Effectiveness of Interventions Designed to Prevent Female Genital Mutilation/Cutting: A Systematic Review

Rigmor C. Berg and Eva Denison

Female genital mutilation/cutting (FGM/C) is widely considered a human rights infringement, although communities that practice the tradition view it as an integral part of their culture. Given these vastly different views, the effectiveness of efforts to abandon FGM/C is uncertain. We conducted a systematic review of the best available evidence regarding evaluations of interventions to prevent FGM/C, including eight controlled before-and-after studies with 7,042 participants from Africa. Findings indicate that 19 of 49 outcomes (with baseline similarity) were significantly different at study level, mostly favoring the intervention, but results from four meta-analyses showed considerable heterogeneity. The limited effectiveness and weak overall quality of the evidence from the studies appear related to methodological limitations of the studies and shortcomings in the implementation of the interventions. Nevertheless, the findings point to possible advantageous developments from the interventions. (Studies in Family Planning 2012; 43[2]: 135–146)

The practice of female genital mutilation/cutting (FGM/C), or female circumcision, is complex and controversial. Physically, it involves the partial or total removal of, or injury to, the female external genital tissue for nontherapeutic purposes. Cutting varies for the three most common procedures—clitoridectomy, excision, and infibulation—from partial removal of the clitoris to narrowing the vaginal orifice and creating a covering seal by cutting and appositioning the labia minora and majora after excision of the clitoris and prepuce (WHO 2008). Although FGM/C is considered by many to be a human rights infringement (WHO 2008), practicing communities consider it an integral part of their culture, deeply rooted in their social systems and beliefs (WHO 1999; UNFPA 2007).

WHO has calculated that 100–130 million girls and women have been subjected to some form of FGM/C in Africa, Asia, and the Middle East. Each year, more than a million girls are subjected to FGM/C, as estimated by population-based surveys from 28 African countries in which the practice is common and prevalence rates are high (WHO 1997 and 2008). Recent figures show a prevalence of more than 70 percent in Burkina Faso, Egypt, Ethiopia, Mali, Sierra Leone, and Somalia (Yoder and Kahn 2008).

FGM/C is recognized as a harmful practice with negative health consequences for women ranging from pain, bleeding, and shock to chronic infections and difficulty passing urine and feces (WHO 2008). A review and study of the health complications associated with the practice (WHO 2000; WHO Study Group 2006) identified several obstetric problems, the most common of which are prolonged and/or obstructed labor, episiotomies, and perineal tears. A more recent systematic review concluded that women who have undergone FGM/C are more likely to experience pain during intercourse and reduction in sexual satisfaction and desire (Berg and Denison 2011). For many girls and women, undergoing FGM/C is a traumatic experience that leaves a lasting psychological mark (WHO 2008; HRP 2011).

In addition to statements against FGM/C by international and professional organizations (WHO 2008 and 2010), laws explicitly prohibiting the practice exist in many African countries, including Burkina Faso, Egypt, Eritrea, Ethiopia, and Senegal. In some countries, general provisions of existing criminal codes can be applied to FGM/C (Mali and North Sudan), whereas no laws against the practice exist in Somalia and the Gambia (PRB 2008). Considering the omnipresence of FGM/C in Africa, continued abandonment efforts are needed. According to Muteshi and Sass's analysis (2005), abandonment efforts

Rigmor C. Berg is Adjunct Assistant Professor and Eva Denison is Professor, Norwegian Knowledge Centre for the Health Services, P.O. Box 7004, St. Olavs Plass, N-0130 Oslo, Norway. E-mail: rigmor.berg@nokc.no.

are based on human rights and legal mechanisms, health-risk approaches, training health workers and converting circumcisers, and comprehensive social development approaches. The Population Reference Bureau conducted an extensive survey of "best practices" interventions in African countries and identified 27 evaluated projects (PRB 2007). Much of the identified research used observational designs—that is, designs in which the investigators do not seek to intervene (Higgins and Green 2011)—making the drawing of causal inferences difficult.

Evaluation studies and surveys of best practices can assess the effectiveness of some FGM/C prevention interventions in achieving desired outcomes. What is lacking is an updated systematic appraisal of the best scientific evidence that draws valid conclusions about the effectiveness of interventions. This article fills this gap by presenting a systematic review of the best available evidence on interventions designed to prevent FGM/C. This study is based on and is an update of a 2009 report (Denison et al. 2009).

Methods

We conducted a systematic review in accordance with guidelines set forth in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins and Green 2011). The main literature identification strategy consisted of searches in the following 13 international electronic databases: African Index Medicus, Anthropology Plus, British Nursing Index and Archive, CENTRAL, EMBASE, EPOC, MEDLINE, PILOTS, POPLINE, PsycINFO, Social Services Abstracts, Sociological Abstracts, and WHOLIS. The searching of databases was conducted by an information search specialist using a strategy incorporating prespecified subject headings (for example, MeSH terms in MEDLINE) and text words, in title and abstract, adapted for each database. The search strategy for MEDLINE is shown in endnote 1.[1] We supplemented the database searches with searches in the reference lists of included studies and relevant reviews, and in databases of international organizations engaged in FGM/C projects. We also contacted experts engaged in FGM/C-related projects. These processes were repeated until no new references were identified. The searches were completed in March 2011.

For the search and screening processes, we applied the (S)PICO model, which directs attention to the study design, population, intervention, comparison, and outcomes (Sackett et al. 2000). Unpublished reports, briefs, and preliminary reports were considered for inclusion on the same basis as published reports. Study participants included girls and women at risk of FGM/C, specific members (such as health-care personnel) of communities where FGM/C is practiced, and other residents of communities where FGM/C is practiced. Analyses of any interventions or programs designed to prevent FGM/C were considered, including education about the practice, training of health workers as agents of change toward alternative rites, and the positive deviance approach, which identifies individuals who have challenged or deviated from societal expectations of FGM/C by publicly declaring opposition, abandoning the practice, or urging others to do so (Muteshi and Sass 2005). Such interventions could be compared to any other program or to none. We accepted for inclusion only those studies with designs containing a comparison group: randomized controlled trials and controlled before-and-after studies (Higgins and Green 2011). Outcomes of interest were quantitatively reported rates of FGM/C, proportions of the sample in favor of abandoning the practice, related behaviors, awareness of rights, knowledge of harmful consequences, and beliefs and attitudes related to FGM/C. We included all publication years and languages. When considered likely to meet the inclusion criteria, studies were translated into English.

Literature screening was carried out in a three-stage process, with each level consisting of increased scrutiny of the records based on the inclusion criteria of the review. At each level, we evaluated the identified records independently of one another using a predeveloped inclusion form, and then evaluated the records jointly. The final determination to include or exclude was made together. Throughout the process of screening titles, abstracts, and full texts, differences were resolved through discussion and consensus. After obtaining and assessing the full text of records promoted to level-three screening (the full text could not be obtained for two studies), only those studies meeting all inclusion criteria were included. We were not blinded to the authors or other information when assessing the records.

