Education leads to public discussions/social interactions (mobilization) which leads to public commitment; Education empowers people
Ouoba 2004	Burkina Faso. Tostan education program	Education leads to increased knowledge, understanding; Education leads to improved motivation, skills; Discussions and meetings leads to increased knowledge, engagement, and triggers confidence (in own abilities); Education empowers participants

5.13 TRAINING IN MALI

Information concerning mechanisms was collected from: Diop et al., 1998.

P14 The long term objective of this study is to contribute to the reduction of FGC in Mali through the use of clinical personnel trained in FGC issues. There were five short-term objectives: 1. Increase the level of knowledge of FGC issues and improve attitudes of health providers regarding FGC-related health problems. 2. Reduce the number of clinician willing to perform FGC and the demand of the parents. [3. Health facility clients] 4. Build the capacities of health personnel in the identification and treatment of FGC-related problems. [5. Health facility clients]. Hypothesis 1: Health care providers who have been trained in FGC issues will be better informed of health-related problems and more likely to oppose the practice than those who have not been trained. Hypothesis 2: Health care providers who have been trained are less likely to agree to cut girls than those who have not. [Hypothesis 3: Health facility clients] Hypothesis 4: Health care providers who have been trained will be more able to identify and to treat health problems related to FGC than those who have not.

P35 Although IEC tools were available in the experimental sites, few activities were conducted. Providers cited several reasons for this, including the lack of an appropriate space, the lack of time, and discomfort in publicly broaching a taboo subject. The time span provided for IEC training was apparently too short and should be expanded.

5.14 EDUCATION IN EGYPT

Information concerning mechanisms was collected from: Mounir et al., 2003.

P437 Aim of the work: To assess the impact of a short term health education program about reproductive health on knowledge and attitude of female Alexandra university students.

P439 Objectives of the program were to increase knowledge, remove misconceptions about RH [reproductive health] aspects, and to build a positive attitude toward RH aspects.

5.15 COMMUNICATION PROGRAM IN NIGERIA

Information concerning mechanisms was collected from: Babalola et al., 2006.

P1595 While there have been considerable efforts to stop FGC in Enugu State in the past, support for the practice persists in many parts of the state. The factors underpinning the practice in the state are cultural and support the convention

hypothesis ... For example, in the areas where the practice persists, women who have not undergone the procedure are generally believed to be unmarriageable, unclean and potentially promiscuous Indeed, the convention theory predicts that practices informed by convention will either persist indefinitely or stop abruptly when a significant number of people in the affected communities simultaneously decide to discontinue the practice. Against this background, the Health Communication Partnership, in collaboration with the National Association of Women Journalists (NAWOJ) and Women Action Research Organization (WARO), designed a multitiered and multimedia programme called Ndukaku (Igbo word for 'health is better than wealth'), with funding from the United States Agency for International Development (USAID). The programme focused on three local government areas (LGAs) in Enugu State (Uzo-Uwani, Isi-Uzo and Enugu South).The programme ...aimed to contribute to the elimination of FGC in the project communities by challenging individual and communities to examine their beliefs and values around FGC and encouraging action towards the elimination of the practice. Specifically, the program was designed to raise awareness of the possible negative effects of FGC, increase community dialogue about the practice, address the cultural and socio-economic props that reinforce the practice, and mobilize community members to abandon the practice and advocate in favor of its elimination among their peers.

P1597 We measured program exposure through recall of messages provided on any of the program materials or during any of the programme-related community events. More than three-fifths (63.4%) of the respondents in Enugu state were exposed to at least one programme material or participated in at least one program activity.

P1600 The use of multiple and credible communication channels in the FGC programme helped to make community-generated dialogue (involving local leaders, local men and women, and local radio personalities) accessible to the majority of the target audience. ... analyses consistently indicate that programme exposure is associated with significant improvement in FGC-related ideation, including increased perceived self-efficacy to resist the pressure to perform FGC, decreased personal approval of FGC, increased perceived social support for the abandonment of FGC and increased personal advocacy for the elimination of FGC. Programme methodology of facilitating group interactions among women and between men and women as well as promoting advocacy with peers is especially effective for improving both self-efficacy and perceived social support. There is also evidence from the data that exposure to the programme led to more widespread intention not to perform FGC. While exposure to the programme through either the mass media or the community activities alone had a significant impact on intention, the data document the marked added advantage of exposure through a combination of both sources.

P1601 the importance of exposure to consistent messages from a variety of sources for ideational and behavioral change. The ideational factors were jointly significant

in predicting intention, indicating that the ideation model is relevant to understanding FGC behaviors.

5.16 OUTREACH AND ADVOCACY IN KENYA

Information concerning mechanisms was collected from: Chege et al., 2004.

P4 The following general hypotheses were tested during the study: Community-level advocacy: trained civil society entities (organizational and individual advocates) will advocate for changes in community norms, helping to lead to the reduction of FGC. Community-level information and education: Systematically developed and implemented information and education activities using BCC approaches will lead to increased awareness of individuals, particularly women, as decision-makers who can influence decisions regarding their daughters' being cut, helping to lead to a reduction in FGC. In addition, country-level hypotheses were tested. Kenya: The comparison group population that is exposed to the IEC intervention alone will have 1. Lower knowledge of the negative effects of FGC, 2. Less awareness of human rights as they relate to FGC, 3. Be less supportive of FGC abandonment, 4. Will express less often their intention to not cut their daughters than the intervention group population exposed to IEC [information, education, communication] and advocacy activities. The project facilitated individual change and supported collective action as it emerged.

P7 The study design was contaminated in both Kenyan and Ethiopian sites, due to population movements into and out of the intervention and comparison areas. Refugees can move freely between the refugee camps in Dadaab, although not all people move between camps. Initially this was to be counteracted by having a third camp, situated between the intervention and comparison camps, to serve as a buffer. However, this changed when one of the camps refused to participate in the study, forcing the project to work in two adjacent camps. In Ethiopia, the movements of the Afar population with their animals during the rainy season may have also put people in intervention villages in contact with people in comparison villages. In addition, early in the project implementation stage a traditional system of communication between Afar clans, the Dagu system, was "discovered" by the project staff. Information on the project and consequent community discussions was relayed to different clansmen through the Dagu system, including people in the control areas (as the survey results demonstrate). / As the project progressed, a social change framework gradually emerged that guided the overall project activities and took into account the health, psychosexual, and social factors that influence both individual and larger community to choose to continue the practice, as well as efforts to support collective action as it emerged.

P10 Project staff had to take a "back seat" role in the process, guided by a belief that it was the communities' decision whether or not to change the practice, not the project's. The project's relationship with local leaders also shifted – from

coordinating to working with leaders to become advocates on FGC abandonment issues. Even with this thoughtful planning, staff faced challenges as the information and advocacy interventions began. In a few instances, especially in Kenya, staff were actually threatened by community members because they publicly dared to bring up the subject and provide more complete information on negative health and social consequences relating to FGC. Once community reactions started to occur and individuals began to publicly come out against the continuation of the practice, the issue of protecting those being marginalized and threatened arose, and with that, the question of who had responsibility to protect these people. / Several program design factors were essential to the success of the FGC abandonment activities. First, the program shifts described above could only occur in RH projects that were community-based. Facility-based programs by definition are limited in their scope of activities and ways of interacting with communities. A second critical factor was that the topic of FGC abandonment was part of a larger set of issues addressed by the RH projects. If FGC had been tackled as a stand-alone issue, and not within a larger project context such as health and social well-being, it is likely that the community would have viewed the interventions as part of an FGC abandonment agenda being imposed by outsiders. Related to this, the project position of not imposing its values on the community was likely critical. Recognition of CARE from earlier projects in the area as a "neutral" actor, with no religious or other agenda, allowed for dialogue.

P11 (figure 1.1) Psychological realities, physical realities, social realities. Project inputs: Information: health, socio-cultural-religious, rights to good health and of women and girls. Fostering dialog: creating "spaces" for sharing and discussion. Fostering advocacy: religious and other community leaders. Supporting change: support groups, protection from stigma by elders. → New understanding. New ways of believing → Individuals change. / More mutual understanding. More mutual agreement → Collective action by groups. → Social change. / Eventual change in social norms.

P19 In Kenya, radio was ranked as the second highest source of anti-FGC messages, especially in the comparison site. It is important to note that although advocacy through religious leaders was one of the key strategies used in the intervention sites in Ethiopia and Kenya (though in Kenya the work with religious leaders was never fully realized), an almost equal proportion of respondents from the intervention and control/comparison sites (23% versus 30% in Kenya), cited religious leaders as their source for anti-FGC messages, which raises questions about the possibility of contamination of the intervention strategies in the control/comparison sites and the weakness of the intervention.

P24 Prevailing attitudes and beliefs around the practice of FGC help to drive the continuation of the practice; therefore, they were addressed in FGC abandonment interventions as points for discussion and debate.

P33 The entry was aided by the fact that reproductive health activities were already operating prior to this project and communities already trusted and respected the

CARE staff. The project staff and volunteers' roles in the intervention as trusted facilitators of a community process: For example, several religious leaders –already reached by CARE/NCCK outreach prior to this project - continued to advocate for an end to the practice in the comparison camp. The resulting effect was exacerbated when planned work with local religious leaders to develop a consensus on the position of Islam on FGC and to ask them to work as advocates never occurred, and so religious leaders continued to provide mixed messages to the faithful in both camps.

