Effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting

kunnskapssenteret
Norwegian Knowledge Centre for the Health Services

Background: In November 2008, the Norwegian Knowledge Centre for Violence and Traumatic Stress Studies commissioned the Norwegian Knowledge Centre for the Health Services to do a systematic review about the effectiveness of interventions to prevent female genital mutilation/cutting (FGM/C). The review would answer the question: What is the effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting compared to no or any other intervention? **Method:** We searched systematically for relevant literature in international scientific databases, in databases of international organisations that are engaged in projects concerning FGM/C, and in reference lists of relevant reviews and included studies. We selected studies according to prespecified criteria and appraised the methodological quality using checklists. We summarized the results using tables and calculated effect estimates in outcomes for which pre- and post scores for both intervention and comparison groups were reported. **Results:** We included and summarized results from six controlled before-after-studies. All studies were set in Africa and compared an

kunnskapssenteret

(continued from page one) intervention with no intervention. There was great variation in prevalence, ethnicity, religion, and education among these settings. All studies were judged to have weak methodological quality and the quality of the evidence was low. **Conclusion:** There is a paucity of high quality evidence regarding the effectiveness of interventions to prevent FGM/C and the evidence base is insufficient to draw solid conclusions. While first generation anti-FGM/C intervention studies are informative, there is an urgent need for additional studies. Such second generation studies should be randomized or at a minimum secure similar distribution of prognostic factors in the intervention and comparison groups; long term to ensure viability and reliable assessment of changes in prevalence; take into account regional, ethnic and socio demographic variation in the practice of FGM/C; focus on prevalence – assessed by physical examinations – behaviours, and intentions; and they should be cross disciplinary, if possible through international collaborative initiatives.

List of abbreviations

AR	Alternative rights.
FGM/C	Female Genital Mutilation/Cutting.
HIV	Human Immunodeficiency Virus.
NKVTS	Nasjonalt Kunnskapssenter om Vold og Traumatisk Stress/ The Norwegian Knowledge Centre for Violence and Traumatic Stress studies.
NOKC	Nasjonalt Kunnskapssenter for Helsetjenesten/ Norwegian Knowledge Centre for the Health Services.
PRB	Population Reference Bureau.
RCT	Randomized Controlled Trial.
UNFPA	United Nations Population Fund.
UNICEF	United Nations Children's Fund.
USAID	U.S. Agency for International Development.
WHO	World Health Organization.

Key messages

Effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting

Background: In November 2008, the Norwegian Knowledge Centre for Violence and Traumatic Stress Studies (NKVTS) commissioned the Norwegian Knowledge Centre for the Health Services (NOKC) to do a systematic review about the effectiveness of interventions to prevent female genital mutilation/cutting (FGM/C). The review would answer the question: What is the effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting compared to no or any other intervention?

Methods: We searched systematically for relevant literature in international scientific databases, in databases of international organisations that are engaged in projects concerning FGM/C, and in reference lists of relevant reviews and included studies. We selected studies according to pre-specified criteria and appraised the methodological quality using checklists. We summarized the results using tables and calculated effect estimates (adjusted absolute risk difference and risk ratio) in outcomes for which pre- and post scores for both intervention and comparison groups were reported.

Results: We included and summarized results from six controlled before-and-after studies. All studies were set in Africa and compared an intervention with no intervention (except one study which included an educational module). There was great variation in prevalence, ethnicity, religion, and education among these settings. All studies were judged to have weak methodological quality and the quality of the evidence was low. The effect estimates suggest that 1) training health personnel likely produced no effects in knowledge or beliefs/attitudes about FGM/C; 2) educating female students may possibly have led to a small increase in knowledge/awareness about FGM/C; 3) multifaceted community activities may possibly have increased the proportion of participants having favourable cognitions and intentions about FGM/C; 4) community empowerment through education may possibly have positively affected prevalence of FGM/C, participants' knowledge about the consequences of FGM/C, and regrets about having had daughter cut. However, the low quality of the body of evidence affects the interpretation of results and raises doubts about the validity of the findings.

Conclusion: There is a paucity of high quality evidence regarding the effectiveness of interventions to prevent FGM/C and the evidence base is insufficient to draw

solid conclusions. While first-generation anti-FGM/C intervention studies are informative, there is an urgent need for additional studies. Such second-generation studies should be randomized or at a minimum secure similar distribution of prognostic factors in the intervention and comparison groups; long-term to ensure viability and reliable assessment of changes in prevalence; take into account regional, ethnic and sociodemographic variation in the practice of FGM/C; focus on prevalence – assessed by physical examinations – behaviours, and intentions; and they should be cross-disciplinary, if possible through international collaborative initiatives.

Executive summary

Female genital mutilation/cutting (FGM/C) is a traditional practice that involves the partial or total removal of or other injury to the female genital organs for cultural or other non-therapeutic reasons. FGM/C is practised in more than 28 countries in Africa and in immigrant communities in a number of countries, including Australia, Canada, France, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The practice of FGM/C is rooted in religious, personal and societal beliefs within a frame of psycho-sexual and social reasons such as control of women's sexuality and family honour which is enforced by community mechanisms.

FGM/C is recognized as a harmful practice which violates human rights. It is prohibited by law in several African and Western countries. The 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, *infibulation*, 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. Type 4, *other*, involves all other harmful procedures to the female genitalia for non-medical purposes, for example: pricking, piercing, incising, scraping and cauterization. There is great variation in prevalence, reflecting ethnicity, tradition, and sociodemographicfactors. Countries with very high prevalence, over 70%, include Egypt, Ethiopia, Mali, and Somalia. FGM/C is associated with several health risks such as severe pain, bleeding, and shock, difficulty in passing urine and faeces, and infections. Caesarean section, blood loss, low birth weight and increased perinatal mortality are associated birth risks. Several psychological, social, and sexual consequences such as anxiety, depression, memory loss, loss of libido, and dyspareneuia are associated with FGM/C.

Efforts to abandon the practice of FGM/C in Africa have used several different approaches, including those based on human rights frameworks, a health risk approach, training health workers as change agents, and the use of comprehensive social development approaches. Although there are indications of the effectiveness of some anti-FGM/C interventions in achieving changes in knowledge, beliefs, attitudes, behaviours and practices related to FGM/C, systematic appraisal of the

evidence is lacking. Further, much research has used observational designs that make it difficult to draw causal inferences, thus hampering valid conclusions about the effects of these interventions.

We asked the following question: What is the effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting compared to no or any other intervention?

METHODS

In February 2009, we searched systematically for literature in the following scientific databases: African Index Medicus, Anthropology Plus, British Nursing Index and Archive, The Cochrane Library (CENTRAL, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects), EMBASE, EPOC, MEDLINE, PILOTS, POPLINE, PsychINFO, Social Services Abstracts, Sociological Abstracts, and WHOLIS. We also searched in databases of international organisations that are engaged in projects concerning FGM/C, and manually in reference lists of relevant reviews and studies included in this systematic review. We searched for studies that used the following study designs: systematic reviews, randomized controlled studies, and controlled before-and-after studies.

Two of the authors independently assessed studies for inclusion according to pre-specified criteria and the methodological quality of the studies using checklists. We summarized the results in text and tables and calculated effect estimates (adjusted absolute risk difference and risk ratio) in outcomes for which pre- and post scores for both intervention and comparison groups were reported.

RESULTS

We identified 3,667 publications and after having assessed titles, abstracts, and articles in full text we included six studies that fulfilled the inclusion criteria. All studies were controlled before-and-after studies that were carried out in African countries. We failed to obtain two potentially relevant publications, despite extensive retrieval efforts.