Data Extraction and Analysis

We independently extracted data from the published sources using a predesigned data recording form. Differences in extracted data were resolved through reexamination of the reports and discussion. We used the McMaster University, Effective Public Health Practice Project, Quality Assessment Tool for Quantitative Studies (Thomas n.d.) to assess methodological quality. The tool considers eight factors rated according to prespecified criteria. The factors are: selection bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, intervention integrity, and analyses. We applied a global rating of strong, moderate, or weak methodological quality for each study after resolving any

discrepancies in our assessments with respect to component ratings. Further, we applied the instrument Grading of Recommendations Assessment, Development, and Evaluation (GRADE) to assess the extent to which we could have confidence in the estimates of effects (Higgins and Green 2011). We used the four standard definitions in grading the quality of the evidence (high, moderate, low, very low). (For more information about the GRADE system, see the publications by the GRADE Working Group [Guyatt et al. 2008 and 2011].)

We present data for outcomes when pre- and post-scores for both intervention and comparison groups were reported by study authors, allowing for comparison. We used two measures to estimate the effects of interventions: 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, and relative risk (RR) based on postintervention data. We present continuous data with mean differences and p-values. If studies were sufficiently similar, we grouped them together and applied the statistical techniques of meta-analysis to estimate the effect. To be pooled, the same outcome had to be assessed in a similar manner, in similar populations, across similar intervention studies. We used Mantel-Haenszel random effects meta-analysis for dichotomous outcomes, and inverse-variance random effects meta-analysis for continuous outcomes. We also examined between-study heterogeneity using Chi-square and I-square tests. We used RevMan version 5.0.25, the latest version of the Cochrane Collaboration's meta-analysis software.

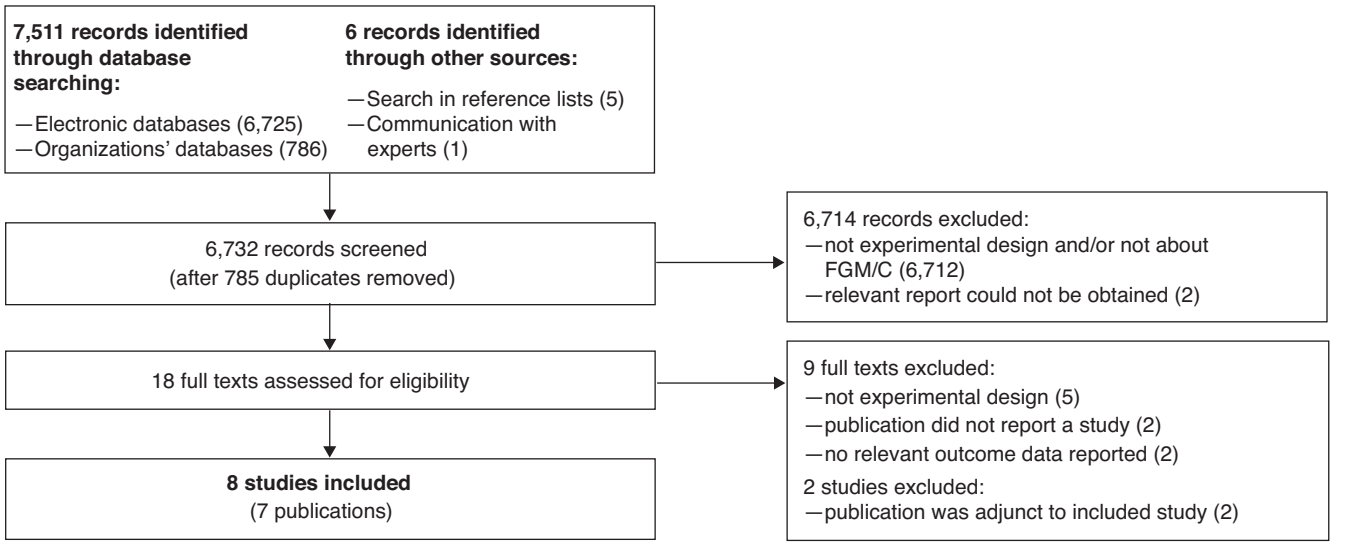
Results

The PRISMA flow diagram (Figure 1) shows that the search yielded 7,517 records. Of the 18 potentially relevant publications, the full text for two could not be obtained: Babalola and Adebajo (1985) and Akweongo and colleagues (2007). Nine publications were excluded because they either were not experimental designs, did not report a study, or reported no relevant outcome data. Two studies were not included because they overlapped with selected studies (Monkman, Miles, and Easton 2007; Diop and Askew 2009). Finally, we summarized and included results from eight studies (Diop et al. 1998; Easton, Miles, and Monkman 2002; Mounir, Mahdy, and Fatohy 2003; Chege et al. 2004 [Ethiopia and Kenya]; Diop et al. 2004; Ouoba et al. 2004; Babalola et al. 2006). Two of the final studies were published in peer-reviewed journals (Mounir, Mahdy, and Fatohy 2003; Babalola et al. 2006); the other six were reports to funding agencies.

Quality Assessment

Based on the component ratings of the McMaster Quality Assessment Tool, we arrived at a final assessment of "weak study quality" for all eight included studies. None were randomized controlled studies. Rather, the eight studies employed a controlled before-and-after design. The six studies that evaluated community-based interventions appear to have used cross-sectional and independent samples at baseline and follow-up (except possibly Easton, Miles, and Monkman 2002). Aside from the lack of randomization in all studies, the design was likely

Figure 1 PRISMA flow diagram of the literature review process (7,517 records)



contaminated by population movements into and out of the intervention and comparison areas in three studies (Easton, Miles, and Monkman 2002; Chege et al. 2004 [two countries studied]) and by the presence of another FGM/C awareness-raising program in two intervention areas (Easton, Miles, and Monkman 2002; Diop and Askew 2009). Additionally, differences regarding similarity of prognostic factors at baseline existed between the intervention and comparison groups in five studies (Easton, Miles, and Monkman 2002; Chege et al. 2004 [Kenya and Ethiopia]; Ouoba et al. 2004; Babalola et al. 2006). Prognostic indicators (education, prevalence of FGM/C, and religious affiliation) were based on suggestions made by several agencies (UNICEF 2005a; PRB 2008; WHO 2008).

We assessed the quality of the evidence through GRADE. Because all the studies were non-randomized, with weak methodological quality, the evidence was graded as low quality, meaning further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate (GRADE Working Group 2004; Guyatt 2008).