P36 The significant increases in knowledge and attitude in support of FGC as a rights issue in Kenya did not translate into increases in support of FGC abandonment and the intention not to cut their daughters in the future. One plausible reason could be the intensity of exposure to the intervention. The proportion of respondents reporting that they were exposed to anti-FGC messages in the Ethiopian intervention site was much higher than in the Kenyan intervention and comparison sites. Both in Ethiopia and in Kenya, it is important to note that multiple logistic regression models underline exposure to anti-FGC messages as a significant predictive factor for supporting FGC abandonment, believing that FGC violates the rights of women and of girls, and intention not to cut their daughters in the future. Therefore, lower levels of exposure to the intervention might translate to lower increases in positive attitudes and intended behaviour. It is also possible that there were other contextual factors in the Kenya refugee camps that were not captured in the quantitative surveys. For example, qualitative data collected during the endline alluded to the possibility of a backlash against the project activities due to the enforcement of the law criminalizing FGC in Kenya during the intervention period. Attempts to enforce this law among some Somali refugee parents, who were discovered to have cut their daughters, enraged the community.

P37 Changing practices that are deeply entrenched in the culture and social life of individuals and their communities, such as FGC, requires individuals to identify and acknowledge these practices as problematic, consider their importance, evaluate their own behaviour, attitudes and beliefs, and begin to make changes in their lives. However, since individual behaviour is strongly linked to and reinforced by prevailing social norms and belief systems in their community, getting the larger community and those influential people who are charged with upholding social norms to contemplate and question these norms and belief systems is critical to create an enabling environment for individual behaviour change.

P39 Although both groups are of Islamic faith and highly religious, CARE was much more successful in engaging the religious leaders in Ethiopia than in Kenya. ... It is also plausible that the fact that the FGC intervention in Kenya was associated with a Christian group (NCCK) might have raised distrust among the predominantly Islamic group.

5.17 OUTREACH AND ADVOCACY IN ETHIOPIA

Information concerning mechanisms was collected from: Chege et al., 2004.

Pix To test the effectiveness of education activities using behavior change communication (BCC) approaches and advocacy activities by religious and other key leaders...

P4 ... in the intervention site hypotheses were tested. Ethiopia: The control group population that is *not* exposed to IEC [information, education, communication] and advocacy interventions will have. 1. Lower knowledge of negative effects of FGC, 2. Less awareness of human rights as they relate to FGC, 3. Be less supportive of FGC abandonment, 4. Will express less often their intention to not cut their daughters, than the intervention group population that is exposed to IEC and advocacy activities.

P7 As the project progressed, a social change framework gradually emerged that guided the overall project activities and took into account the health, psychosexual , and social factors that influence both individual and larger community to choose to continue the practice.

P10 Even with this thoughtful planning, staff faced challenges as the information and advocacy interventions began. In a few instances, especially in Kenya, staff were actually threatened by community members because they publicly dared to bring up the subject and provide more complete information on negative health and social consequences relating to FGC. Once community reactions started to occur and individuals began to publicly come out against the continuation of the practice, the issue of protecting those being marginalized and threatened arose... Several program design factors were essential to the success of the FGC abandonment activities. First, the program shifts described above could only occur in RH projects that were community-based. Facility-based programs by definition are limited in their scope of activities and ways of interacting with communities. A second critical factor was that the topic of FGC abandonment was part of a larger set of issues addressed by the RH projects. If FGC had been tackled as a stand-alone issue, and not within a larger project context such as health and social well-being, it is likely that the community would have viewed the interventions as part of an FGC abandonment agenda being imposed by outsiders. Related to this, the project position of not imposing its values on the community was likely critical. Recognition of CARE from earlier projects in the area as a “neutral” actor, with no religious or other agenda, allowed for dialogue.

P11 (figure 1.1) Psychological realities, physical realities, social realities. Project inputs: Information: health, socio-cultural-religious, rights to good health and of women and girls. Fostering dialog: creating “spaces” for sharing and discussion. Fostering advocacy: religious and other community leaders. Supporting change: support groups, protection from stigma by elders. → New understanding. New ways of believing → Individuals change. / More mutual understanding. More mutual

agreement → Collective action by groups. → Social change. / Eventual change in social norms.

P19 During the endline survey, respondents who reported that they had heard anti-FGC messages in the past two years were asked to name the sources of their information and the places where messages were heard. CARE and partner agency staff were mentioned by the highest proportion of respondents in both Ethiopia and Kenya. In Ethiopia, religious leaders were the second highest source of anti-FGC messages.

P24 Prevailing attitudes and beliefs around the practice of FGC help to drive the continuation of the practice; therefore, they were addressed in FGC abandonment interventions as points for discussion and debate.

P33 The project staff and volunteers' roles in the intervention as trusted facilitators is perceived as critical in allowing the project to go as far as it did in terms of fostering change on the issue.

P37 Changing practices that are deeply entrenched in the culture and social life of individuals and their communities, such as FGC, requires individuals to identify and acknowledge these practices as problematic, consider their importance, evaluate their own behaviour, attitudes and beliefs, and begin to make changes in their lives. However, since individual behaviour is strongly linked to and reinforced by prevailing social norms and belief systems in their community, getting the larger community and those influential people who are charged with upholding social norms to contemplate and question these norms and belief systems is critical to create an enabling environment for individual behaviour change.

P39 CARE was much more successful in engaging the religious leaders in Ethiopia than in Kenya. Many of these influential leaders in Ethiopia started to speak with one voice on the position of Islam vis-à-vis FGC, themselves becoming a critical mass that could advocate for change and in so doing, give permission to the faithful to discuss and change a practice that no longer was considered relevant for reasons of religion.

5.18 TOSTAN EDUCATION PROGRAM IN MALI

Information concerning mechanisms was collected from: Easton et al., 2002; Gillespie & Melching, 2010; Monkman et al., 2007.

Easton et al., 2002

P4 The main objectives of the program address change at both the individual and the collective level. *At the individual level*, the program is designed to transform the perspectives and activities of individual participants, both men and women, by two principal means. (1) The village empowerment curriculum developed by Tostan provides information on women's rights and on critical health issues in a human rights framework. It was to be presented in evening sessions, over a 6-month period,

to participants from the five experimental villages. (2) The program also gives participants the chance to practice active learning and problem solving methodologies and to apply them to their own lives (3) Participants are encouraged to take responsibility for sharing the information they have received and behavioral changes they have experienced with non-participants: other women and men from their own and neighboring villages. (4) They should also be given the tools, desire and ability to initiate and to sustain collective action focused on securing respect of human rights and improving living conditions in their communities. *Program strategy* Although these stipulations are not explicit laid out in the curriculum, it is implicitly expected that the process of empowerment set in motion by the program will involve (a) participants making behavioral and attitudinal changes in their own lives; (b) non-participants who hear the information, and/or observe the changed behaviors and practices of participants, making related changes in their lives; and (c) participants and non-participants coming together to share information and critical analyses and to initiate collective actions that will improve the living conditions in the community as a whole. While the classes are still in session, the group "dynamic" nurtured in class, the regular meetings and the problem-solving habits practiced should lay the foundation for ongoing and "empowered" collective action that will benefit the community as a whole.

P6 The different arguments for and against FGC were presented in the very words used by Malian women and men who had testified during previous village workshops run by IEP. This "grounding" in the local context seems to have been both very effective and highly consistent with Tostan's approach to village empowerment. In Senegal, the original curriculum was devised in an intensively participatory and iterative manner, incorporating language and cultural forms familiar to participants.

P7 Staff of Tostan and of IEP ran into some major disagreements during the course of implementation, due in large part to the modifications in the originally proposed methodology that the Malian personnel felt it necessary to adopt.... This touchy negotiation took place at the same time as a significant breakdown in communication between Tostan and IEP regarding activities during the last three months of the program.

P16 many other things were taking place in the environment at the same time, so that it is altogether possible that changes might come about independently of the program.

P28 In at least two of the villages, Sebeninkoro and Konobougou, there was already in place a program of awareness-raising about FGC and its negative consequences, carried out by a Malian NGO.

P37 IEP did not turn anyone away from the classes; at first, a majority of males signed up, and in some villages, only men. Many quit after they saw that the sessions included mostly topics related to women; some quit when they realized that this wasn't a literacy program. IEP pursued different strategies to ensure significant and

meaningful participation by women. For example, they carried out a prise de position (public meeting devoted to "taking a stance") on the subject of women's participation. This had the effect of selecting out those men who were not comfortable with women's participation; in a sense it forced men to break with the traditional position. Some men quit at that point; the remaining male participants were a self-selected group and were valuable allies.