All the included studies were judged to have weak methodological quality and the quality of the evidence was low. Collectively, the studies involved a total of 6,803 participants at entry. All studies compared an intervention with no intervention (except one which included an education module). Each study was set in a different country in Africa: Burkina Faso, Egypt, Ethiopia / Kenya, Mali, Nigeria, and Senegal. There was great variation in prevalence, ethnicity, religion, and education among these settings. Two of the studies were directed at the individual level, and four at the community level. The first individually-based study consisted of educational activities delivered to health personnel in Mali, who learned about context and local rationale of FGM/C as well as the different types of cutting and its health complications. The other individually-based study took place in Egypt and

involved female university students, who received information about reproductive health aspects, including FGM/C. The multifaceted, community-based intervention in Kenya was delivered in a Somali refugee camp, and six village communities in Ethiopia received a nearly identical intervention, consisting of community meetings, theatre performances, video sessions, and mass media activities. In Nigeria, the multifaceted community activities, involving programmes such as multimedia and development of action plans to improve women's situation, was delivered at three community levels. The community empowerment intervention took place first in Senegal and then it was replicated in Burkina Faso. It consisted of educational sessions in human rights, problem solving, environmental hygiene, and women's health.

The most frequently reported outcomes were beliefs/attitudes, knowledge/awareness, and intentions concerning FGM/C. Less frequently reported outcomes were self-reported prevalence, behaviours such as talking to others about FGM/C, perceptions regarding spouse's disapproval of FGM/C, and regrets of having had daughter cut. The effect estimates suggest that 1) training health personnel likely produced no effects in knowledge or beliefs/attitudes about FGM/C; 2) educating female students may possibly have led to a small increase in knowledge/awareness about FGM/C; 3) multifaceted community activities may possibly have increased the proportion of participants having favourable cognitions and intentions about FGM/C; 4) community empowerment through education may possibly have positively affected prevalence of FGM/C, participants' knowledge about the consequences of FGM/C, and regrets about having had daughter cut. However, the low quality of the body of evidence affects the interpretation of results and raises doubts about the validity of the findings.

CONCLUSION

There is a paucity of high quality evidence regarding the effectiveness of interventions to reduce the prevalence of FGM/C and the evidence base is insufficient to draw solid conclusions. While first-generation anti-FGM/C intervention studies are informative, there is an urgent need for additional studies. Such second-generation studies should be randomized or at a minimum secure similar distribution of prognostic factors in the intervention and comparison groups; long-term to ensure viability and reliable assessment of changes in prevalence; take into account regional, ethnic and sociodemographic variation in the practice of FGM/C; focus on prevalence – assessed by physical examinations – behaviours, and intentions; and they should be cross-disciplinary, if possible through international collaborative initiatives.

The Norwegian Knowledge Centre for the Health Services (NOKC) summarizes and disseminates evidence concerning the effect of treatments, methods, and interventions in health services, in addition to monitoring health service quality. Our

goal is to support good decision making in order to provide patients in Norway with the best possible care. The Centre is organized under The Directorate for Health and Social Affairs, but is scientifically and professionally independent. The Centre has no authority to develop health policy or responsibility to implement policies.

Norwegian Knowledge Centre for the Health Services
PB 7004 St. Olavs plass
N-0130 Oslo, Norway
Telephone: +47 23 25 50 00
E-mail: post@kunnskapssenteret.no
Full report (pdf): www.kunnskapssenteret.no

Hovedfunn

Bakgrunn: I november 2008 gav Nasjonalt kunnskapssenter om vold og traumatisk stress (NKVTS) Nasjonalt kunnskapssenter for helsetjenesten (NOKC) i oppdrag å utføre en systematisk kunnskapsoppsummering av effektiviteten av tiltak for å redusere forekomsten av kjønnslemlestelse. Oppsummeringen skulle besvare spørsmålet: Hva er effekten av tiltak for å redusere forekomsten av kjønnslemlestelse sammenlignet med ingen eller annet tiltak?

Metode: Vi søkte systematisk etter relevant litteratur i internasjonale databaser, i databaser til internasjonale organisasjoner som driver prosjekter om kjønnslemlestelse, og i referanselistene til kunnskapsoppsummeringer og de inkluderte studiene. Vi valgte ut studier som oppfylte våre predefinerte inklusjonskriterier og brukte sjekklister for å vurdere den metodiske kvaliteten til studiene. Vi oppsummerte resultatene i tabeller og beregnet effektestimat (justert absolutt risikoforskjell og relativ risiko) for utfall som hadde før- og etterskår for både intervensjons- og kontrollgruppene.

Resultater: Vi inkluderte og oppsummerte resultatene for seks kontrollerte før- og-etter studier. Alle var gjennomført i afrikanske land og sammenlignet et tiltak med ingen tiltak (med unntak av én studie som inkluderte en utdanningsmodul). Det var stor variasjon i forekomst, etnisitet, religion og utdanning blant disse stedene. Alle studiene ble vurdert å ha lav kvalitet etter to forskjellige kvalitetsvurderings-systemer, og kvaliteten på evidensen var lav. Effektestimatene tyder på at 1) trening av helsepersonell muligens ikke gav noen forbedringer av kunnskaper om eller holdninger til kjønnslemlestelse 2) undervisning av kvinnelige studenter kan muligens ha ført til litt økt kunnskap om kjønnslemlestelse 3) sammensatte samfunnsaktiviteter kan muligens ha økt andelen deltakere som hadde holdninger og intensjoner om å oppgi kjønnslemlestelse 4) samfunn "empowerment" gjennom undervisning kan muligens ha minsket andelen mødre som rapporterte omskjæring av datter og økt andelen deltakere som hadde kunnskaper om følger av kjønnslemlestelse, og som følte anger over å ha omskåret datter. Imidlertid påvirker den lave kvaliteten på kunnskapsgrunnlaget tolkningen av resultatene og reiser tvil om gyldigheten av funnene.

Konklusjon: Det mangler dokumentasjon av høy kvalitet når det gjelder effekten av tiltak for å redusere forekomsten av kjønnslemlestelse, og kunnskapsgrunnlaget er utilstrekkelig for å dra sikre konklusjoner. Førstegenerasjonsstudiene som er oppsummert her er informative, men det er et presserende behov for ytterligere

studier. Slike andregenerasjonsstudier bør være randomiserte eller minimum sikre lignende fordeling av prognostiske faktorer i tiltaks- og sammenligningsgruppene; være langsiktige for å sikre levedyktighet og reliabel måling av forekomst; ta i betraktning regional, etnisk, sosial og demografisk variasjon i utøvelsen av kjønnslemlestelse; fokusere på forekomst – bedømt ved fysisk undersøkelse – atferd og intensjoner; og tverrfaglige, gjerne gjennom internasjonale samarbeid på flere felter.

Sammendrag

BAKGRUNN

Kjønnslemlestelse er en tradisjonell praksis som innebærer att hele eller deler av de kvinnelige kjønnsorganene fjernes eller skades av ikke-terapeutiske grunner. Kjønnslemlestelse praktiseres i mer enn 28 land i Afrika og i immigrantsamfunn i vestlige land som Australia, Canada, Frankrike, Norge, New Zealand, Storbritannia, Sveits, Sverige, og USA. Kjønnslemlestelse er grunnet i religiøse, personlige, og sosiale forestillinger, der kontroll av kvinnelig seksualitet og vern av familiens ære er viktige normer. Kjønnslemlestelse er anerkjent som en skadelig praksis som krenker menneskerettigheter og er uttrykkelig forbudt i mange afrikanske og vestlige land. Kjønnslemlestelse klassifiseres i fire kategorier: Type 1, *klitoridektomi*, delvis eller total fjerning av klitoris og/eller forhuden; Type 2, *eksisjon*, delvis eller total fjerning av klitoris og de små kjønnsleppene; Type 3: *infibulasjon*, delvis eller fullstendig fjerning av ytre kjønnslepper og gjensying slik at urinrørsåpningen dekkes av et hudseil og skjedeåpningen forsnevres til ca. 1 cm i diameter; Type 4: *alle andre former*, inklusive prikking og stikking i klitoris, strekking av klitoris og/eller kjønnslepper, etsning eller brenning av klitoris og omliggende vev, og innføring av etsende stoffer i skjeden for å minske hulrommets størrelse.