Descriptions of Included Interventions

The eight studies collectively included 7,042 participants, all from African countries: Burkina Faso, Egypt, Ethiopia, Somalia/Kenya, Mali (two studies), Nigeria, and Senegal (see Table 1). The mean age of participants was 32 years. Most had little education, and illiteracy was common. Six interventions were delivered at the community level; two (Diop et al. 1998; Mounir, Mahdy, and Fatohy 2003) consisted of educational activities at the individual level. The interventions included training, educational sessions, advocacy, empowerment, and mass media campaigns. The duration of the intervention varied from two weeks at the individual level to an average of 11 months at the community level. With the exception of the Kenya study by Chege and colleagues (2004), only one category of comparison was used: those who did not participate in the intervention. None of the studies collected biological data. All information was self-reported knowledge/awareness, beliefs/attitudes, intentions, behaviors, and prevalence of FGM/C, col-

Table 1 Descriptions of the eight included FGM/C intervention studies

Topic/study author/ year (country)	Population	Intervention	Comparison	Outcomes
Training of health personnel				
Diop et al. 1998 (Mali)	N = 108, health personnel (obstetricians, gynecologists, family planning providers). Majority ethnic group Bambara (34%), median age 38, 96% Muslim, 92% FGM/C prevalence.	Training clinic staff about FGM/C, supervising trained clinic staff (8 sites). Intervention duration = 2 months in 1997.[a]	6 health sites/ clinics with no intervention	Knowledge, beliefs/attitudes, experiences
Education of female students				
Mounir, Mahdy, and Fatohy 2003 (Egypt)	N = 682, female Alexandria University students. Mean age 19, lived in the university hostels, 47% from low-social-class families, 87% FGM/C prevalence.	Two 60-minute educational sessions about reproductive health, including FGM/C (Ezbet Saad hostel). Intervention duration = 2 weeks in 2001.[a]	El-Shatby hostel with no intervention	Knowledge, beliefs/attitudes
Multifaceted community activities				
Chege et al. 2004 (Kenya)	N = 1,440 (360 women), Somali community members in refugee camps. Mean age 31, 61% below primary education level, primarily Muslim, 74% unemployed, 100% FGM/C prevalence.	Behavior-change communication activities and advocacy. 1 refugee camp (Ifo). Intervention duration = 18 months in 2001–02.	1 refugee camp with no intervention	Knowledge, beliefs/attitudes, intentions
Chege et al. 2004 (Ethiopia)	N = 819 (202 women), community members in 12 villages. Afar people, mean age 26, 82% below primary education level, primarily Muslim, 91% FGM/C prevalence.	Behavior-change communication activities and advocacy. 6 villages. Intervention duration = 15 months in 2001–02.	6 villages with no intervention	Knowledge, beliefs/attitudes, intentions
Babalola et al. 2006 (Nigeria)	N = 957 (531 women), community members in eastern Nigeria. Mean age 34, 66% less than secondary education, 58% protestant, ~100% FGM/C prevalence.	Community mobilization, advocacy, mass media activities (Enugu State). Intervention duration = 12 months[a] in 2003–04.	Ebonyi State with no intervention	Behaviors, beliefs/attitudes, intentions
Village empowerment				
Easton, Miles, and Monkman 2002 (Mali)	N = 239 (132 women), community members in 9 villages. Majority ethnic group Bambara, mean age 36, 69–87% no education, primarily Muslim, majority FGM/C.	Education in hygiene, problem solving, women's health, human rights. 5 villages. Intervention duration = 6 months in 2000.	4 villages with no intervention	Beliefs/attitudes
Diop et al. 2004 (Senegal)	N = 1,332 (775 women), community members in 40 villages. Mainly Pulaar and Mandingo ethnic groups, mean age 35, 78% no education, 99% Muslim, ~ 88% FGM/C prevalence.	Education in hygiene, problem solving, women's health, human rights. 20 villages. Intervention duration = 6 months in 2001.	20 villages with no intervention	Prevalence, knowledge, beliefs/attitudes, intentions
Ouoba et al. 2004 (Burkina Faso)	N = 1,465 (718 women), community members in 46 villages. Mossi ethnic group, mean age 36, 90% no education, 45% Muslim, > 80% FGM/C prevalence.	Education in hygiene, problem solving, women's health, human rights. 23 villages. Intervention duration = 8 months in 2001–02.	23 villages with no intervention	Prevalence, behaviors, knowledge, beliefs/attitudes, intentions

[a] Not explicitly stated, but text suggests it.

lected through face-to-face structured interviews or paper questionnaires.

Training of Health Personnel

Diop and colleagues (1998) examined the effectiveness of training and supervising Mali health workers regarding attitudes toward FGM/C, experiences and skill caring for women, and knowledge of complications. Health clinic staff at eight sites in southern Mali received three group training sessions that 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 [were] emphasized" (Diop et al. 1998: 16). Obstetricians, gynecologists, and family planning providers were subsequently supervised during clinic duty. The study was conducted among ethnic groups with high FGM/C prevalence rates (92 percent). Study authors did not detail program fidelity (the degree to which procedures/components of the intervention were implemented as planned [Higgins and Green 2011]), but report that loss to follow-up was only 4 percent.

Education of Female Students

The study of the intervention in Egypt (Mounir et al. 2003) that targeted female students examined the effectiveness of two educational sessions in shaping participants' beliefs and knowledge. Intervention participants received information about reproductive health, including FGM/C, in two one-hour sessions consisting of talks, group discussions, role playing, and educational aids. Study authors did not report information about program delivery.

Multifaceted Community Activities

Somali Refugees in Kenya. Although the first study in the report by Chege and colleagues (2004) took place in Kenya, Somali refugees comprised the intervention group. Among the Somali, all women of reproductive age have undergone some type of FGM/C. The 18-month intervention at the Ifo refugee camp consisted of community-level educational outreach: behavior-change communication and community-level advocacy (educational events, community meetings, theater performances, video sessions, mass media campaigns, and advocacy activities). According to the study description, implementation was problematic. The comparison group, the Hagadera refugee camp in Kenya, participated in educational activities (details were not provided). Fifty-nine percent of the participants in the intervention group reported that they had been exposed to FGM/C abandonment messages, compared with 47 percent of participants in the comparison group. The message remembered most was that the practice was harmful to health.

Afar People in Ethiopia. Chege and colleagues (2004) also assessed a multifaceted intervention that took place among the Afar population in six Ethiopian villages. Six other villages with similar FGM/C prevalence (91 percent) served as the comparison site. The intervention was nearly identical to that among Somali refugees in Kenya. Approaches for addressing human rights and gender differences varied between the two sites because of dissimilar cultural contexts; the Ethiopian intervention was also three months shorter. As in the Somali refugee camp in Kenya, the measures of the effect of the Ethiopian intervention were awareness of human rights, knowledge of the harmful effects of FGM/C, and various beliefs and intentions regarding the practice. Seventy-one percent of Ethiopian study participants recalled being exposed to FGM/C abandonment messages.