P37-38 The final evaluation suggests that the participation of men – or at least of these self-selected male allies – was critical to the success of the program. They were active participants in encouraging women's participation and in mobilizing for the applications of session concepts; and they are now strongly represented on all the committees that grew out of the program. ... In general, as the questionnaire data reviewed in Section II reveal, men tended at the outset to be somewhat more informed and "pro-active" on issues of health than were women (excepting the topic of human rights, where both groups were ill-informed and at a virtual par), a gap that closed among participants during the program. They could, therefore, if they chose, serve as very useful catalysts in the courses. .. It is not surprising that men also represent the majority on all but one of the committees interested in working to stop the practice of FGC (Konobougou constituted the one exception), given that it is women who have traditionally been the implementers and strong supporters of the practice. What is encouraging, however, is that on all the FGC committees except one (Noumabougou), there is at least one woman.

P39 Some facilitators had more experience with village-level facilitation than others ... The facilitators in Sebeninkoro, Bananbanin and Doneguebougou had been school teachers, two of them with IEP; all three were dedicated to their work but less comfortable with the modules, less effective at engaging participants and increasing their participation. The observations of the final evaluation team also suggest that more training up front could have made a significant difference.

P43 Overall it seems clear that the Tostan/IEP program was quite effective – despite significant problems of implementation – in conveying new information about human rights and women's rights to participants, in connecting these concerns with those of personal and community health, in raising the problems of FGC (or helping participants themselves raise these issues) in a *human rights'* context, and in promoting increased activism among the group with respect to individual and community health and the abandonment of FGC.

P43-44 It seems the approach advocated by both Tostan and IEP is what is needed at this time in Mali. It opens up communication between people, gets them sharing experiences and information in an environment of attention to human rights, and encourages collective action. It also includes significant numbers of both men and women. This alliance is particularly useful in the movement to stop FGC given that women are the main implementers and supporters of the practice.

P44 The human rights framework around which the Tostan curriculum is organized is compelling. It was clear from our interviews that the concepts were meaningful to

participants and offered them keys to solving some of their problems. The absence in the Tostan curriculum of any discussion of *human rights for men and children*, however, needs to be addressed.

P45 The findings of the final evaluation suggest that the mix of men and women in the sessions was advantageous; men helped deepen the analysis in the sessions, and helped with problem solving. In the early stages they helped encourage the participation of women. To make sure that significant participation by men does not reduce the opportunity for women to become empowered, the program should consider regularly scheduled in-depth sessions with men and women separately to give women more opportunity to hone their leadership skills and men a chance to focus on their own gender issues. Already as part of the program, separate women's circles were created informally in the villages in the mornings by the facilitators as they circulated among the women. It is clear that these were important reinforcement mechanisms for women. ... Improve program scheduling and pre-service facilitator training. There is no doubt that it would have been better – under ideal circumstances – to implement Module 4 of the Tostan curriculum (the one on reproductive and sexual health and FGC) over the longer period initially designed and intended. This would have given time for the facilitators to support the ongoing work of the committees as they engaged in "harder" activities.

P46 It should also include by design a significant period of facilitator training before the program starts, including instruction in the content of the modules themselves and hands-on practice in different methods of active learning and different ways of promoting adult participation. Such a session should also include more instruction about keeping accurate monitoring records and assessing participants' progress in acquiring key concepts and translating them into action. This kind of in-depth training held "up-front" would remove the need for regular meetings in Kati during implementation and would strengthen the skills of inexperienced facilitators in particular. .. Clarify mutual expectations at the outset. The problems between IEP and the facilitators could have been avoided had expectations been clear at the beginning (and not changed midstream) and had there been some form of official *contract* between IEP and the facilitators.

P47 In a similar vein, the major misunderstanding between Tostan and IEP could have been avoided had there been more clarity about what each party expected in the initial agreement... The goals of the collaboration as stated in written documents are actually in conflict: Nowhere is it clear the extent to which the program could be adjusted to fit the local context -- yet adjustments were inevitably necessary since the Malian context is quite different from Senegal.

Gillespie & Melching, 2010

P478 nonformal adult education program in Sub-Saharan Africa empowered learners and positively transformed their communities.

P491 educational practices – use of experiential, interactive pedagogy and integration of socio-cultural context into the learning process. / Engaging

participants in imagining and discussing their future aspirations at the beginning of the program allowed village community members to express their own understanding of their 'indigenous conceptions' of human rights.

P492 The discussions revealed gaps in knowledge or generative themes. ... Collectively identifying the gap between what is and what could be motivated participants to act, and everyone contributed to the solution, equalizing participation.

P493 The availability of a larger discourse community, however, emboldened community members to share their new understandings with friends, family members, and neighbors.

P494 the modules on democracy and human rights led to a more vigorous praxis: to refection and action that has created large-scale social transformation. / Tostan's introduction of visioning created space for imaginative explorations of possibilities for their communities. / When communities learned about international human rights groups and their work, they saw them as resources (and inspiration) for their own struggles / Tostan did not confine learning to the traditional classroom. ... created new spaces (the CMC [Community Management Committees], adopt-a-learner, and intervillage meetings) so that participants could practice in ever-expanding public venues what was learned and practiced in the classroom.

Monkman et al., 2007

P451 Non-formal education 'village empowerment program' designed to empower individuals and communities through knowledge acquisition and critical awareness leading to community decision-making and action, particularly around issues related to women's health and human rights.

P453 The VEP is designed to create experiences of power for both individual and communities. / The VEP approach to empowerment is a participatory model of education and social change that enables people to identify and address their own problems individually and collectively through a systematic process of critically examining life conditions, constructing awareness of contextual and influential factors, and planning and implementing change.

P457 The combination of appropriate facilitation techniques and solid emancipator knowledge cast in a human rights framework helped build confidence so people could speak out and mobilize to obtain their rights.

P458 collective initiatives to bring about the abandonment of FGC began to take shape only after the 3-day intensive workshop on reproductive health and FGC. Interviews with the facilitators and the supervisors suggest participants were initially reluctant to discuss the practice; they did not even speak about FGC among themselves before the program, and were even less inclined to do so publicly. / Yiriba and Kofane already had a program of awareness-raising about FGC carried out by another Malian NGO. The midwife who worked in the villages for another

NGO believed their approach had given people new information and gotten them talking to one another and to her and her colleagues about FGC.

P459 Exposure to new information is a critical factor in motivating change in attitudes and behaviors related to reproductive health and human rights

P460 sessions to build awareness, to generate discussion, to gain experience in problem-solving processes. As the participants' knowledge and skill increased, mediation gave way to more locally-driven socio-political forms of empowerment.

5.19 TOSTAN EDUCATION PROGRAM IN SENEGAL

Information concerning mechanisms was collected from: Diop et al., 2004; Diop et al., 2008a; Diop & Askew, 2009; Gillespie & Melching, 2010; UNICEF, 2008.

Diop et al., 2004

Pi The basic education program consists of four modules: hygiene, problem solving, women's health, and human rights. Through these four themes, emphasis was placed on enabling the participants, who were mostly women, to analyze their own situation more effectively and thus find the best solutions for themselves.

Pii Those from the Mandingo ethnic group seemed more willing to abandon the practice than those from the Pulaar ethnic group, which may be due to the influence of the Pulaar religious leader who was not in favor of abandonment.

Pv The education program consists of four modules: hygiene, problem solving, women's health and human rights. Through these four themes, emphasis was placed on enabling the participants, who were mostly women, to analyze their own situation more effectively and thus find the best solutions for themselves. TOSTAN's approach was based on peaceful social change through a basic community education program and a process of social mobilization.

P2 Informing the traditional and religious leaders and the elected politicians of the area about project activities;/Selection of facilitators and supervisors from the communities that participate in the program.

P3 The overall objective of the evaluation was to provide information to support a strategy for improving women's health and for abandonment of the practice of FGC.

P3 Hypotheses. 1. Women who have attended a basic education program will have higher levels of knowledge about, and more positive attitudes towards, reproductive health (including FGC) and human rights than women who have not attended a basic education program. 2. Partners of women who have attended a basic education program will have higher levels of knowledge about the negative consequences of female genital cutting and be more in favor of abandoning the practice. 3. Women and men in communities in which women have attended a basic education program will have higher levels of knowledge about the negative consequences of FGC and

will be more likely to make a public declaration abandoning female genital cutting than women and men in communities where no such program has been undertaken. 4. Women and men in communities in which women have undertaken a basic education program and in which a community mobilization process has been completed will be more likely to abandon the practice.

P8 Participation in the program. Before the study, TOSTAN provided a listing of the women and men in the 20 villages who had expressed interest in participating in the program. By the time the program started, however, only 64 percent of the women and 50 percent of men who had expressed interest actually did so. People didn't attend the program for several reasons: men failed to participate because they found that no economic compensation was provided and women because classes were too overcrowded, they got sick or pregnant, or because their husband forbade it. The lower level of participation by the men was the result of a decision by TOSTAN, as the education program was developed originally to empower women. However, TOSTAN recognized not including men created problems among them, who felt their exclusion was not fair. With the exception of one village (Karcia), all villages experienced men dropping out of the program, including one village where all nine men who said they would participate in the program dropped out. According to an informant in this village, this was because: "the facilitator did not choose the best way of introducing the theme about the practice of FGC. The village head didn't like this theme, and he called for the center to be closed. The idea of destroying the hut and attacking the supervisor came up on the day a baptism was being celebrated in the village. The people thought the program was coming to fight against the traditional culture, which had come down from their forefathers, and even more to fight against the principles of Islam and the purification of the woman." According to the facilitator "several meetings had been held to discuss carrying out the TOSTAN program in the village, up to the day when the son of the president of the women's group turned up and demanded that the hut should be torn down at once and the facilitator expelled." Participation and Interest in the Program's Modules. Overall, 69 percent of the women and 57 percent of the men who participated in the program completed all four modules, with the remainder completing part of the program.