Forekomsten av kjønnslemlestelse i Afrika varierer sterkt, avhengig av kulturelle, etniske, sosiale og demografiske forhold. Land med meget høy forekomst, over 70 %, er for eksempel Egypt, Etiopia, Mali og Somalia. Kjønnslemlestelse settes i sammenheng med flere helsefarer som for eksempel alvorlige smerter, blødning, og sjokk, vanskeligheter med avføring, og infeksjoner. Keisersnitt, blodtap, lav fødselsvekt og økt perinatal dødelighet er risikofaktorer i sammenheng med fødsel. Flere psykologiske, sosiale og seksuelle konsekvenser som engstelse, depresjon og hukommelsestap, tap av libido og smerter ved samleie er også dokumentert i sammenheng med kjønnslemlestelse.

Mange prosjekter med hensikt å redusere forekomsten av kjønnslemlestelse er blitt gjennomført. Tiltakene har hatt ulike tilnærminger, for eksempel undervisning om menneskerettigheter, informasjon om helsefarer, opplæring av helsepersonell for å fremme forandring, og omfattende tiltak for sosial utvikling på samfunnsnivå. Selv om det finnes indikasjoner på at noen tiltak er effektive for å oppnå forandring av kunnskaper, holdninger og praksis når det gjelder kjønnslemlestelse, så mangler det en systematisk vurdering av kunnskapsgrunnlaget. I tillegg er det meste av

forskningen gjennomført med metoder som gjør det vanskelig å trekke sikre konklusjoner om årsak og virkning.

Vi stilte følgende spørsmål: Hva er effekten av tiltak for å redusere forekomsten av kjønnslemlestelse sammenlignet med ingen eller annet tiltak?

METODE

I februar 2009 søkte vi systematisk etter litteratur i følgende 13 internasjonale databaser: African Index Medicus, Anthropology Plus, British Nursing Index and Archive, The Cochrane Library (CENTRAL, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects), EMBASE, EPOC, MEDLINE, PILOTS, POPLINE, PsychINFO, Social Services Abstracts, Sociological Abstracts, og WHOLIS. Vi søkte også i databaser til internasjonale organisasjoner som driver prosjekter om kjønnslemlestelse, og manuelt i referanselister til oversiktsartikler og inkluderte studier. Vi søkte etter litteratur med følgende studiedesign: systematiske kunnskapsoppsummeringer, randomiserte kontrollerte studier og kontrollerte før-og-etter studier.

To prosjektmedarbeidere gjorde uavhengige vurderinger av studier for inklusjon, og av de inkluderte studienes metodiske kvalitet. Til dette brukte vi et inklusjonsskjema og sjekklister. Vi oppsummerte resultatene i tekst og tabeller og i kalkulerte effektestimater der det var mulig.

RESULTAT

Vi identifiserte 3667 publikasjoner og etter å ha vurdert titler, sammendrag og artikler i fulltekst fant vi seks studier som oppfylte inklusjonskriteriene. Alle studiene var kontrollerte før-og-etter studier og gjennomført i afrikanske land. To muligens relevante publikasjoner fikk vi ikke tak i, på tross av omfattende forsøk. Alle inkluderte studier ble vurdert å ha lav metodisk kvalitet.

Alle studiene sammenlignet et tiltak med ingen tiltak, unntatt en studie der et tiltak ble sammenlignet med et annet tiltak i en del av studien. Til sammen deltok 6803 personer i de seks studiene, som var gjennomførte i syv ulike afrikanske land: Burkina Faso, Egypt, Etiopia / Kenya, Mali, Nigeria, og Senegal. Det var stor variasjon i forekomst av kjønnslemlestelse, etnisitet, religion og utdanning mellom de ulike studieområdene. To av studiene var rettet mot individnivå og fire var rettet mot samfunnsnivå. Det første individrettede tiltaket inkluderte helsepersonell i Mali som fikk opplæring i lokale årsaker til kjønnslemlestelse, samt de ulike typene av kjønnslemlestelse og behandling av fysiske skader som følge av kjønnslemlestelse. Det andre individrettede tiltaket inkluderte kvinnelige studenter i Egypt som mottok undervisning om reproduksjon fra et helseperspektiv, som inkluderte kjønnslemlestelse. Tiltaket som kan kalles sammensatte samfunnsaktiviteter ble gitt til Somaliere i Kenya, seks landsbyer i Etiopia, og en delstat i Nigeria. Intervensjonstiltakene varierte fra holdningskampanjer, teateroppsetninger, video, fellesmøter, til bruk av massemedier. Samfunn "empowerment" gjennom

undervisning er et tiltak som inneholdt undervisning i menneskerettigheter, problemløsning, miljøhygiene, og kvinners helse. Tiltaket ble først gitt til flere samfunn i Senegal og dernest i Burkina Faso.

De hyppigst rapporterte utfallsmålene var holdninger, kunnskaper/bevissthet, og intensjoner angående kjønnslemlestelse. Mindre hyppig rapporterte utfallsmål var forekomst, atferd som å snakke sammen med andre om kjønnslemlestelse, oppfatning om ektefelles uvilje mot kjønnslemlestelse, og anger over å ha latt datter bli omskåret. Effektestimatene antyder at 1) opplæring av helsepersonal muligens ikke hadde noen effekt på kunnskaper eller holdninger om kjønnslemlestelse; 2) undervisning av kvinnelige studenter kan muligens ha ført til en liten økning av kunnskaper om kjønnslemlestelse; 3) sammensatte samfunnsaktiviteter kan muligens ha økt andelen deltakere som hadde holdninger og intensjoner om å oppgi kjønnslemlestelse; 4) samfunn "empowerment" gjennom undervisning kan muligens ha minsket andelen mødre som rapporterte omskjæring av datter og økt andelen deltakere som hadde kunnskaper om følger av kjønnslemlestelse, og som følte anger over å ha omskåret datter. Imidlertid påvirker den lave kvaliteten på kunnskapsgrunnlaget tolkningen av resultatene og reiser tvil om gyldigheten av funnene.

KONKLUSJON

Kunnskapsgrunnlaget i denne oppsummeringen er av lav kvalitet og ikke tilstrekkelig for å trekke sikre konklusjoner om effekten av tiltak for å redusere forekomsten av kjønnslemlestelse. Det er et presserende behov for studier som er randomiserte eller som et minimum sikrer lignende fordeling av prognostiske faktorer i intervensjons- og sammenligningsgruppene; bruker tilstrekkelig lang oppfølgingsperiode til å kunne bedømme forekomst på en sikker måte; tar i betraktning regional, etnisk, sosial og demografisk variasjon med hensyn til kjønnslemlestelse; fokuserer på forekomst – målt ved fysisk undersøkelse – atferd, og intensjoner; er tverrfaglig, gjerne gjennom internasjonale samarbeid på flere felter.

Table of contents

Preface

In November 2008, the Norwegian Knowledge Centre for Violence and Traumatic Stress Studies (NKVTS) contacted the Norwegian Knowledge Centre for the Health Services (NOKC) with a request for assistance in establishing a competence centre on the topic of FGM/C. Specifically, the NKVTS commissioned the NOKC to conduct a series of systematic reviews about female genital mutilation/cutting (FGM/C), one of which assessed the effectiveness of interventions designed to reduce the prevalence of FGM/C.

The project team was composed of:

- Project coordinator, researcher, Eva Denison, NOKC
- Researcher, Rigmor C Berg, NOKC
- Senior researcher, Simon Lewin, NOKC
- Research librarian, Sari Ormstad, NOKC
- Statistician Jan Odgaard-Jensen, NOKC
- Research director, Atle Fretheim, NOKC

We are grateful for peer review by two internal and two external reviewers:

- Brynjar Landmark, researcher, NOKC
- Gunn Vist, researcher, NOKC
- Liselott Dellenborg, Ph.D., University of Gothenburg (Sweden)
- Owolabi Bjälkander, FGM/C expert, Socialstyrelsen (Sweden)

The aim of this report is to support well-informed decisions in health promotion that lead to reductions in FGM/C prevalence and improved quality of services related to FGM/C. The evidence should be considered together with other relevant issues, such as clinical experience and patient preferences.