Residents in Enugo State, Nigeria. The multifaceted Nigerian program (Babalola et al. 2006) was delivered at three community levels: hamlet ("village"), local government area, and state. At the state level, the multimedia program *Ndukaku*, which is the Igbo word for "health is better than wealth," dominated (for example, in newspaper columns and radio call-in shows), but other activities, such as development of action plans to improve conditions for women, also took place. Outcomes measured to assess effectiveness included beliefs, intentions, self-efficacy, and whether the participants encouraged others not to perform FGM/C. In Nigeria, virtually 100 percent of women in the intervention and comparison sites had experienced FGM/C, but considerable variation in prevalence was found by local government area. Program exposure was measured through recall of messages. More than 36 percent of participants were not exposed to any program components.

Empowerment through Education (TOSTAN Program)

Villagers in the Kati Region, Mali. The TOSTAN empowerment intervention took place in five Kati-region villages (Easton et al. 2002; Monkman et al. 2007). Four demographically similar comparison villages were chosen in the same geographic area. According to the study's authors, the majority of Mali families practice FGM/C. Thus, the intervention, which consisted of four educational modules covering hygiene, problem solving, women's health, and human rights, was designed to empower women by providing them with "a participatory approach to stopping FGC" (Easton et al. 2002: 3). The

focus was on enabling participants to analyze their situation and find the best solution. The only outcome measured with respect to FGM/C was participants' position regarding the practice (support, mixed, oppose). In large part because of disagreements between the organizations involved, the program ended abruptly, which complicated accurate measurement of implementation and effectiveness. (For a description of the TOSTAN program, see Diop et al. 2008; Gillespie and Melching 2010.)

Kolda Villagers in Senegal. The empowerment intervention in Senegal was delivered in 20 villages in the Kolda region, which is characterized by 88 percent FGM/C prevalence (Diop et al. 2004; Diop and Askew 2009). Twenty similar but geographically distant villages served as the comparison group. To assess intervention effectiveness, changes were measured in knowledge of harmful consequences, various beliefs and intentions, and prevalence. Female respondents were asked whether their daughters younger than 11 years of age had been subjected to FGM/C. Intervention organizers encountered some obstacles, including that only about half of those who had expressed interest in the program enrolled, and that many community members, especially men, expressed discontent toward the intervention. Sixty-nine percent of participating women and 57 percent of men attended all four educational modules, with the remainder completing part of the program.

Mossi People in Burkina Faso. The Senegal empowerment intervention was carried out in 23 villages in Burkina Faso (Ouoba et al. 2004). Although FGM/C is practiced by most ethnic groups in Burkina Faso, greatest prevalence (>80 percent) is in the central Mossi plateau, where the study took place. As in Senegal, recruitment was difficult, but although not all of those who had expressed interest in the program enrolled, others participated too. Overall, 63 percent of participating women and 60 percent of men attended all four educational modules, with the remainder completing part of the program. Outcomes for assessment of effectiveness were the same as in the Senegal study.

Results of Interventions at Study Level

Table 2 shows the results for each outcome in each study for which we could calculate effect estimates. In the majority of cases, prevalence was not measured, and other data, such as behaviors and intentions regarding FGM/C, were often not presented in such a way that calculation of effect estimates was possible. Therefore, they are not presented here. For the study evaluating the effectiveness of reproductive health education (Mounir et al. 2003), the effect estimate for only one outcome could be calculated; thus, we excluded this study from the table. The mean knowledge score about dangers of FGM/C increased by 0.47 points in the intervention group and by 0.02 points in the comparison group, making the mean difference at endline significantly different (MD = 0.75 points on a 0–3 scale; $p < 0.05$).

Results of Meta-Analyses

Results of the meta-analyses are presented in Table 3. The two studies contained in Chege and colleagues (2004)—of interventions among Somalis in Kenya and among Ethiopians—were nearly identical and included the same outcomes, warranting the pooling of effect sizes in meta-analysis. One outcome was reported in both studies with baseline similarity between the groups. The meta-analysis result for this outcome, belief that FGM/C compromises the human rights of women, was not significant (RR = 1.30; $p > 0.05$) and considerable heterogeneity existed ($I^2 = 98$ percent).

Two studies that evaluated effectiveness of community empowerment through education (Diop et al. 2004; Ouoba et al. 2004) were sufficiently similar to warrant pooling of effect sizes in meta-analysis. We were able to assess three dichotomous outcomes. The results of the meta-analyses for the outcome "knowledge of harmful consequences of FGM/C" were significant for men (RR = 2.11; $p < 0.05$) but not for women (RR = 1.85; $p > 0.05$), and the results revealed considerable heterogeneity ($I^2 = 95$ percent and 98 percent, respectively). The last outcome for which meta-analysis was possible was prevalence of FGM/C among girls. The result was significant (RR = 0.77, $p < 0.05$), but one study had a higher number of events than the other and therefore assumed the most weight and contributed disproportionately to the pooled-effect size. For all four meta-analysis results, the considerable heterogeneity and unequal weight affect the interpretation of effect sizes, raising doubt about the validity of the results.

Discussion

The overall quality of the evidence from the FGM/C abandonment interventions included here was weak. Results indicate that effectiveness of training health personnel in Mali is slight, at best, which may result from the limited time provided for training. The study's authors concede that "training was apparently too short and should be expanded" (Diop et al. 1998: 35). WHO's position in regard to the role of healthcare providers in FGM/C abandonment is that they "can play a key role in preventing female genital mutilation and in supporting and informing patients and communities about the