P25 at the time of the baseline over half of the men and women in both the intervention and comparison groups reported having already received some information about FGC.

P26 When asked the source of their information, 87 percent of women participants identified the TOSTAN program, and one quarter of non-participating women did so. However, only 36 percent of participating men and 19 percent of non-participating men obtained their information from TOSTAN; the main source cited was the radio. ...This indicates that the program achieved its intention of encouraging communities to openly discuss FGC following the TOSTAN educational sessions.

P32 This successful experience encouraged TOSTAN to consider the educational program, the community-wide discussions that follow, and the public declaration to be a model for enabling social change. Those who took part in the educational program organize a process of social mobilization. Public discussions are held with the community in general on subjects of interest identified by those who participated in the program so that the whole village is exposed to the intervention. The discussion leaders seek the support of the community for denouncing certain harmful practices (especially FGC and early marriage). The direct exposure of women and men to the program intervention dramatically increased their knowledge, improved their attitudes and changed their behavior on most indicators.

P33 Of critical importance was the expectation that women would spread the information learned within the villages. They shared the information they had obtained with their friends, their children and their husbands. This was less effective among the men who participated, as they were less inclined to share the information they acquired.

P34 It was much more difficult, however, to organize for subjects concerned with harmful practices, such as FGC, as evidenced by less discussion and less willingness to organize during the implementation process. Overall, the campaign against FGC made less of an impression on the villages. Some of the village committees set up to support the social action found problems of planning and organization.

P36 The TOSTAN educational program was successfully implemented in all 20 experimental villages; on average, most women who participated in the program attended all four modules. Women who didn't complete all the sessions attended mainly the human rights and the hygiene modules. Only half of the men participating followed the entire program, but most attended the hygiene module. The problem-solving module was the least well attended by both women and men, followed by the health module. For nine of 10 women, their expectations were fully satisfied, and communities also reported high levels of satisfaction during focus group discussions.

Diop et al., 2008a

P8 The communities themselves generally contributed a great deal to programme introduction and implementation. This is especially evident in the creation of the workspace known as the classroom...

P9 Villagers also worked collectively to implement the program*: "Everyone worked as one and put themselves out to help Tostan so that there would be no difficulty with the courses and no problems among the participants."* However, this widespread enthusiasm for accepting and setting up the programme did not carry over to class attendance. The perception was that the programm's usefulness had to be weighed against the obligations of day-to-day survival, a case some made to justify irregular attendance. For many women, domestic duties were paramount and justified missing classes.

P26 Villages with strong leaders and the ability to support a facilitator are both factors that affected village selection. After the first public declaration, the name of Tostan became synonymous with abandoning FGM/C, as the event was widely covered by the media. A transition was also seen as villages began requesting the programme through the influence of inter-village migrations and the programme's popularity. For this reason, the possibility of selection bias should be considered. If an agreement to abandon FGM/C appears to be a condition for introducing the programme into a village, the most resistant villages will naturally exclude themselves from the selection process. The most destitute villages may also be excluded because they are unable to support a facilitator.

P27 Inconsistent attendance and dropouts over time may also result in bias. Even if the reasons were justified, only the most motivated women participated in the complete programme. Does this mean that only the most determined women, who probably have the most progressive ideas, benefit fully from the program? In other words, do only the most convinced, those already predisposed to change, end up building their capacities? A previous evaluation has already shown a higher level of knowledge among these women before the programme began (Diop et al. 2004). The less independent and perhaps busiest women with household chores and children (and therefore usually the youngest women) end up being less involved. However, one interesting result of the study is that any differences between participants and non participants tends to be erased over time, with other women living in the village eventually reporting that they have adopted the movement toward change.

Diop & Askew, 2009

P308 By first learning about human rights, women and men were expected to better understand and defend their rights related to health, and particularly the rights to bodily integrity for themselves and their children, especially their daughters, and to bring about positive transformations in their own mental and physical well-being. This education served as an empowering foundation for the remaining modules and was constantly referred to throughout the program. (2) The problem-solving process: Participants learned a generic process for problem solving to help them better identify and analyze their own human rights, health, and hygiene issues so as to find solutions appropriate to their situations and implement them at the individual, family, and community levels. (3) Basic hygiene: Information about the transmission of germs and the importance of hygiene was imparted to deepen participants' knowledge of health issues, in general, and of the harmful effects of FGM/C, in particular. (4) Women's health: Women and men learned about women's health and how to improve women's physical and mental well-being, including family planning and other reproductive health issues, within the context of a woman's right to health. Positive traditions that lead to women's improved health were stressed, and negative traditions such as FGM/C were discouraged.

P309 The rationale for the TOSTAN strategy as a model for changing behavior concerning FGM/C has been explained by reference to social convention theory.

P313 This last point highlights two 'contamination' possibilities. First, TOSTAN had supported a radio broadcast throughout the region; thus, radio messages with FGM/C information may have influenced those living in comparison villages. Two, this change in attitude may represent the outcome of social interactions and discussions among residents of the comparison and intervention villages...

P315 the outcomes achieved could have been greater had more people attended the classes as planned and had the participants attended all four modules of instruction. ... The potential exists for the study's findings to reflect selectivity bias. Participants in the intervention differed to some degree from those living in the comparison communities in terms of ethnicity, marital status, and age, and at baseline their attitudes toward FGM/C and its continuation were less favorable than the attitudes of those living in the comparison communities, although the two groups reported at baseline roughly the same level of cutting among daughters aged ten and younger. Because participants may have been more likely to be predisposed to change their attitudes and behavior than nonparticipants and than those in the comparison-group sample...

Gillespie & Melching, 2010

P478 nonformal adult education program in Sub-Saharan Africa empowered learners and positively transformed their communities.

P491 Engaging participants in imagining and discussing their future aspirations at the beginning of the program allowed village community members to express their own understanding of their 'indigenous conceptions' of human rights.

P492 The discussions revealed gaps in knowledge or generative themes. ... Collectively identifying the gap between what is and what could be motivated participants to act, and everyone contributed to the solution, equalizing participation.

P493 The availability of a larger discourse community, however, emboldened community members to share their new understandings with friends, family members, and neighbors.

P494 the modules on democracy and human rights led to a more vigorous praxis: to reflection and action that has created large-scale social transformation. / Tostan's introduction of visioning created space for imaginative explorations of possibilities for their communities. / When communities learned about international human rights groups and their work, they saw them as resources (and inspiration) for their own struggles / Tostan did not confine learning to the traditional classroom. ... created new spaces (the CMC [Community Management Committees], adopt-a-learner, and intervillage meetings) so that participants could practice in ever-expanding public venues what was learned and practiced in the classroom.

UNICEF, 2008

P2 programme of informal basic education which aims to provide populations, especially in rural areas, with skills and knowledge (factual and behavioral) that will

enable them to be the actors in their own development. ... skills acquired by the beneficiaries of both sexes at the module level – especially in the "human rights and responsibilities" module – have enabled populations to develop greater awareness of their multiple problems, mostly in the health area.

P16 local people themselves requested the programme after they received good feedback about it, or following the migration of a person who knew the programme.

P28 the programme was not well implemented in all villages, nor did it consistently reach its objectives. Two of the villages visited met with problems which limited the learning of the residents there.

P69 First, we can question the choice of villages. All the information gathered by Tostan as well as by the villagers shows that some pre-conditions are necessary for a village to be included in the survey. Some villages have strong leaders who polarize other villages; a village's obligation to support facilitators is another factor that can influence its selection. After the first public declaration, the name of the NGO was strongly associated with the abandonment of circumcision in the media coverage that this programme enjoyed. We also saw that a transition occurred as villages started asking for the programme due to its popularity and because of the effect of inter-village migrations. We therefore wonder about choice biases. Since abandoning circumcision is a theme that appears to be a prerequisite for establishing a programme, the more intractable villages can exclude themselves from the process. The more deprived villages can also exclude themselves from the process because they cannot afford the facilitator. ... One can also wonder about participants in the programme. The low attendance level of some participants and drop-outs over time suggest a possible bias. Even if the reasons for dropping out appear justified, the fact is that only the most motivated women thoroughly participate in the programme.

5.20 TOSTAN EDUCATION PROGRAM IN BURKINA FASO

Information concerning mechanisms was collected from: Ouoba et al., 2004; Gillespie & Melching, 2010; Diop et al., 2003.

Ouoba et al., 2004

Pi Although one of the ultimate goals of this model is the eradication of female genital cutting (FGC), Tostan has found that an effective means to achieve this is by empowering community members to understand and act on their situation through a broad-based functional educational program. TOSTAN's educational model includes four modules on human rights, problem solving, environmental hygiene and women's health. TOSTAN's hypothesis is that communication of technical information, discussion of human rights issues, and development of strategies for social transformation improve the confidence and self-determination of women through participatory educational methods.