Anne Karin Lindahl   Atle Fretheim   Eva Denison
Director   *Research Director*   *Researcher*

Objective

The present systematic review summarizes extant literature about interventions designed to reduce the prevalence of FGM/C and evaluates the effectiveness of such interventions.

NKVTS commissioned the NOKC to conduct a systematic review to support the organization's health promotion work concerning FGM/C among women subjected to and at risk for the practice in Norway, but the systematic review is of relevance in all countries where FGM/C may occur. The overall aim of the systematic review is to support well-informed decisions in health promotion that lead to the reduction of FGM/C prevalence and improved quality of services related to FGM/C.

The main research question was:

* What is the effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting compared to no or any other intervention?

Background

The terminology used for the cutting of female genital organs varies. It has been referred to as "female circumcision", "female genital mutilation", "female genital cutting" and "female genital mutilation/cutting" (1). We adopt the official terminology used by UNICEF and UNFPA "female genital mutilation/cutting" (FGM/C) throughout this report. FGM/C is a traditional practice that involves the partial or total removal or other injury to the female genital organs for cultural or other non-therapeutic reasons (2). FGM/C is practised in more than 28 countries in Africa, usually on girls under the age of 15 years, and in some countries in the Middle East and Asia (3;4). FGM/C is also practised by immigrant communities in a number of countries, including Australia, Canada, France, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States (4). As Western governments become more aware of FGM/C among the immigrant communities, legislation has come into effect and European Union institutions and Member States have taken steps towards ending FGM/C (5).

The practice of FGM/C is rooted in religious, personal and societal beliefs within a frame of psycho-sexual and social reasons such as control of women's sexuality and family honour which is enforced by community mechanisms (6). While reasons for the practice vary across cultural groups, social reasons may include FGM/C as an initiation act for girls into womanhood, as an act of social integration and for the maintenance of social cohesion. 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. Religious reasons rest on the belief that it is a religious requirement. Hygienic and aesthetic reasons for FGM/C include beliefs that the female genitalia are dirty and unsightly, and health reasons include beliefs that FGM/C enhances fertility and child survival. FGM/C may also be an important source of income for circumcisers (7).

FGM/C is recognized as a harmful practice which violates the human rights – civil, cultural, economic, political and social – of girls and women (1). Further, FGM/C is a manifestation of gender inequality and discrimination "related to the historical subjugation and suppression on women" (8). By extension, it is hypothesized that changing beliefs about women's rights is a key to its abandonment (1;9). There are laws that explicitly prohibit the practice of FGM/C in many African countries, e.g. in Burkina Faso, Egypt, Eritrea, Ethiopia, Kenya, and Senegal. In other countries, existing general provisions of criminal codes can be applied to FGM/C, e.g. in Mali and North Sudan, while there are no laws against FGM/C in Somalia and Nigeria

(10). There are laws prohibiting FGM in several Western countries including Australia, Canada, New Zealand, USA and at least 13 countries in Western Europe (4), among these Denmark, Norway, and Sweden (11). However, the implementation of anti-FGM laws and their impact on eliminating the practice has so far not been extensively studied (12).

According to the WHO and other leading health organizations, there are no known health benefits to FGM/C (1). Male circumcision on the other hand, provides some protection from certain infections, such as human immunodeficiency virus (HIV). Evidence from randomized, controlled clinical trials in South Africa, Kenya, and Uganda confirms that male circumcision can be efficacious for men in reducing their risk of HIV acquisition through sex with women (13-15). The implications of African trials on circumcision for HIV prevention programmes in the other parts of the world are less clear, due to the differences surrounding the HIV epidemics in Africa and other regions, such as North America and Europe (16). A 2007 WHO and Joint United Nations Programme on HIV/AIDS technical consultation on male circumcision resulted in a summary document which provides conclusions and recommendations relating to policy and programmeming on male circumcision and HIV prevention. The report explicitly states that, based on limited available data, promoting circumcision for HIV-positive men is not recommended (17). In sum, one main distinction between male circumcision and FGM/C is the potential medical benefits of male circumcision (18;19). Another important distinction between FGM/C and male circumcision is the degree of harm to the genital organs. Whereas in male circumcision the foreskin is cut off from the tip of the penis without damaging the organ itself, the degree of cutting in FGM/C is anatomically much more extensive (20). Toubia (20) writes that the male equivalent of clitoridectomy (in which all or part of the clitoris is removed) would be the amputation of most of the penis. The male equivalent of infibulation (which involves cutting the labia minora and/or the labia majora and the closing off of the tissue around the vagina) would be the removal of the penis; its root of soft tissue, and part of the scrotal skin.

CLASSIFICATIONS

The 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, *infibulation*, 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 (2).

PREVALENCE

Recent figures for African countries show a prevalence of FGM/C of more than 70% in Burkina Faso, Djibouti, Egypt, Eritrea, Ethiopia, Guinea, Mali, Mauretania, Northern Sudan, and Somalia (21). However, national rates do not reveal the magnitude of FGM/C among certain ethnic groups; there is great variation in prevalence between and within countries, reflecting ethnicity and tradition. Prevalence figures that are presented by geographical area show differing variation between areas with the highest and lowest prevalence within countries. UNICEF (22) has proposed that countries be categorized in three groups according to prevalence rates, which vary greatly between and within countries. In Table 1, we present prevalence data from countries which will be discussed in this systematic review. Group 1 consists of countries where prevalence rates are high (80% or more). In Group 2 countries, the prevalence rates are at intermediate levels (25-79%) and usually only certain ethnic groups practice FGM/C, at varying levels. Group 3 countries have low prevalence rates (1-24%) and only some ethnic groups practice FGM/C.

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	100	0.8	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 (22). 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 (21).

CONSEQUENCES

FGM/C is associated with numerous health risks and consequences, as well as psychological, social, and sexual consequences (1).

Physical consequences

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 (1). In general, the consequences are similar for FGM/C Type I, II, and III, but they tend to be more severe and more prevalent the more extensive the procedure (1).

A systematic review of the health complications of FGM/C (23) 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 28 obstetric centres in several African countries (24) concluded that women with FGM/C are 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.

Research has found that medicalization of FGM/C has increased dramatically in recent years in some countries, such as Mali (22). The medical profession, led by the WHO and the World Medical Association, has condemned the medicalization of FGM/C (22). Although FGM/C that is performed by medical personnel in health clinics may reduce some short-term complications regularly seen when it is performed by traditional practitioners, medicalized FGM/C is not necessarily less severe or conditions sanitary, and there is no evidence that medicalization reduces obstetric or other long-term complications associated with FGM/C (1).

Psychological consequences

For many girls and women, undergoing FGM/C is a traumatic experience that leaves a lasting psychological mark and 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. Disordered eating habits include loss of appetite, weight loss or excessive weight gain, and disordered sleeping habits include sleeplessness and recurring nightmares (4). There are also reports of post-traumatic stress disorder, anxiety, depression, and memory loss associated with FGM/C (1).

Social consequences

FGM/C is a deeply entrenched social convention among some ethnic groups and as such carries consequences both when it is and when it is not practised. When girls and families conform to the practice they acquire social status and respect. For girls, undergoing FGM/C promotes honour and her full acceptance in the community, as well as imparts a sense of pride and of coming of age (22). 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 (25). For families, fulfilling the cultural expectation that girls should be cut assigns status and community membership. Conversely, failure to conform leads 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 (22).

Sexual consequences

Sexual consequences of FGM/C were recently summarized in a non-systematic literature review (26), which concluded that the available evidence does not support the notion that FGM/C automatically precludes sexual activity or the enjoyment of

sexual relations. However, adverse sexual consequences of FGM/C, such as loss of libido and dyspareunia are documented (27).