Table 2 Percentage of respondents, by intervention and study outcomes

Intervention/study/outcome	Intervention			Comparison			Absolute risk difference	Relative risk	95% confidence interval
	Pre	Post	Change	Pre	Post	Change			
Training of health personnel									
Diop et al. 1998 (Mali)									
Could name any type of FGM/C[a]	76	95	19	47	81	34	–15	1.12	(0.97–1.30)
Could name at least 3 types of long-term FGM/C complications[a]	50	72	22	61	73	12	10	0.99	(0.79–1.26)
Believed FGM/C poses no health risk if done in hygienic environment[a]	29	44	15	9	29	20	–5	1.54	(0.91–2.60)
Wished to play role in educating health clinic clients about FGM/C	93	91	–2	89	86	–3	–1	1.07	(0.93–1.23)
Multifaceted community activities									
Chege et al. 2004 (Kenya)									
Knew of harmful consequences of FGM/C[a]	57	91	34	71	89	18	16	1.02	(0.99–1.06)
Believed FGM/C compromised human rights of women	25	31	6	27	40	13	–7	0.77*	(0.67–0.89)
Supported abandonment of FGM/C in own community[a]	23	23	0	11	19	8	–8	1.21	(0.99–1.48)
No intention to perform FGM/C on daughter	14	17	3	10	18	8	–5	0.94	(0.75–1.17)
Chege et al. 2004 (Ethiopia)									
Knew of harmful consequences of FGM/C	33	87	54	36	71	35	19	1.37*	(1.26–1.49)
Believed FGM/C compromised human rights of women	8	42	34	10	19	9	25	2.21*	(1.75–2.79)
Supported abandonment of FGM/C in own community[a]	22	54	32	15	25	10	22	2.16	(1.78–2.62)
No intention to perform FGM/C on daughter[a]	8	34	26	14	13	–1	27	2.62*	(1.96–3.49)
Babalola 2006 (Nigeria)									
Women									
Had encouraged someone not to perform FGM/C on daughter	16	24	8	11	9	–2	10	2.68*	(1.76–4.08)
Did not believe there were benefits to FGM/C[a]	58	75	17	66	72	6	11	1.04	(0.95–1.15)
Disapproved of FGM/C[a]	63	88	25	71	73	2	23	1.21	(1.11–1.13)
Believed most community members favored discontinuation of FGM/C[a]	36	49	13	21	14	–7	20	3.50	(2.58–4.76)
Believed had self-efficacy to resist spousal pressure to perform FGM/C on daughter[a]	57	72	15	40	42	2	13	1.71	(1.47–1.99)
No intention to perform FGM/C on daughter	59	76	17	64	67	3	14	1.13*	(1.02–1.26)
Men									
Had encouraged someone not to perform FGM/C on daughter	12	14	2	10	12	2	0	1.19	(0.71–2.01)
Did not believe there were benefits to FGM/C	53	76	23	58	65	7	16	1.17*	(1.02–1.33)
Disapproved of FGM/C[a]	67	77	10	30	30	0	10	2.57	(2.06–3.20)
Believed most community members favored discontinuation of FGM/C	24	35	11	25	20	–5	16	1.76*	(1.25–2.47)
Believed had self-efficacy to resist spousal pressure to perform FGM/C on daughter[a]	72	88	16	51	40	–11	27	2.20	(1.85–2.61)
No intention to perform FGM/C on daughter[a]	53	73	20	64	66	2	18	1.11	(0.97–1.27)
Village empowerment									
Easton, Miles, and Monkman 2002 (Mali)									
Was opposed to FGM/C[a]	8	82	74	25	28	3	21	2.25	(1.62–5.34)
Diop et al. 2004 (Senegal)									
Women									
Girls aged 10 years and younger who had been cut	54	40	–14	52	52	0	–14	0.77*	(0.64–0.93)
Knew at least two consequences of FGM/C	7	73	66	3	25	22	44	2.92*	(2.28–3.74)
Believed FGM/C was unnecessary[a]	30	85	55	12	39	27	28	2.18	(1.82–2.61)
Disapproved of FGM/C[a]	28	84	56	11	40	29	27	2.10	(1.76–2.51)
Believed that husband disapproved of FGM/C[a]	35	86	51	14	46	32	19	1.87	(1.60–2.18)
Regretted having had daughter cut[a]	66	90	24	52	58	6	18	1.55	(1.37–1.76)
No intention to perform FGM/C on daughter[a]	29	88	59	11	46	35	24	1.91	(1.64–2.23)
Men									
Knew at least two consequences of FGM/C	11	66	55	14	21	7	48	3.10*	(2.28–4.23)
Believed that FGM/C was supported by religion[a]	14	34	20	25	48	23	–3	0.71	(0.51–0.99)
No intention to perform FGM/C on daughter[a]	34	87	53	22	44	22	31	1.97	(1.65–2.36)
Ouoba et al. 2004 (Burkina Faso)									
Women									
Girls aged 10 years and younger who had been cut	6	3	–3	4	4	0	–3	0.74	(0.33–1.66)
Discussed FGM/C with others[a]	54	91	37	64	65	1	36	1.40	(1.27–1.55)
Knew at least two consequences of FGM/C	52	86	34	57	73	16	18	1.18*	(1.08–1.29)
Believed that FGM/C was unnecessary	93	99	6	92	97	5	1	1.02	(1.00–1.05)
Disapproved of FGM/C	89	98	9	90	94	4	5	1.04*	(1.01–1.08)
Believed that husband disapproved of FGM/C	97	99	2	97	96	–1	3	1.03	(1.00–1.06)
Regretted having had daughter cut	53	81	28	50	64	14	14	1.26*	(1.14–1.40)
No intention to perform FGM/C on daughter	97	99	2	96	98	2	0	1.01	(0.99–1.03)
Men									
Discussed FGM/C with others[a]	62	92	30	79	75	–4	34	1.22	(1.13–1.33)
Knew at least two consequences of FGM/C	64	88	24	66	60	–6	30	1.47*	(1.31–1.64)
Believed that FGM/C was unnecessary	97	98	1	97	92	–5	6	1.06*	(1.02–1.11)
Disapproved of FGM/C	90	98	8	93	89	–4	12	1.10*	(1.01–1.09)
No intention to perform FGM/C on daughter	97	98	1	98	93	–5	6	1.05*	(1.01–1.09)

*Statistically significant at $p < 0.05$.

[a]Differences of 10 percent or more existed between the intervention and comparison groups at baseline for many of these variables, which casts doubt on the validity of any differences in their post scores.

Notes: Pre- and post-scores are reproduced from the study publication. We calculated percent changes, adjusted absolute risk differences, and relative risks with p-values and 95 percent confidence intervals.

Table 3 Results of meta-analysis, by outcome variable and study, according to events and risk ratio

Intervention/outcome/study	Intervention Events	Intervention Total	Comparison Events	Comparison Total	Weight (%)[a]	Risk ratio	95% confidence interval
Belief that FGM/C violates human rights							
Chege et al. 2004 (Kenya)	223	720	288	720	50.4	0.77	(0.67–0.89)
Chege et al. 2004 (Ethiopia)	168	400	76	400	49.6	2.21	(1.75–2.79)
Total	391	1,120	364	1,120	100.0	1.30	(0.46–3.66)
Heterogeneity: $Tau^2 = 0.55$; $Chi^2 = 57.60$, df = 1 ($p < 0.00001$); $I^2 = 98\%$							
Test for overall effect: Z = 0.50 (p = 0.62)							
Knowledge of harmful consequences of FGM/C							
Women							
Diop et al. 2004 (Senegal)	243	333	50	200	49.4	2.92	(2.28–3.74)
Ouoba et al. 2004 (Burkina Faso)	497	578	166	228	50.6	1.18	(1.08–1.29)
Total	740	911	216	428	100.0	1.85	(0.65–5.22)
Heterogeneity: $Tau^2 = 0.55$; $Chi^2 = 62.35$, df = 1 ($p < 0.00001$); $I^2 = 98\%$							
Test for overall effect: Z = 1.16 (p = 0.25)							
Men							
Diop et al. 2004 (Senegal)	54	82	42	198	48.1	3.10	(2.28–4.23)
Ouoba et al. 2004 (Burkina Faso)	394	448	137	229	51.9	1.47	(1.31–1.64)
Total	448	530	179	427	100.0	2.11	(1.00–4.42)
Heterogeneity: $Tau^2 = 0.27$; $Chi^2 = 20.27$, df = 1 ($p < 0.00001$); $I^2 = 95\%$							
Test for overall effect: Z = 1.97 (p = 0.05)							
Prevalence of FGM/C among girls aged 10 and younger							
Diop et al. 2004 (Senegal)	143	358	103	199	95.0	0.77	(0.64–0.93)
Ouoba et al. 2004 (Burkina Faso)	16	519	9	217	5.0	0.74	(0.33–1.66)
Total	159	877	112	416	100.0	0.77	(0.64–0.92)
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.01$, df = 1 ($p = 0.93$); $I^2 = 0\%$							
Test for overall effect: Z = 2.84 (p = 0.004)							