Pii Mwangaza encouraged a system of formally expecting participants to share the information learned with other individuals in the community as this had been effective in Senegal, but it did not work well in this context, in that it limited the spread of information throughout the whole of the community.

P1 the strategy focuses on three areas: basic education, public debates and public statements. In addition, this program supports strategies from the PLA (participatory training approach) which indicates that the abandonment of the practice of female circumcision through education of the population and the involvement of the entire community organizing activities.

P3 the education program initiated the principle of sponsorship is for each participant to choose a godchild (or godchildren) of the village whom he shares his knowledge acquired.

P17 In implementing the program, a mentoring system has been developed to transmit to others the information received. Each participant is to choose a friend in the village (this person is not participants) with whom he / she will share all information received in the classroom.

P45 The various activities in the experimental area during the implementation of the program allowed the inhabitants of these villages to receive information on FGM.

Gillespie & Melching, 2010

P478 nonformal adult education program in Sub-Saharan Africa empowered learners and positively transformed their communities.

P491 Engaging participants in imagining and discussing their future aspirations at the beginning of the program allowed village community members to express their own understanding of their 'indigenous conceptions' of human rights.

P492 The discussions revealed gaps in knowledge or generative themes. ... Collectively identifying the gap between what is and what could be motivated participants to act, and everyone contributed to the solution, equalizing participation.

P493 The availability of a larger discourse community, however, emboldened community members to share their new understandings with friends, family members, and neighbors.

P494 the modules on democracy and human rights led to a more vigorous praxis: to refection and action that has created large-scale social transformation. / Tostan's introduction of visioning created space for imaginative explorations of possibilities for their communities. / When communities learned about international human rights groups and their work, they saw them as resources (and inspiration) for their own struggles / Tostan did not confine learning to the traditional classroom. ... created new spaces (the CMC [Community Management Committees], adopt-a-learner, and intervillage meetings) so that participants could practice in ever-expanding public venues what was learned and practiced in the classroom.

Diop et al., 2003

P1 The specific objectives were: a) To empower participants through education in human rights, the problem solving process, hygiene and women's health. b) To actively engage the populations of these zones in reproductive health issues by organizing meetings and public discussions. To achieve these objectives and attain the expected programme outcome, both men and women would acquire:

- An understanding of human rights and how they can be applied to improve living conditions
- Skills for solving family and community problems
- Technical information relating to hygiene and disease prevention
- Knowledge about the development and functions of the human body
- Information on the methods of promoting health at each stage of the development of the human body
- Knowledge of traditional practices that are harmful to health and the strategies to end them
- Communication skills to effectively share information with other members of the community and other villages in the zone
- Motivation to undertake research in order to find common solutions to problems facing the village
- Renewed confidence in their ability to initiate positive social change.

P2 new understanding gained in this module serves as an empowering foundation for the remaining themes and is constantly referred to throughout the programme. 2. The Problem-Solving Process: Participants learn a generic process for problem solving that helps them better identify and analyze their own problems related to human rights and responsibilities, health, the environment and project implementation. This process helps them find solutions appropriate to their situation and to implement them on an individual, family and community level. 3. Basic Hygiene: Participants understand the importance of personal and community hygiene by learning about germ transmission and the benefits of prevention over the costs of treatment. Participants immediately apply their new problem-solving skills for the improvement of community hygiene problems. 4. Women's Health: Through open and honest dialogue, participants gain information to improve women's physical and mental health in the community. Focusing on the importance of respecting women's human rights, participants talk about their bodies and reproduction, healthy development and illnesses affecting women and girls. Participants also weigh the merits of certain positive traditional practices (breastfeeding and infant massage) against the dangers of others such as early marriage and FGC.

P19 The second part of the programme had to be postponed until the following year when participants are available to study. The loss of these experienced and trained

facilitators disrupted normal programme progress. Class attendance problems: Some social events (baptisms, marriages, funerals), which occurred when the classes were in session, disrupted the programme because of numerous absences and sessions had to be repeated, causing inconvenience to facilitators in the preparation and teaching of the sessions. Lack of tangible incentives such as small income generating and micro credit projects to motivate programme participants was frequently mentioned as a drawback. This led to a lack of interest in certain topics and irregular attendance.

Appendix 7: Results from context studies

Appendix 7 presents results of factors found to perpetuate and hinder FGM/C in Egypt and in Nigeria.

5.21 EGYPT

We included five studies from Egypt which presented results of multivariate regression models examining predictors perpetuating and hindering FGM/C (table 8).

Table 8. Results of context studies from Egypt, predictors perpetuating and hindering FGC

Author	Results
Allam 1999	Predictors of supporting/condoning FGM Male gender OR=1.52 Do not know a case who suffered complications following FGM OR=2.13 Non-medical literature as source of information about FGM OR=1.85 Disagree that FGM has no religious basis OR=2.53 Disagree that FGM is not equivalent to virtue OR=2.04 Disagree that FGM is a custom OR=2.59 Disagree that FGM is a harmful procedure OR=4.11 Disagree that FGM may negatively affect future sexual relations OR=1.55 Disagree that FGM is usually followed by complications OR=5.14
Dandash 2001b	Predictors of positive attitude to female circumcision Rural residence OR=10.4 First year student OR=1.8
El-Gibaly 2002	Predictors of believing female circumcision is necessary (women's view) Current age is 13-15 OR=0.60 Current age is 16-19 OR=0.38 Attended school in the past OR=0.34 Currently in school OR=0.21 Residency in urban, lower Egypt OR=0.62 Predictors of believing women should be circumcised before marriage (women's view) Residency in urban, lower Egypt OR=0.5 Currently in school OR=0.45 Predictors of believing women should be circumcised before marriage (men's view) High socioeconomic status OR=0.38 Predictors of being circumcised Mother's education is beyond preparatory school OR=0.66 Residency in urban, lower Egypt OR=0.72

Author	Results
Yount 2004	**Predictors of intending to circumcise daughter**[1] Mother has education: primary OR=0.44 (-.81), preparatory OR=0.34 (-1.08), secondary OR=0.44 (-.81), higher than secondary OR=0.15 (-1.93) Urban residence OR=0.27 (-1.32) Family is Muslim OR=5.99 (1.79) Father has education: preparatory OR=0.38 (-.97), secondary OR=0.37 (-.99) **Predictors of having circumcised daughter**[1] Mother is circumcised OR=13.07 (2.57) Mother has education: primary OR=0.68 (-.39), secondary OR=0.20 (-1.59), higher than secondary OR=0.06 (-2.75) Mother is over 35 years: age 35-44 OR=5.37 (1.68), age 45+ OR=7.46 (2.01) Family is Muslim OR=4.48 (1.50) Age of daughter is over 5 years: age 5-9 OR=7.54 (2.02), age 10-14 OR=21.76 (3.08), age 15+ OR=36.97 (3.61)
Afifi 2009	**Predictors of <u>not</u> intending to perform female genital cutting on daughter** Low empowerment, high education OR=6.73 High empowerment, high education OR=8.06 Urban residence OR=3.87 Working OR=1.24 Ever married woman with FGC OR=115.44

Note: [1] The results of the multivariate analyses were presented as log odds, which we converted to OR for consistency purposes, but we include the original log odds in parentheses.

We conducted six meta-analyses of reasons for FGM/C and four meta-analyses of reasons against FGM/C (figures 1-10) (Note: in the figures, El-Gibali 1999 should be El-Gibaly 2002).