A systematic review of the psychosocial (psychological, social and sexual) consequences of FGM/C is underway at the NOKC.

INTERVENTIONS TO REDUCE THE PREVALENCE OF FGM/C

Efforts to abandon the practice of FGM/C in Africa have used several different approaches which, in turn, have had implication for interventions. These approaches include those based on human rights frameworks, legal mechanisms, a health risk approach, training health workers as change agents, training and converting circumcisers, an alternative rites approach, the positive deviance approach, and the use of comprehensive social development approaches. Interventions based on these approaches have targeted stakeholders at individual, interpersonal, community and national levels (28).

Recently, the Population Reference Bureau (PRB) carried out an extensive survey of current intervention projects taking place in African countries (29). By applying criteria of 'best practice' defined by the United Nations, the authors arrived at a small number of projects (n=3) that 1) were evaluated by scientific methods, 2) had a demonstrable and tangible impact on improving people's quality of life, 3) were the result of effective partnerships between the public, private, and civic sectors of society, and 4) were socially, culturally, economically, and environmentally sustainable. The projects had in common a comprehensive community approach, building on the involvement of human resources in the community and the communication of messages relating to FGM/C through multiple channels. In total, the PRB identified 92 projects, 27 of which were evaluated, mostly by observational designs. Only four of the 27 evaluated projects used a controlled before-and-after design, and about a dozen of the evaluations used cross-sectional or pre-post intervention questionnaires or interviews without a control group. Results of such before-and-after studies (without a control group) were generally positive. For example, an assessment by Chege, Askew, and Liku (30) of the alternative rites (AR) approach for encouraging abandonment of FGM/C in Kenya suggested there were positive differences from the pretest to the posttest among community members who participated in the programme. At endline, more girls knew about reproductive health issues and expressed gender egalitarian attitudes; more AR families stated there were no benefits to FGM/C and had knowledge about health and social/psychological problems associated with FGM/C; and fewer girls and boys in AR households intended to circumcise their daughters. Similarly, a 2008 impact evaluation of a health education intervention in the Shao community in Nigeria found that at endline, there was an increase in the proportion of intervention participants who wanted the practice to be stopped and the proportion who intended to cut their daughters in the future decreased significantly among both Christians and Muslims (31).

Although such studies indicate the effectiveness of some anti-FGM/C interventions in achieving desired outcomes such as changes in knowledge, beliefs, attitudes, behaviours and practices related to FGM/C (28;29), systematic appraisal of the evidence is lacking. Further, much research has used observational designs that make it difficult to draw causal inferences, thus hampering valid conclusions about the effects of these interventions (32).

Method

We conducted a systematic review of the effectiveness of anti-FGM/C interventions in accordance with the NOKC handbook for summarizing evidence (33) and most of the guidelines in the Cochrane Handbook for Systematic Reviews of Interventions (34).

LITERATURE SEARCH

The database search strategy was designed and executed February 4-9 2009, by research librarian Sari Ormstad at the NOKC. The search in Anthropology Plus was executed February 19 2009, by Hege Oswald at NKVTS. We searched systematically for relevant literature in the following 13 international databases:

- African Index Medicus
- Anthropology Plus
- British Nursing Index and Archive
- The Cochrane Library (CENTRAL, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects)
- EMBASE
- EPOC
- MEDLINE
- PILOTS
- POPLINE
- PsycINFO
- Social Services Abstracts
- Sociological Abstracts
- WHOLIS

The search strategy is detailed in Appendix 1. We supplemented the database search with searches in databases of six international organizations that are engaged in projects regarding FGM/C (see Appendix 1), as well as searching literature lists of relevant reviews and included studies. Unpublished reports, abstracts, brief and preliminary reports were considered for inclusion on the same basis as published reports.

INCLUSION CRITERIA

We accepted several study designs:

1. systematic reviews,
2. randomized controlled trials,
3. controlled before-and-after studies.

Population: Girls/young women at risk of FGM/C, other members of communities practicing FGM/C, communities practicing FGM/C.

Interventions: Any intervention or programme designed to reduce the prevalence of FGM/C, including but not limited to 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, and comprehensive social development.

Outcomes: Rates of FGM/C, public declarations to abandon FGM/C, proportion of sample in favour of abandoning FGM/C, behaviours related to FGM/C, awareness of rights, knowledge of harmful consequences of FGM/C, and beliefs and attitudes related to FGM/C.

Languages: We included all languages. When considered likely to meet the inclusion criteria, studies were translated to English.

EXCLUSION CRITERIA

We excluded all studies not meeting our pre-specified inclusion criteria. Specifically, we excluded non-systematic reviews, empirical evaluation studies without a comparison group, empirical evaluation studies without both pre- and postmeasures, as well as cross-sectional studies.

We excluded studies where the participants were not part of a community/ethnic group in which FGM/C is practiced. We also excluded programmes that were not designed to change cognitive or behavioural factors related to FGM/C and studies that did not assess cognitions, behaviours or prevalence of FGM/C.

SELECTION OF STUDIES

Two authors (Denison and Berg) independently read all titles/and or abstracts resulting from the search process and eliminated any obviously irrelevant studies. We obtained full copies of the remaining potentially relevant studies (one study could not be obtained in full text and one study is forthcoming). The same pair of authors, acting independently, classified these as clearly relevant, that is, met all inclusion criteria and therefore to be included, or clearly irrelevant and therefore to be excluded. Pre-designed inclusion/exclusion forms were used for each screening level. All titles and abstracts were also read by researchers at the NKVTS who have knowledge in the subject area. The results of the NKVTS screening were used as validation of relevance of the studies selected by the NOKC researchers.

It was not necessary to contact the authors of any studies to aid the decision process. Decisions were made on inclusion criteria outlined, i.e. types of studies, types of

participants, intervention and outcome measures used. Differences in opinion in the screening process were few and were resolved through consensus. Studies formally considered but excluded are listed in Appendix 2 and reasons for exclusion are provided.

DATA EXTRACTION AND ANALYSIS

Two authors independently extracted data from the published sources using a pre-designed data recording form. Where differences in data extracted occurred, this was resolved through discussion.

With respect to quality of included studies, we used the McMaster University, Effective Public Health Practice Project, Quality Assessment Tool for Quantitative Studies (35) to assess the methodological quality. The tool includes the following sections: A) selection bias, B) study design, C) confounders, D) blinding, E) data collection methods, F) withdrawals and drop-outs, G) intervention integrity, and H) analyses. Sections A – F are each given a rating of strong, moderate or weak according to pre-specified criteria. Finally, a global rating of strong, moderate or weak is given, according to the following criteria: strong (four strong ratings with no weak ratings); moderate (less than four strong ratings and one weak rating); weak (two or more weak ratings). A final decision of strong, moderate or weak methodological quality was agreed upon by Denison and Berg after discussing whether there was a discrepancy between the two reviewers with respect to the component (A – F) ratings.

To assess the quality of the evidence, we used the Cochrane Collaboration's tool for assessing risk of bias, as described in chapter 8 of the Cochrane Handbook for Systematic Reviews of Interventions, version 5.0.1 (34). Denison and Berg independently assessed the following five specific domains at study level: similarity of prognostic factors at baseline, blinding of assessor, incomplete outcome data, selective reporting, and other potential threats to validity, such as study design. We judged assessment at study level rather than outcome level as appropriate because data were collected in the same manner, by interview/questionnaire in all studies, and all outcomes can be considered 'cognitions'. Also prevalence data regarding young girls were self-reported by mothers and not based on physical examination.

For non-randomized study designs, the degree of equivalency between comparison groups is uncertain and must be assessed. We chose education, ethnicity, prevalence of FGM/C, and religion as prognostic factors that should be similar in the intervention and comparison groups, based on suggestions by several agencies (1;10;22). For example, based on multivariate analyses of demographic survey data, UNICEF (22) suggests that educational attainment, a woman's own cutting status, and ethnicity have the greatest influence in determining support or opposition to the practice. Because cross-sectional and independent samples appear to have been used in four of the included studies (36-39), similarity of prognostic factors needed to be assessed at both the baseline and the follow-up assessments in these studies.