[a] Weight refers to the amount of contribution assigned each study to the pooled effect size (based on the inverse of the variance).

benefits of eliminating it" through education and outreach (WHO 2008: 19). Unfortunately, results from this one study evaluating the effectiveness of an intervention program for health personnel suggest that after the intervention, fewer wished to play a role in educating clients about the practice. A sense of advocacy among participating health personnel appeared low. This group could play a key role in halting the prevalence of FGM/C; thus, it is important to encourage advocacy in an effort to gain their active contribution.

Moreover, because medical personnel are caretakers of girls and women who have been subjected to FGM/C and have experienced complications, and because such personnel are figures of authority in many communities, their possession of extensive, accurate knowledge about the practice is essential. Unfortunately, Diop and colleagues' 1998 study showed that the knowledge level of health personnel remained suboptimal, despite training. Only 72 percent could name long-term complications associated with FGM/C. These findings, coupled with the misperception that FGM/C is safe if performed in a hygienic environment, demonstrate the importance of counteracting misconceptions and convictions about the practice among medical and healthcare personnel. Whereas FGM/C performed by medical personnel in hospitals and health clinics may reduce some short-term complications regularly seen when the procedure is performed by traditional practitioners, even when carried out by trained professionals, FGM/C is not necessarily less severe, nor are conditions necessarily sanitary. WHO (2008) offers no evidence that medicalization reduces obstetric or other long-term complications associated with FGM/C. According to Budiharsana and colleagues (2004), the involvement of medical professionals may have the unintended effect of conferring legitimacy upon FGM/C as a medically sound procedure, thereby contributing to the misconception that the practice is acceptable and medically safe. Diop and colleagues' intervention study (1998) has low methodological quality and is more than a decade old. Thus, FGM/C knowledge among today's health personnel in Mali may be greater, and they may be more willing to abandon the practice. Nevertheless, the findings are compelling and warrant the attention of future researchers.

According to the results from Mounir and colleagues (2003), the second type of included intervention shows that two sessions of reproductive health education, which address the dangers of FGM/C, improve female students' knowledge about the practice. This is encouraging, and the role of education within FGM/C prevention initiatives is indicated (see also Asekun-Olarinmoye and Amusan 2008).

Concerning the effectiveness of multifaceted community activities, the meta-analysis result for belief that FGM/C compromises the human rights of women

was not significant, and considerable inconsistency was found between the studies. Whereas a greater proportion of Ethiopian intervention participants than comparison group participants held this belief at endline, the reverse situation occurred among Somali participants. Given that the absence of rights is considered to be a fundamental force sustaining FGM/C (UNICEF 2005b; WHO 2008), empowerment and women's rights issues should be incorporated into prevention programs. A causal link between beliefs about women's rights and practicing FGM/C has not been fully explored, however, and the current results suggest that multicomponent community programs that adopt a women's rights framework fall short of altering individuals' beliefs about women's rights.

Basing the dismissal of the potential for change through a focus on rights on the present study's findings would be premature, however. For example, the Somali comparison group received some education, and thus is an imperfect assessment. Approaches to address human rights were culturally adapted to the two sites because of their different cultural contexts. Whereas the multicomponent intervention among Ethiopians and Somalis sought to present FGM/C as a women's and girls' rights issue, participants' actual exposure to such messages may have been low. For both sites, intervention fidelity was uncertain in general, and intervention exposure uncertain in particular. No data about exposure to women's rights messages were available for the Ethiopian site, and in the Somali site only 1 percent of respondents recalled the message that FGM/C compromises the human rights of women.

In contrast to the other interventions, two of the TOSTAN empowerment interventions included measures (via mothers' report) of whether girls had been cut. In the Senegal study, 12 percent fewer girls aged 10 and younger were cut in the intervention group, compared with the comparison group, whereas the difference in the Burkina Faso study was 1 percent at endline. Although the validity of the meta-analysis result was uncertain, the observed decrease in cutting among daughters of intervention participants is an encouraging finding and is the most relevant finding with respect to the effectiveness of FGM/C abandonment interventions.

One of the meta-analysis results from the empowerment interventions, men's knowledge of harmful consequences, favored the intervention, but high heterogeneity raised doubts about the validity of the result. Analyses were completed separately for female and male participants, and collapsing across gender resulted in both a nonsignificant result and an I^2 of 99 percent. Concerning inconsistency across studies, the TOSTAN replication study in Burkina Faso showed different results from the study in Senegal. This may be related to relevance, a point we address below.

Although the meta-analyses results show that the effectiveness of the empowerment interventions on community members' knowledge of harmful consequences is uncertain, study-level results were encouraging. Similar to the multifaceted community interventions, the empowerment interventions appeared to affect participants' awareness of FGM/C. The community-based programs appeared to affect beliefs about benefits of FGM/C, approval of the practice, regrets about having had a daughter cut, and intentions for the future regarding having daughters subjected to FGM/C. Collectively, results from the community-based interventions support the proposition that information dissemination may form a central component for successful change in FGM/C prevalence.

Implications

The low quality of the research designs of the studies evaluating FGM/C interventions points to the need for additional and stronger research. Eight controlled FGM/C abandonment studies, which can be considered first-generation program evaluations, show a paucity of reliable evidence regarding effectiveness of interventions to halt FGM/C. To stem the practice, future intervention studies should be developed in partnership with local communities (and with the particular categories of individuals and institutions most appropriate for the setting) and be situated within appropriate historical, cultural, and policy contexts. Programs to stem FGM/C should marshal local resources—for example, by drawing on the authority of key opinion leaders. Because the practice is strongly reinforced by social norms and belief systems, encouraging the larger community and the authorities who uphold social customs to question unhealthy norms is essential.