Figure 1. Forest plot, reason for FGM/C is custom/tradition/good habit

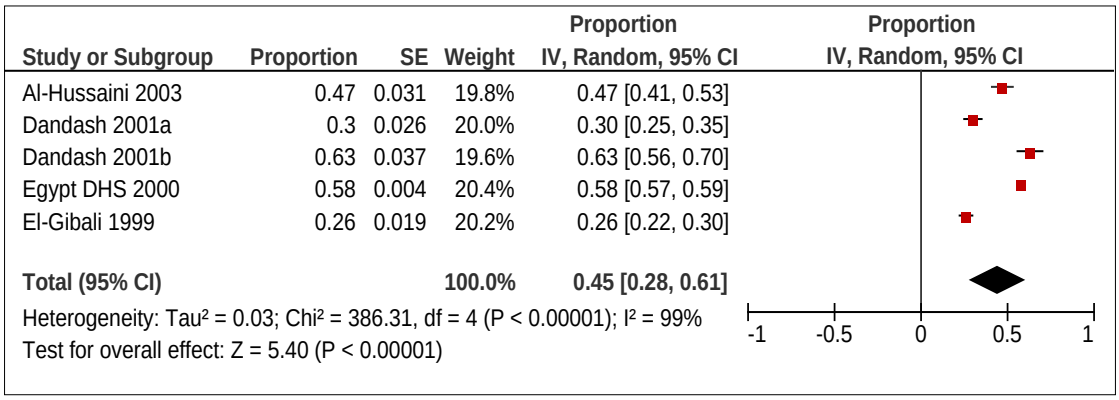


Figure 2. Forest plot, reason for FGM/C is religious tradition or significance

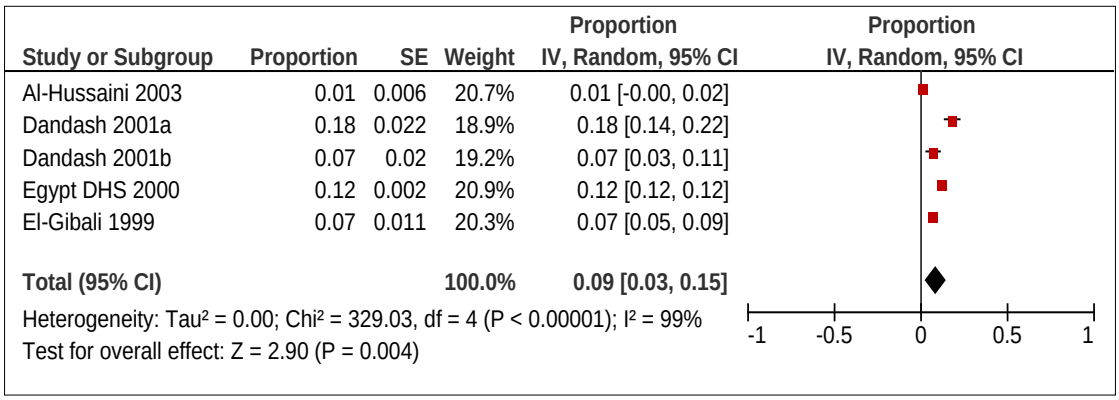


Figure 3. Forest plot, reason for FGM/C is hygiene/cleanliness

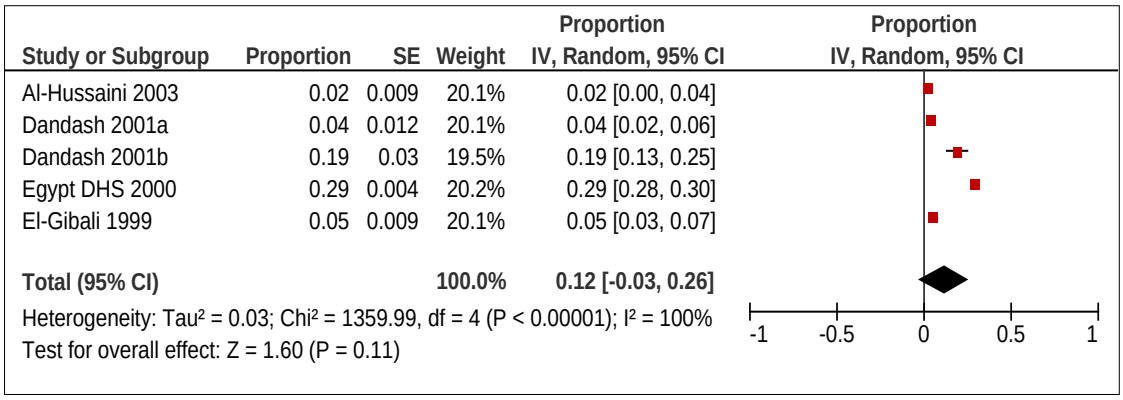


Figure 4. Forest plot, reason for FGM/C is reduce sexual desire/preserve virginity

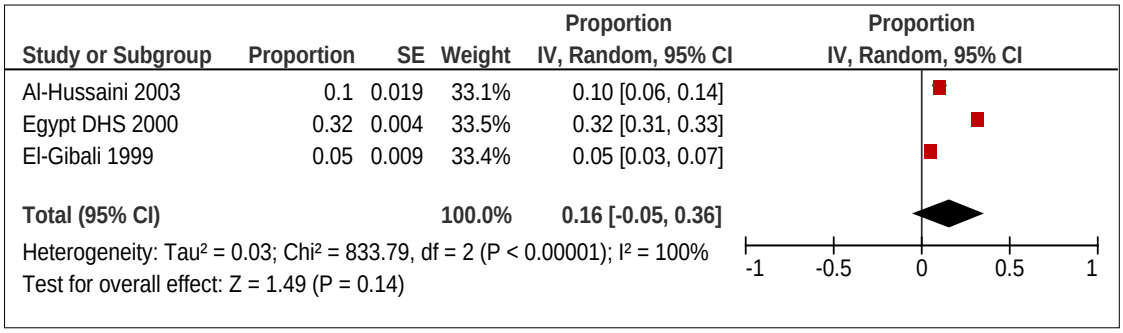


Figure 5. Forest plot, reason for FGM/C is (sexual) morals

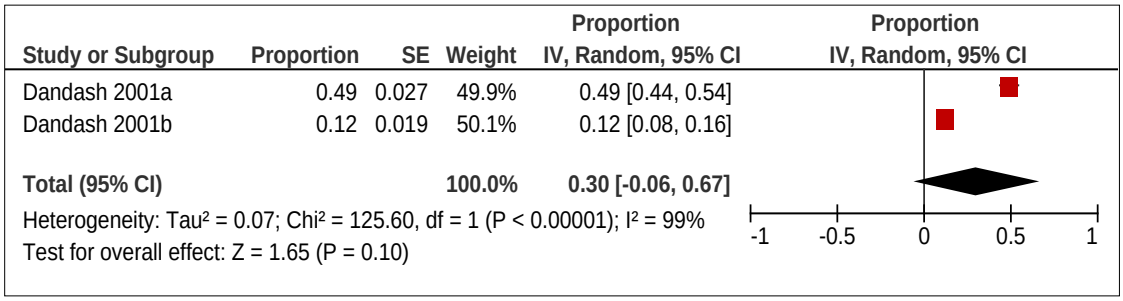


Figure 6. Forest plot, reason for FGM/C is marriageability

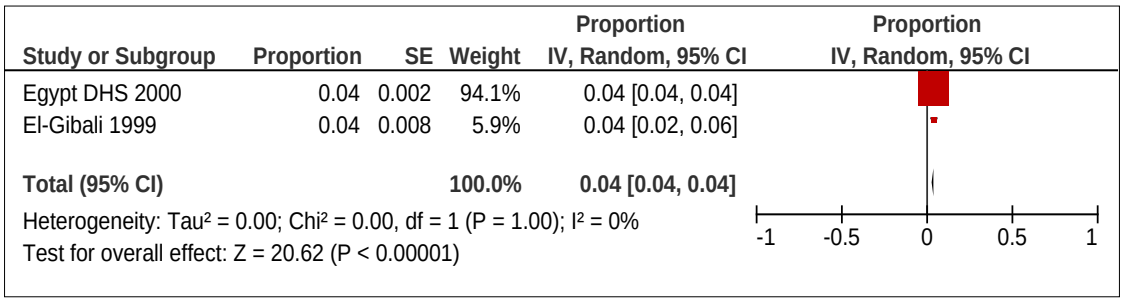


Figure 7. Forest plot, reason against FGM/C is it's against religion / not a religious requirement

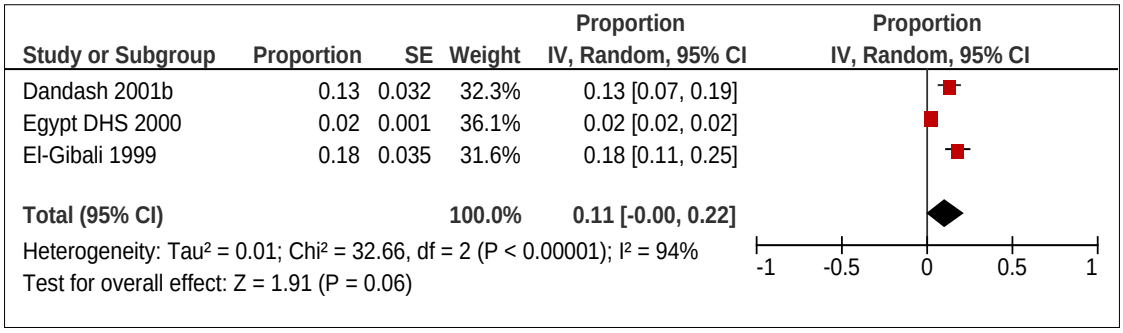


Figure 8. Forest plot, reason against FGM/C is fear of complications / harmful health outcomes