We discussed and agreed about the adequacy of each risk of bias study domain by assigning a judgement of 'Yes' indicating low risk of bias, 'No' indicating high risk of bias, and 'Unclear' indicating unclear or unknown risk of bias. Criteria indicated by the Cochrane handbook and adapted to the health promotion field were used to make these judgements.

Further, we planned to apply the instrument Grading of Recommendations Assessment, Development and Evaluation (GRADE) (34) to assess the extent to which we could be confident that estimates of effect were correct.

With respect to analyses, we present dichotomous data for the outcomes listed in the inclusion criteria in the results tables when pre- and post scores for both intervention and comparison groups are reported by study authors, allowing comparison. We estimated effects of interventions in two ways. One, 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 was deemed to be large was a judgement whereby we also took baseline difference into account. Two, we estimated effect by the relative risk (RR) and 95% confidence interval (95%CI) based on post-intervention data. We present continuous data with mean difference and 95%CI.

We also decided, *a priori*, to perform meta-analyses to estimate effect. We decided to use Mantel-Haenszel random effects meta-analyses because it was assumed that the studies would estimate different, but related, intervention effects. Further, we used RevMan 5, the latest version of the Cochrane Collaboration's meta-analysis software (34).

Results

DESCRIPTION OF INCLUDED LITERATURE

Results of the search

The electronic search resulted in 3,667 individual publications and the manual search in one relevant publication (Figure 1). After removal of duplicates, Denison and Berg screened the publications by reviewing titles and abstracts. We eliminated obviously irrelevant publications based on titles and where available, abstracts.

We excluded 3,649 publications, leaving 18 potentially relevant publications, two of which could not be obtained in full text: One (40) could not be obtained despite extensive library retrieval efforts, and one (41) is forthcoming. We read the full text for 16 publications. We excluded ten publications (characteristics of excluded studies are presented in Appendix 2) and included six studies presented in nine publications.

Figure 1: Flowchart of the reviewing process

4,432 publications identified	
African Index Medicus	12
Anthropology Plus	200
British Nursing Index	135
CENTRAL/Cochrane	3
EMBASE	570
EPOC Register	0
MEDLINE	1100
PILOTS	14
POPLINE	1566
PsycINFO	401
Social Services Abstracts	40
Sociological Abstracts	325
WHOLIS	65
Manual search	1

765 duplicates removed

3,667 titles/abstracts read

3,649 irrelevant publications excluded

- 1 relevant report has not been obtained in full text
- 1 relevant report is forthcoming

16 publications read in full text

7 publications excluded due to:
- no control group (4)
- qualitative data only (1)
- publication did not report a study (2)

6 studies included
(9 publications)

Included studies

Six primary studies were included in this review:

- Diop (1998). Study of the effectiveness of training Malian social and health agents in female genital cutting issues and in educating their clients. (42;43)

- Mounir (2003). Impact of health education about reproductive health on knowledge and attitude of female Alexandria university students. (44)

- Chege (2004). Testing the effectiveness of integrating community-based approaches for encouraging abandonment of female genital cutting into CARE's reproductive health programmes in Ethiopia and Kenya. (36)

- Diop (2004). The TOSTAN programme evaluation of a community-based education programme in Senegal. (37;45)

- Ouoba (2004). Experience from a community-based education programme in Burkina Faso. (38;46)

- Babalola (2006). Impact of a communication programme on female genital cutting in eastern Nigeria. (39)

Two of the studies were published in peer-reviewed journals, *Tropical Medicine and International Health* (39) and *The Journal of the Egyptian Public Health Association* (44). The other four studies were published as reports to their funding agencies (36-38;42). These four studies were all funded by the U. S. Agency for International Development (USAID).

All the included studies employed a controlled before-and-after study design, involving a total of 6,803 participants at entry from seven different African countries (Table 2).

Table 2: Included studies (N=6)

Author, year	Population	Intervention	Comparison	Outcomes
Diop, 1998 (42)	N=108 male and female health personnel, Mali	Training clinic staff about FGM/C, supervising trained clinic staff (8 sites)	No intervention (6 sites)	Knowledge; Beliefs/Attitudes; Experiences
Mounir, 2003 (44)	N=682 female university students, Egypt	Two educational sessions of 60 minutes about reproductive health, including FGM/C (Ezbet Saad hostel)	No intervention (El-Shatby hostel)	Knowledge; Beliefs/Attitudes
Chege, 2004 (36)	N=2,259 male and female community members, Ethiopia and Kenya	Education through behaviour change communication activities, and advocacy (Kenya 1 refugee camp; Ethiopia 6 villages)	Kenya: education (1 refugee camp) Ethiopia: no intervention (6 villages)	Knowledge/Awareness; Beliefs/Attitudes; Intention
Diop, 2004 (37)	N=1,332 male and female community members, Senegal	Education in hygiene, problem solving, women's health, human rights (20 villages)	No intervention (20 villages)	Prevalence; Knowledge/Awareness; Beliefs/Attitudes; Intention
Ouoba, 2004 (38)	N=1,465 male and female community members, Burkina Faso	Education in hygiene, problem solving, women's health, human rights (23 villages)	No intervention (23 villages)	Prevalence; Behaviours; Knowledge/Awareness; Beliefs/Attitudes; Intention
Babalola, 2006 (39)	N=957 male and female community members, Nigeria	Community mobilization; advocacy; mass media activities (Enugu State)	No intervention (Ebonyi State)	Behaviours; Beliefs/Attitudes; Intentions

All of the studies were based in Africa, specifically in the band of countries stretching from Senegal in West Africa to Ethiopia on the East Coast, where prevalence of FGM/C is estimated to be highest. Each study was set in a different country: Burkina Faso, Egypt, Ethiopia / Kenya, Mali, Nigeria, and Senegal (Figure 2).

Figure 2: Map of Africa showing countries where studies took place
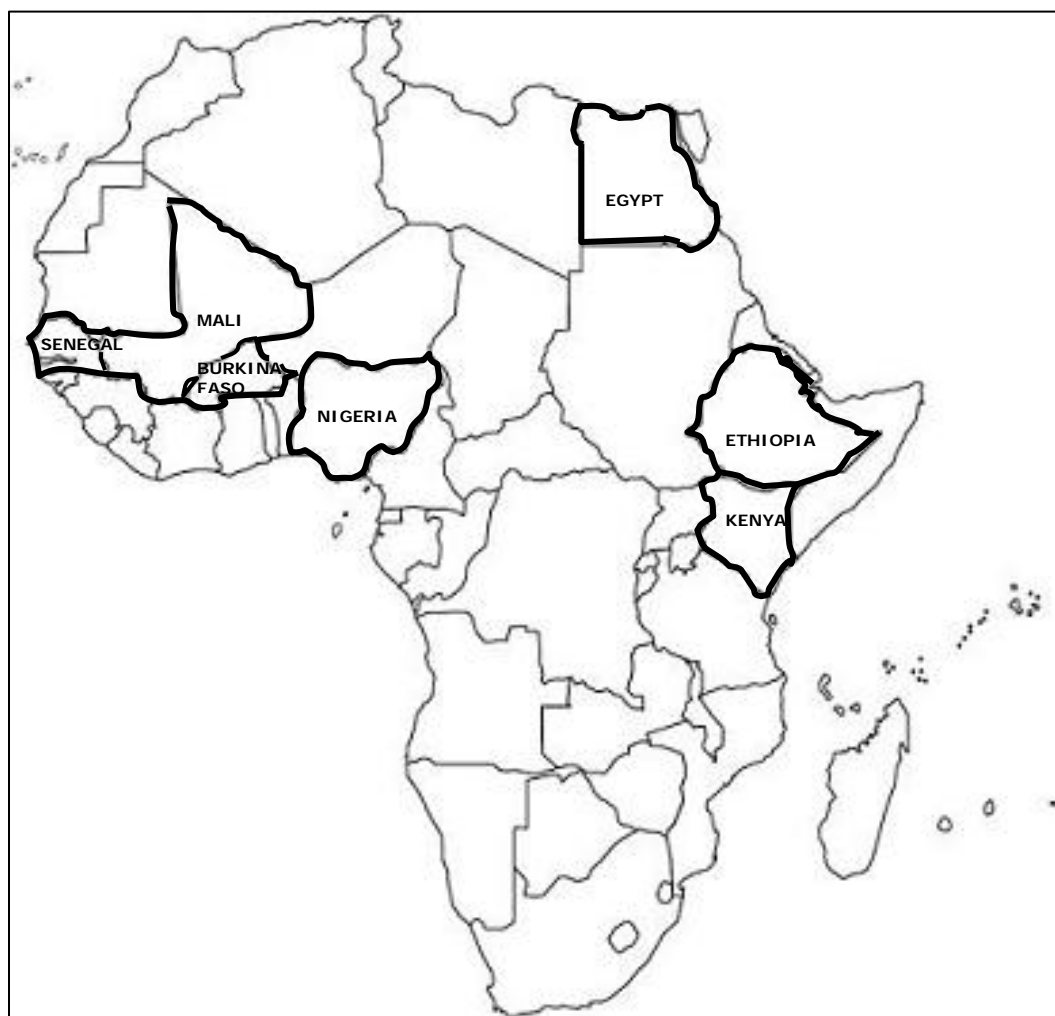
Population