The reasons for the limited effectiveness of the eight included interventions are likely many, but imperfect relevance and implementation fidelity seem particularly probable. By relevance we refer to the degree of fit or match between an intervention and the target community's peculiarities, including demographic and psychological characteristics (McKenzie and Smeltzer 2001; Green and Kreuter 2005). Such fit is best accomplished through rigorous program planning (McKenzie and Smeltzer 2001), which includes intensive study of the particularities of the setting before an intervention is initiated. Lack of knowledge and understanding of the target groups' key features makes it difficult to design appropriate, culturally relevant programs. Further, relevance is essen-

tial for the sustainability of health promotion programs (McKenzie and Smeltzer 2001), and as reports show (WHO 2008), sustained action against FGM/C is vital to ensure long-lasting, positive results.

Pre-intervention assessment and planning varied greatly across the studies, from no mention of planning in one (Easton et al. 2002) to extensive pre-intervention research in two others (Chege et al. 2004 [Kenya and Ethiopia]). The extensive research was aimed at understanding the context of the practice, and the results were applied toward designing a context-relevant intervention. Conceivably, even with intervention planning, some of the inconsistent and nonsignificant findings documented in this systematic review were a result of interventions that were imperfect responses to the populations' needs. Specifically, some of the studies seemed to struggle with the fact that FGM/C abandonment efforts challenged local social customs or were ethically unacceptable to the local community, and the host organizations' ideological structure may have been a barrier to intervention implementation. These issues may be what prompted the occasional hostile responses, such as attacks against community members who voiced opposition to FGM/C, abductions for unwanted cutting, and threats against and direct attacks on staff members (see, for example, Chege et al. 2004; Diop et al. 2004). The six community-based prevention programs addressed the local community as setting, but also to varying extents as target, agent, and resource, placing them firmly within an ecological health-promotion perspective. Changing a practice that is deeply entrenched in a community's culture will likely be possible only when an enabling environment for behavior change has been created, which may be best achieved through community-linked programs. UNICEF (2005b) and WHO (2008) note that interventions to end FGM/C should be community-led and adapted to reflect regional, ethnic, and socioeconomic variations. A study assessing the degree to which specific FGM/C abandonment interventions are tailored to their communities is under way (Berg and Denison, forthcoming).

Concerning the second likely reason for the limited effectiveness of included interventions—weak implementation fidelity—some information was available for a few of the included studies. For example, in the empowerment program in Mali, only a "small group of actual participants [was] in the TOSTAN course" (Easton et al. 2002: 18), casting intervention exposure into doubt, and the curriculum was not conducted according to protocol. Although implementation fidelity is typically difficult to achieve and to measure in multicomponent health-promotion interventions (Armstrong et al. 2009), future evaluations should strive to ensure and document fidelity, because it may help explain an intervention's effectiveness or lack thereof (Bellg et al. 2004; Borrelli et al. 2005).

Whereas conducting complex interventions consisting of interconnected parts (as six of the included interventions did) is extremely complicated, the need exists not only for culturally responsive programs but also for interventions that deliver the best science through methodologically rigorous studies. First, similarity of prognostic factors at baseline is a basic prerequisite in effect studies, and is typically achieved by using adequate methods of randomization (Higgins and Green 2011), which was not the case in the studies included in this review. The assessment of the similarity of prognostic factors at baseline indicated that, in each of the community-based interventions included, the results may be biased in favor of the intervention groups, because of dissimilarities in education, prevalence of FGM/C, and religion. Also, effectiveness outcomes were in many instances dissimilar. Second, intention to perform FGM/C on daughters may be a good proxy for prevalence, but the UNICEF-initiated international agreement on standard indicators for situation analysis toward ending FGM/C states that prevalence is the most important indicator (UNICEF 2005b) and is the most clinically meaningful outcome for girls and women at risk for FGM/C. Estimates of prevalence are most reliable as an outcome when medical examinations, rather than self-reporting, are used. None of the studies in this systematic review included biological data, and in the two studies that included prevalence, cutting status of girls under the age of 11 was ascertained from the mothers' self-reports, with no physical examinations to verify the statements.

One strength of this systematic review was the comprehensive and systematic literature search and systematic process for identifying relevant publications. Another strength is that, in keeping with the recommendation of intervention-assessment researchers in general (Eccles et al. 2003) and FGM/C intervention-assessment researchers in particular, we limited our analyses to controlled studies—those that are best designed to isolate the effects of interventions and control for confounding variables. Unfortunately, the full text of two publications (Babalola and Adebajo 1985; Akweongo et al. 2007) were unobtainable. Further, the inclusion of the study by Akweongo and colleagues (reported in PRB 2007) might have affected the results regarding prevalence. This study is a field experiment investigating two community-based FGM/C abandonment strategies, education activities, and livelihood and development activities, and a combination of the two strategies compared to no intervention. The main outcome is prevalence of FGM/C over a five-year period.

Conclusion

This systematic review of the effectiveness of the best available evidence of FGM/C abandonment interventions included eight diverse studies from Africa. Although the studies were characterized by low methodological quality and low quality of documentation, requiring that we view their results cautiously, the results nevertheless point to possible advantageous developments as a result of the interventions. Thus, these studies, which can be considered first-generation evaluations of interventions to prevent FGM/C, offer reason to be optimistic that with sustained efforts, FGM/C may be ended within a few generations.

Note

1 The literature search strategy in Ovid MEDLINE® In-Process and Other Non-Indexed Citations and Ovid MEDLINE® 1950 to Present consisted of the following search terms: "Circumcision, Female/"; "((female$ or wom#n or girl$1) adj3 (mutilation$ or circumcis$ or cutting$)).tw."; " 'fgm/c'.tw."; "((removal$ or alteration$ or excision$) adj6 female genital$).tw."; "pharaonic circumcision$.tw."; "sunna.tw."; "(clitoridectom$ or clitorectom$).tw."; "(infibulat$ or reinfibulat$ or deinfibulat$).tw."; "or/1-8."

References

Akweongo, P.E., J. Jackson, P. Phillips, et al. 2007. "Reducing female genital cutting through community mobilization: The impact of a field experiment in Northern Ghana." Unpublished.

Armstrong, Rebecca, Elizabeth Waters, and Jodie Doyle. 2009. "Reviews in health promotion and public health," in Julian P.T. Higgins and Sally Green (eds.), *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.0.1*. West Sussex: Wiley-Blackwell, pp. 593–606.

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

Askew, Ian. 2005. "Methodological issues in measuring the impact of interventions against female genital cutting," *Culture, Health & Sexuality* 7: 463-477.

Babalola, S. and C. Adebajo. 1985. "Evaluation report of female circumcision eradication project in Nigeria." Presented at the annual meeting of the American Public Health Association, New York, 18 November 1996.

Babalola, Stella, Angela Brasington, Ada Agbasimalo, Anna Helland, Edith Nwanguma, and Nkechi Onah. 2006. "Impact of a communication programme on female genital cutting in eastern Nigeria," *Tropical Medicine and International Health* 11: 1594–1603.