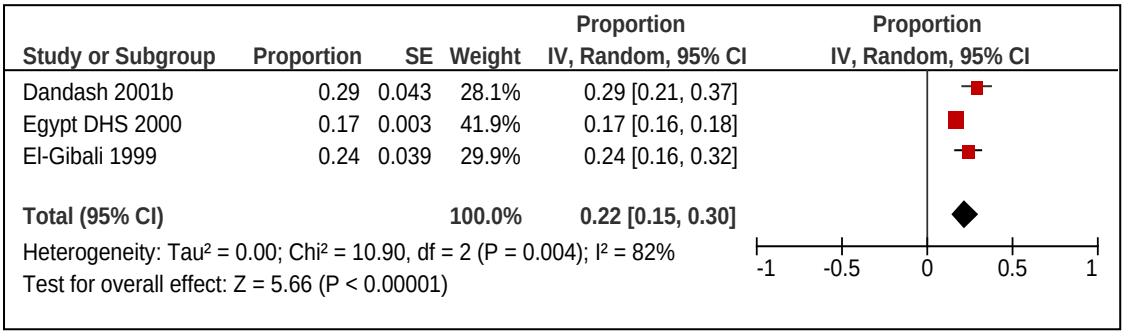


Figure 9. Forest plot, reason against FGM/C is it has no value or benefit

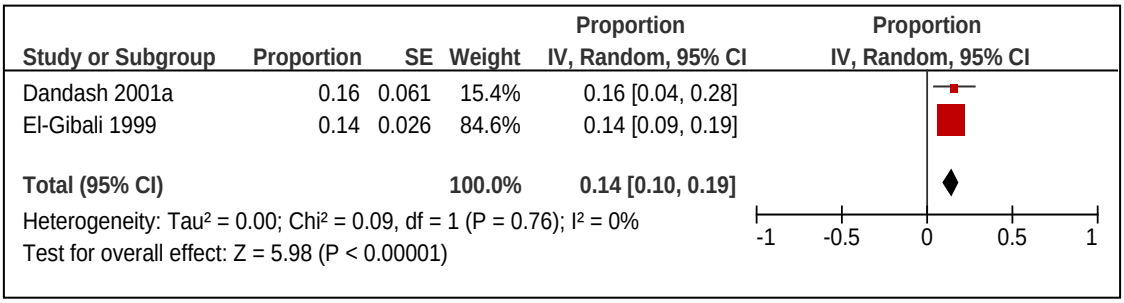
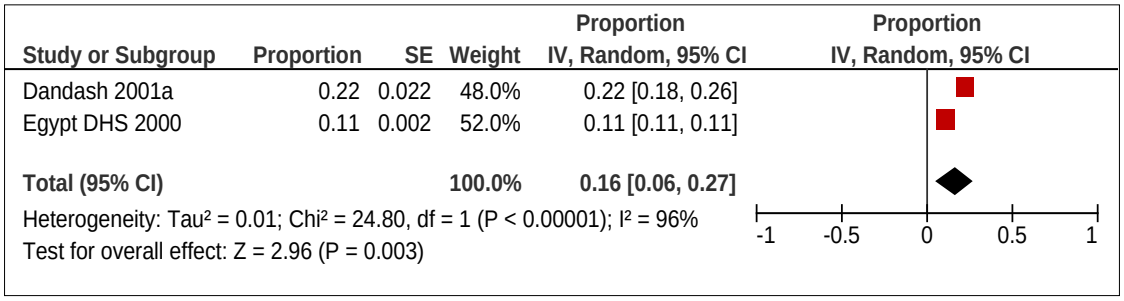


Figure 10. Forest plot, reason against FGM/C is sexual problems



5.22 NIGERIA

We included four studies from Nigeria which presented results of multivariate regression models examining predictors perpetuating and hindering FGM/C (table 9).

Table 9. Results of context studies from Nigeria, predictors perpetuating FGM/C

Author	Results
Abubakar 2004	Predictors of favouring continuation of female genital cutting Formal education OR=0.16 Being cut OR= 2.8
Freyermeyer 2007	Predictors favouring continuation of female genital cutting Being cut OR=39.75 Belonging to ethnic group with >50% of women being cut OR=1.70 Education (high) OR=0.92 Age (young) OR=0.98 Know of community activities to oppose FGC OR=0.42
Kandala 2009	Predictors of being cut Urban residence OR=1.26 Both (marriage) partners make healthcare decisions OR=1.35 Unmarried OR=0.76 Predictors of having cut daughter Urban residence OR=1.26 From high SES/rich household OR=0.56 Mother have some education OR=0.78 Mother is unmarried OR=0.67
Snow 2002	Predictors of being cut Ethnicity other than Bini, Urhobo, Esan, Yoruba OR=0.22 Age (older/45-49 yrs) OR=16.86 Religion Muslim OR=0.23 Educational attainment tertiary OR=0.33

We conducted nine meta-analyses of reasons for FGM/C and three meta-analyses of reasons against FGM/C (figures 11-22).

Figure 11. Forest plot, reason for FGM/C is tradition/culture

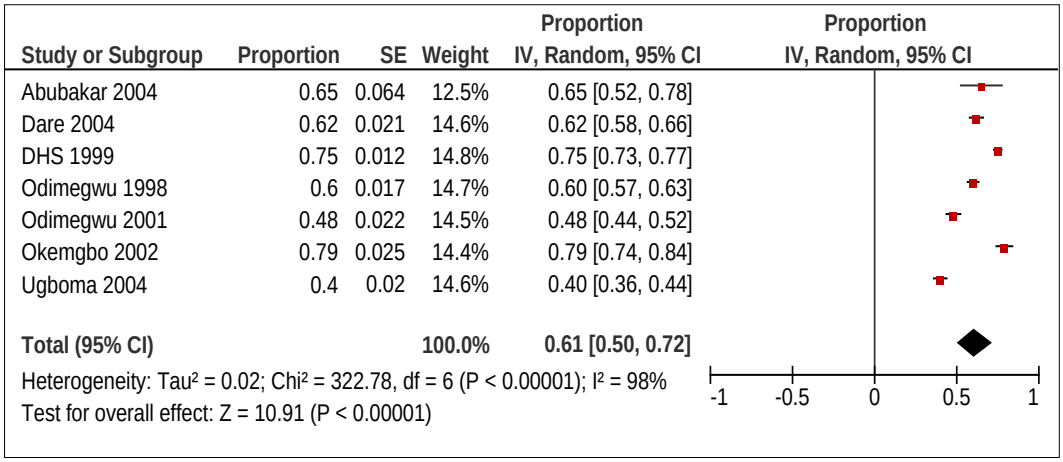


Figure 12. Forest plot, reason for FGM/C is prevention of promiscuity/preserve virginity

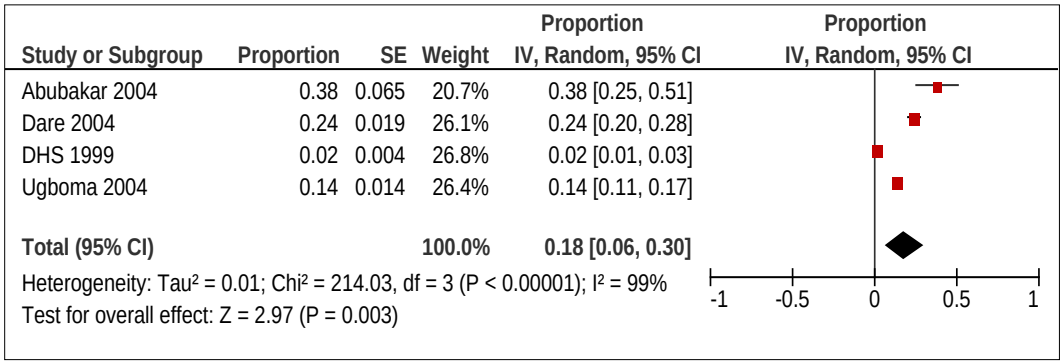


Figure 13. Forest plot, reason for FGM/C is religion

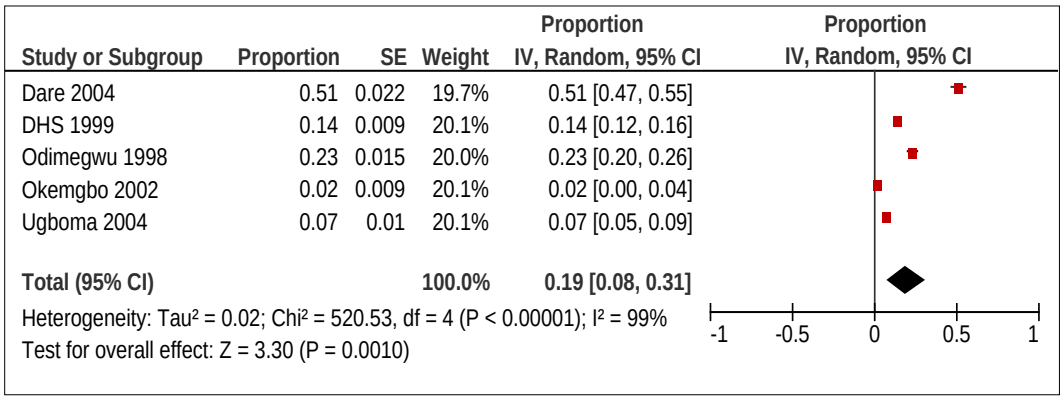


Figure 14. Forest plot, reason for FGM/C is to prevent pregnancy/ delivery problems