All study participants lived in communities in Africa where the documented prevalence of FGM/C is high and a large proportion of women risk FGM/C (Table 3). The study participants were spread across seven different African countries: Burkina Faso, Egypt, Ethiopia, Kenya, Mali, Nigeria, and Senegal.

The sample sizes ranged from 108 to 2,259 participants. The four community-based studies (36-39) had the highest number of participants and included both males and females. In Kenya, Somali refugees received the intervention. The refugee camp, Ifo, was located in Dadaab in the north-eastern province, close to the Somali border. At the time of the intervention, Ifo was inhabited by approximately 44,700 Somali refugees. Six village communities in Awash, Ethiopia, consisting of approximately 18,000 people, received a nearly identical intervention. The empowerment intervention in Senegal (37) was delivered in 20 villages in the Region of Kolda in 2001, and replicated in 23 villages in the Bazega/Zoundweogo province in Burkina Faso (38). One study (44) targeted only female students, and one study (42) did not state how many of the health providers who participated in the study were males and how many were females. The mean age of the study participants across all

studies was about 32. Most of the participants held a low educational level and illiteracy was common.

The religious background of the participants varied. In the Senegal study (37) as well as the study set in Ethiopia and Kenya (36) almost all of the participants were Muslim, about half of the Burkina Faso study participants were Muslim (38), while the majority of the participants from Nigeria considered themselves Protestant (39). Also the ethnic backgrounds of the participants varied, and included groups such as the Mossi in Burkina Faso (38), who hold collectivism in high regard and are hierarchically and patrilineally organized, with the family and land as central cultural elements; Sunna Muslim Somalis in Kenyan refugee camps (36), who feel great loyalty to their clan; and the Ethiopian Afar (36) who are pastoral people.

In Burkina Faso, FGM/C is practiced by most ethnic groups, but it is most prevalent in the central Mossi plateau (over 80% prevalence rate) where the study took place. Type II is the most common type, followed by type I. Infibulation is not practiced (47). In Egypt, prevalence rates are consistently high across the country, at about 97% (22). The prevalence of FGM/C among the Afar in Ethiopia is 91% among women over age 16 and among the Somali in Kenya all women of reproductive age have undergone various types of FGM/C (36). The study in Mali was conducted in areas and among ethnic groups with high FGM/C prevalence rates: 92% of the health centre clients had been cut. Among these, FGM/C type II was the most frequent (73%), followed by type I (21%), and type III (6%) (42). In Nigeria, virtually 100% of women in the two study sites had experienced FGM/C, but according to the study authors there was considerable variation in FGM/C prevalence by local government area, especially in Enugu (39). Lastly, FGM/C is estimated as approximately 88% among women in the Kolda region in Senegal (37).

Table 3: Description of the population in included studies

Author, year	Population
Diop,1998	Obstetricians, gynaecologists and family planning providers (N=108, doctors, certified nurses, nurse aides, mid-wives, traditional birth attendants, health technicians), median age 38, 96% Muslim, 33% polygamous households, majority ethnic groups was Bambara (34%), from the districts of Bamako and Bla in Mali. 92% prevalence of FGM/C.
Mounir, 2003	Female (N=682) Alexandria University (Alexandria, Egypt) students, living in the university hostels, mean age 19, 47% from low social class families, 32% had mothers who were illiterate or who had only primary education. 87% prevalence of FGM/C.
Chege, 2004	Male (n=1,113) and female (n=1126) participants. Ethiopia: Afar people, 82% below primary level of education, mean age 26, primarily Muslim. 91% prevalence of FGM/C. Kenya: Somali refugees, 61% below primary level of education, mean age 31, primarily Muslim, 74% unemployed. 100% prevalence of FGM/C.
Diop, 2004	Male (n=557) and female (n=775) community members, mean age 35, 78% no education, 60% illiterate, 99% Muslim, 46% polygamous household, mainly Pulaar and Mandingo ethnic groups, from Kolda region in Senegal. ~88% prevalence of FGM/C.
Ouoba, 2004	Male (n=1,047) and female (n=718) community members, mean age 36, 90% no education, 73% illiterate, 45% Muslim, 50% polygamous household, Mossi ethnic group, from province of Zoundwégo and Bazèga in Burkina Faso. >80% prevalence of FGM/C.
Babalola, 2006	Male (n=426) and female (n=531) community members, mean age 34, 66% less than secondary education, 58% Protestant , 55% very religious, from Enugy State and Ebonyi State in eastern Nigeria. ~100% prevalence of FGM/C.

Intervention

The interventions consisted of a variety of activities, including trainings, educational sessions, advocacy, empowerment, and mass media (Table 4).

Two interventions (42;44) consisted of educational activities delivered at the individual level, to health personnel and students, respectively. The interventional education content to health personnel was provided in three group sessions, and 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. An entire day was devoted to the use of visual aids. Role playing was used to stimulate counselling sessions." (42). The study authors do not state the exact duration of the training sessions. The students received information about reproductive health aspects through two sessions of health talks, group discussion, role play, and use of educational aids (44).

Four interventions were delivered at the community level (36-39). Chege and colleagues' study (36) consisted of a community education and advocacy intervention, including activities such as community meetings, theatre performances, video sessions, and mass media activities. The multifaceted intervention began with the introduction of expanded (in Kenya) and new (in Ethiopia) FGM/C abandonment activities in 2001 and 2002. The two communities received a nearly identical intervention. The multifaceted programme in Enugu State in Nigeria (39) was delivered at three community levels: hamlet ('village') level, local government area level – which targeted influential decision makers – and Enugu state level. At state level, the multimedia programme 'Ndukaku', which is the Igbo word for "health is better than wealth", dominated, but other activities, such as development of action plans to improve women's situation, also took place.

The empowerment intervention in Senegal (37) took place in 2001 and 2002 and consisted of four educational modules covering hygiene, problem solving, women's health, and human rights. The focus was on enabling the participants to analyze their own situation and find the best solution for themselves. The intervention was replicated in Burkina Faso (38).

The intervention duration, i.e. time period during which the intervention groups and the comparison groups received different treatments, varied from about two weeks at the individual level (44) to an average of about 18 months at the community level (36-39).

Given the extensiveness of the community interventions, several hundred individuals with varying backgrounds – such as educators, communication experts, youth leaders, and journalists – delivered the intervention activities. The individual based interventions were generally provided by educators directly involved in the research project.