Bellg, Albert J., Belinda Borrelli, Barbara Resnick, et al. 2004. "Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium," *Health Psychology* 23(5): 443–451.

Berg, Rigmor C. and Eva Denison. 2011. "Does female genital mutilation/cutting (FGM/C) 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.

———. Forthcoming. "A realist synthesis of controlled studies to understand the effectiveness of female genital mutilation/cutting (FGM/C)." *Campbell Systematic Reviews*.

Borrelli, Belinda, Deborah Sepinwall, Albert Bellg, et al. 2005. "A new treatment tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research," *Journal of Consulting and Clinical Psychology* 73(5): 852–860.

Budiharsana, Meiwita, Lila Amaliah, Budi Utomo, and Erwinia. 2004. "Female circumcision in Indonesia: Extent, implications and possible interventions to uphold women's health rights." Jakarta: Population Council.

Chege, Jane, Ian Askew, Susan Igras, and Jacinta K. Muteshi. 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, DC: Population Council, Frontiers in Reproductive Health.

Denison, Eva, Rigmor C. Berg., Simon Lewin, and Atle Fretheim. 2009. "Effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting." Report from Kunnskapssenteret nr 25. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.

Diop, Nafissatou J. and Ian Askew. 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, Nafissatou J., Modou M. Faye, Amadou Moreau, et al. 2004. "The TOSTAN Program. Evaluation of a community based education program in Senegal." New York: Population Council.

Diop, Nafissatou J., Amadou Moreau, and Hélène Benga. 2008. "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, Frontiers in Reproductive Health.

Diop, Nafissatou J., Fatoumata Traoré, Habibatou Diallo, et al. 1998. "Etude de l'efficacité de la formation du personnel sociosanitaire dans l'educacion des client(e)s sur l'excision et dans le traitemaent de ses complications au Mali." Bamako, Mali: Population Council.

Easton, Peter, Rebecca Miles, and Karen Monkman. 2002. "Final report on the evaluation of the TOSTAN/IEP village empowerment program pilot project in the Republic of Mali." Tallahassee, FL: Florida State University .

Eccles, M., J. Grimshaw, M. Campbell, and C. Ramsay. 2003. "Research designs for studies evaluating the effectiveness of change and improvement strategies," *Quality and Safety in Health Care* 12: 47–52.

Gillespie, Diane and Molly Melching. 2010. "The transformative power of democracy and human rights in nonformal education: The case of TOSTAN," *Adult Education Quarterly* 60(5): 477–498.

GRADE Working Group. 2004. "Grading quality of evidence and strength of recommendations," *BMJ* 328: 1490–1594.

Green, Lawrence W. and Marshall W. Kreuter. 2005. *Health Program Planning: An Educational and Ecological Approach*. Boston: McGraw Hill.

Guyatt, Gordon H., Andrew D. Oxman, Elie A. Akl, et al. 2011. "GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables," *Journal of Clinical Epidemiology* 64: 383–394.

Guyatt, Gordon H., Andrew D. Oxman, Regina Kunz, Gunn E. Vist, Yngve Falck-Ytter, and Holger J. Schünemann. 2008. "What is 'quality of evidence' and why is it important to clinicians?" *BMJ* 336(7651): 995–998.

Higgins, Julian P.T. and Sally Green (eds.). 2011. *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0.* The Cochrane Collaboration.

HRP. 2011. "An update on WHO's work on female genital mutilation (FGM). Progress Report." <http://whqlibdoc.who.int/hq/2011/WHO_RHR_11.18_eng.pdf>. Accessed 9 January 2012.

McKenzie, James F. and Jan L. Smeltzer. 2001. *Planning, Implementing and Evaluating Health Promotion Programs*. Boston: Allyn and Bacon.

Monkman, Karen, Rebecca Miles, and Peter Easton. 2007. "The transformatory potential of a village empowerment program: The TOSTAN replication in Mali," *Women's Studies International Forum* 30: 452–464.

Mounir, Gehan M., Nehad H. Mahdy, and Ibtsam M. Fatohy. 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.

Muteshi, J. and J. Sass. 2005. "Female genital mutilation in Africa: An analysis of current abandonment approaches." Nairobi: PATH.

Ouoba, Djingri, Zakari Congo, Nafissatou J. Diop et al. 2004. "Experience from a community based education program in Burkina Faso, the TOSTAN Program." ("Expérience d'un programme d'éducation à base communautaire au Burkina Faso, le Programme de TOSTAN.") Washington, DC: Population Council, Frontiers in Reproductive Health.

Population Reference Bureau (PRB). 2007. "Abandoning female genital mutilation/cutting: An in-depth look at promising practices." Washington, DC: PRB.

———. 2008. "Female genital mutilation/cutting: Data and trends." Washington, DC.

Sackett, David L., Sharon E. Straus, W. Scott Richardson, William Rosenberg, and R. Brian Haynes. 2000. *Evidence-Based Medicine: How to Practice and Teach EBM*. Edinburgh: Churchill Livingstone.

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

UNFPA. 2007. "A holistic approach to the abandonment of female genital mutilation/cutting." New York.

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

———. 2005b. "Changing a harmful social convention: Female genital mutilation/cutting," *Innocenti Digest*. Florence, Italy: UNICEF.

World Health Organization (WHO). 1997. "Female genital mutilation. A joint WHO/UNICEF/UNFPA statement." Geneva.

———. 1999. "Female genital mutilation. Programmes to date: What works and what doesn't? A review." Department of Women's Health, Health Systems, and Community Health. Geneva.

———. 2000. "A systematic review of the health complications of female genital mutilation including sequelae in childbirth." Department of Women's Health, Health Systems, and Community Health. Geneva.

———. 2008. "Eliminating female genital mutilation: An interagency statement." Geneva: WHO, in partnership with UNAIDS, UNDP, UNECA, UNESCO, UNFPA, UNHCHR, UNHCR, UNICEF, and UNIFEM.

———. 2010. "Global strategy to stop health-care providers from performing female genital mutilation." Geneva: WHO, in partnership with FIGO, ICN, MWIA, UNFPA, UNHCR, UNICEF, UNIFEM, WCPA, and WMA.

WHO Study Group. 2006. "Female genital mutilation and obstetric outcome: WHO collaborative prospective study in six African countries," *Lancet* 367: 1835–1841.

Yoder, P. Stanley and Shane Kahn. 2008. "Numbers of women circumcised in Africa: The production of a total," USAID, DHS Working Paper No. 39.

Acknowledgments

We are grateful to the Norwegian Knowledge Centre for Violence and Traumatic Stress Studies for providing financial support for this systematic review. We thank Owolabi Bjälkander, Liselott Dellenborg, Atle Fretheim, Brynjar Landmark, Jan Odgaard-Jensen, Sari Ormstaad, and Gunn Vist for valuable comments on earlier drafts.