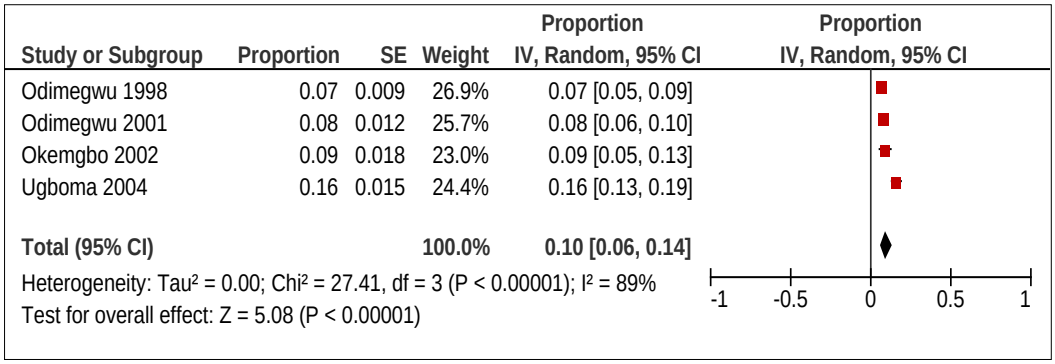


Figure 15. Forest plot, reason for FGM/C is to reduce/control female sexual desire

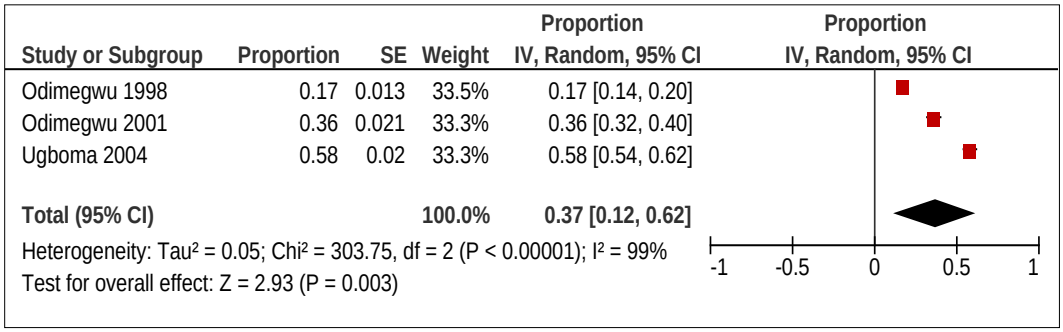


Figure 16. Forest plot, reason for FGM/C is to beautify female organ

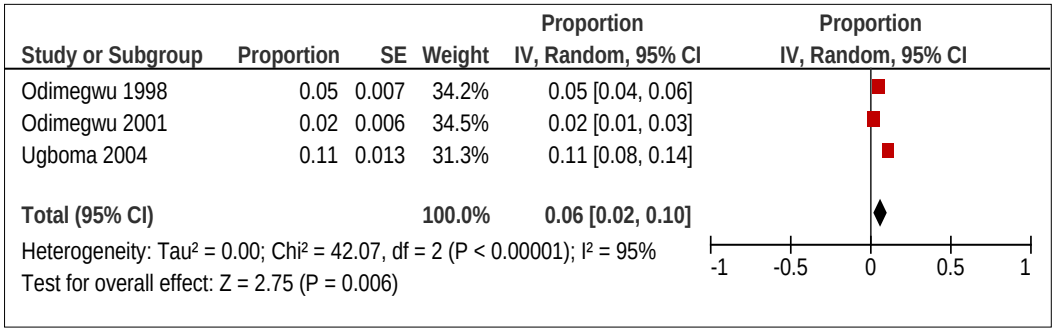


Figure 17. Forest plot, reason for FGM/C is cleanliness/hygiene

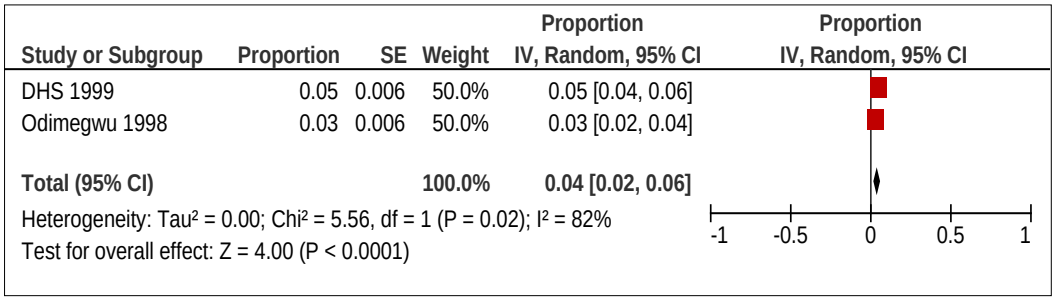


Figure 18. Forest plot, reason for FGM/C is better marriage prospects

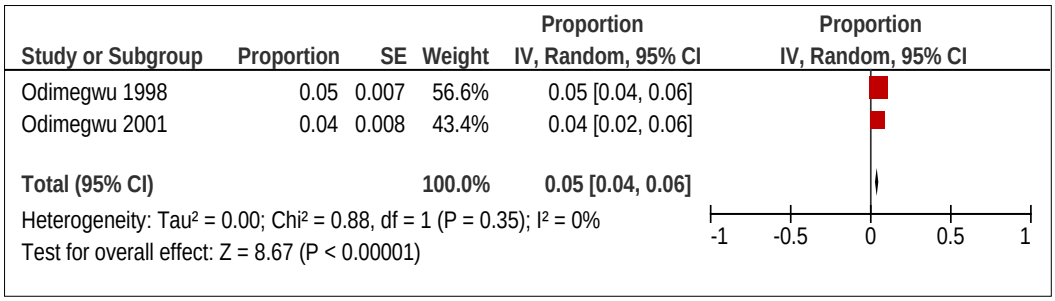


Figure 19. Forest plot, reason for FGM/C is that it is good for health

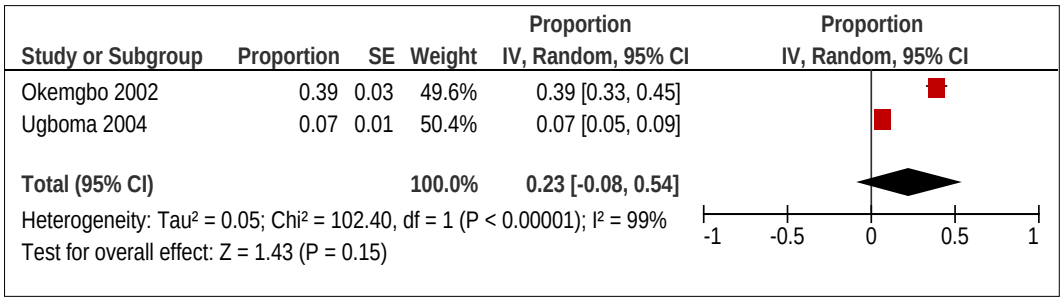


Figure 20. Forest plot, reason against FGM/C is medical complications/health problems

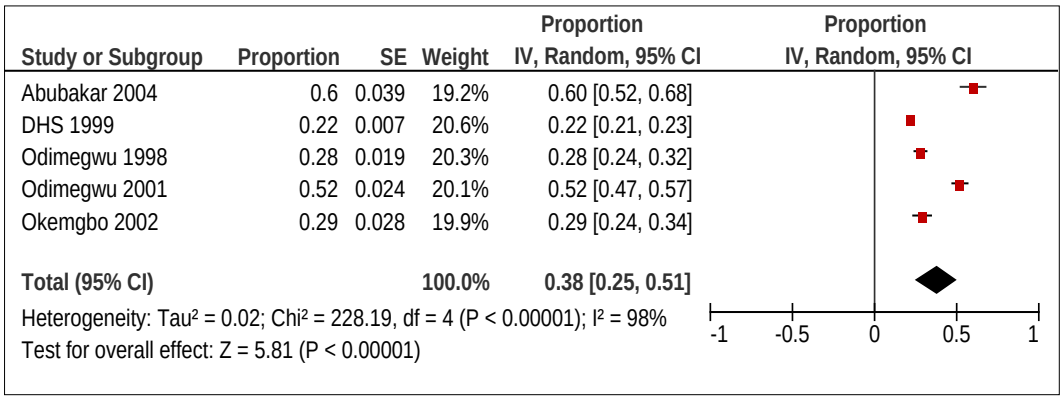


Figure 21. Forest plot, reason against FGM/C is it's a bad tradition

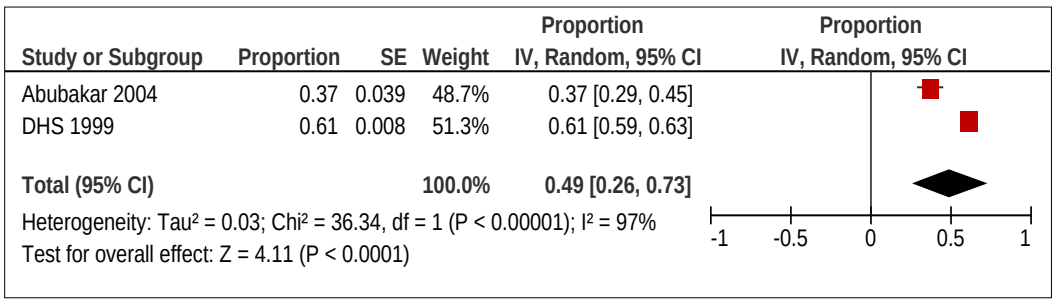


Figure 22. Forest plot, reason against FGM/C is it's not necessary