Table 4: Description of the interventions in included studies

Author, year	Base	Content	Duration & delivery
Diop,1998	Training and supervision of health care workers	3 training sessions (presentation, visual aids, role playing) consisting of recall of female anatomy and FGM/C, its context and local rationale, its prevalence, its health complications and treatment, different types of cutting; supervision of incorporation of FGM/C conversations	2 mo duration, at 8 health care sites, delivered by trainers and clinical personnel from organisations directing the research
Mounir, 2003	Health education programme for students	2 educational sessions (health talk, group discussion, role play, use of educational aids, handouts) about reproductive health: importance of premarital counselling, dangers of FGM/C, antenatal care, family planning, STIs, breast feeding	Two sessions of 60 min duration, at university, delivered by one of the researchers
Chege, 2004	Community education and advocacy	Community-level educational outreach activities using Behaviour Communication Change approaches and community-level advocacy (educational events, community meetings, theatre groups' performances, evening video sessions, mass media activities, support of advocacy activities)	Kenya: 21 mo duration, in refugee camp. Delivered by project staff, Ministry of Health officials, local community and youth leaders. Ethiopia: 17 mo duration. Delivered by project staff and volunteers, members of camp committees, local community and youth leaders
Diop, 2004	Village empowerment through educational programme	Educational sessions in human rights, problem solving, environmental hygiene, women's health	~12 mo duration? 3 classes per week of 120 min. Sessions delivered by facilitators from an NGO
Ouoba, 2004	Village empowerment through educational programme	Educational sessions in human rights, problem solving, environmental hygiene, women's health	~24 mo duration? Sessions delivered by facilitators from an NGO
Babalola, 2006	Community mobilization, advocacy, mass media	Hamlet: community meetings, organized community groups which designed action plans. LGA: visited with traditional leaders, religious leaders, local government officials, school authorities, women's groups, and discussed FGM/C at tribal meetings and town forums. State: newspaper columns, radio call-in shows, public forums	~12 mo duration? Activities delivered by Women Action Research Organization, National Association of Women Journalists

Legend: Hamlet= 'village'; LGA= Local government area; min= minutes; mo= months; NGO= Non-governmental organization; STIs= Sexually Transmitted Infections; ?= duration of intervention cannot be precisely determined from the publication.

Intervention fidelity and reach

Neither Diop and colleagues (42) nor Mounir and colleagues (44) provided programme fidelity details. Neither of the two studies reported any difficulties with respect to programme delivery. Loss to follow-up was 4% (42) and after the baseline assessment 8.5% of the students declined further participation (44).

In the study taking place in Kenya (36), the intervention did not include extensive collaborations with religious leaders, as planned. The authors do not explain why this occurred. They also state that they failed to effectively implement the advocacy strategy. Further, the approaches to address human rights and gender differences differed between the Kenyan and Ethiopian sites because of differences in cultural contexts. During the intervention period, staff experienced hostility from some community members, especially in the Kenyan site. To assess exposure to FGM/C abandonment messages, study participants were asked whether and which messages they could recall. Seventy one percent of Ethiopian study participants and 59% of Kenyan participants (59% of experimental respondents and 47% of comparison

respondents) said that they had been exposed to anti-FGM/C messages. The most commonly remembered message among participants in Kenya was that infibulation was harmful to health. Data about types of messages heard by Ethiopian participants were collected, but the authors write that "data were not available in usable format" (36).

In the Senegal empowerment programme (37), only about half of those who had expressed interest in the programme enrolled. Reasons for failing to participate were varied and included lack of financial compensation, pregnancy, and lack of permission from husband. The programme was originally developed to empower women, but the organizers decided to include men as well because male community members expressed discontent with not being included. Community members expressed some hostility against staff, which included an attack on a supervisor. With respect to programme participation, 69% of participating women and 57% of men attended all four educational modules, with the remainder completing part of the programme. Among those who attended only part of the programme, human rights and hygiene were the most popular modules.

Similarly, in the replication study in Burkina Faso (38), not all of those who had expressed interest in the programme enrolled, but others came to. Overall, 63% of participating women and 60% of men attended all four educational modules, with the remainder completing part of the programme. Reasons for failing to participate were varied and included lack of time, illness, and pregnancy. Attendance varied from over 70% attendance in the women's health and hygiene modules, 25% in the human rights module and 18% in the problem-solving module.

No implementation information was available for the Nigerian study (39). Programme exposure was measured through recall of messages. More than one third (36%) of participants were not exposed to any programme components. Among those who recalled at least one programme component, the majority had been exposed through the radio and about one third had participated in a community event.

Comparison

With the exception of one study, only one category of comparison was used in the six included controlled before-and-after studies: No intervention (Table 5).

All studies included a comparison group which was similar to the intervention group, but that received no intervention activities. However, in the Kenya study (36), the comparison group, Hagadera refugee camp, did receive educational activities. The authors do not provide an explanation for this decision, but make it clear that intervention contamination occurred because refugees can move freely into and out of the camps, although not all people moved between camps.

Table 5: Description of the comparisons in included studies

Author, year	Comparison
Diop,1998	Health care personnel at 6 sites, who received no intervention

Mounir, 2003	Female university students at the El-Shatby hostel, who received no intervention
Chege, 2004	Ethiopia: Community members in 6 villages in Amibara (pop. 38,000), who received no intervention Kenya: Hagadera refugee camp in Dadaab (pop. 44,200), who received education (but not advocacy activities)
Diop, 2004	Community members in 20 villages (similar to the intervention villages), who received no intervention
Ouoba, 2004	Community members in 23 villages, who received no intervention
Babalola, 2006	Community members in Ebonyi State (similar to the intervention state), who received no intervention

Outcomes

Two studies assessed prevalence of FGM/C (37;38) by asking the female respondents whether their daughters under 10 years of age had been subjected to FGM/C. The same two studies included a public declaration to abandon FGM/C as an outcome (37;38). Two studies measured behaviours (38;39); whether the community members discussed FGM/C with others, and whether the respondents discouraged others from performing FGM/C. Five of the studies (36-38;42;44) measured knowledge about the practice. All of the studies measured various beliefs and attitudes related to FGM/C, such as the belief that it is necessary to practise FGM/C (Table 6), and four studies (36-39) measured intentions to have daughter cut.

None of the studies collected biological data. All information was self reported regarding knowledge/awareness, beliefs/attitudes, intentions, behaviours and prevalence and collected from the participants through face-to-face structured interviews or paper-and-pencil questionnaires. In two studies (37;38) circumcision status of girls under the age of 10 was ascertained from statements of mothers, and no physical examinations were conducted to verify the statements.

Table 6: Description of the outcomes in included studies

Author, year	Outcomes
Diop, 1998	**Behaviours**: practice of FGM/C **Knowledge**: types of complications, immediate and late complications **Beliefs/Attitudes**: belief that FGM/C practised in a hygienic environment is safe for health; wished to play a role in educating the population who visited their health facilities about FGM/C; belief of impact on health of mother and child; belief that traditions can change; attitude towards use of term FGM/C; belief about uncut girls; belief that men prefer to marry cut women; belief that excision guarantees virginity; belief that FGM/C can cause infertility; belief that removal of clitoral hood represents no health risk; belief that FGM/C in medical environment must be prohibited **Other outcomes**: personal experience of FGM/C in practice; skill caring for FGM/C The time interval between end of intervention and post-intervention data collection is not reported, but the final data collection is said to have been delayed.
Mounir, 2003	**Knowledge**: dangers of FGM/C **Beliefs/Attitudes**: belief that traditions are the main causes for FGM/C; belief that FGM/C protects the girls from deviations; belief that side effects of FGM/C are serious Post-intervention data collection was done after one month in both groups.
Chege, 2004	**Knowledge/Awareness**: of harmful social, psychosexual, and health effects of FGM/C; awareness of human rights **Beliefs/Attitudes**: support for abandonment of FGM/C; perceived value of continuing FGM/C; perceived benefits of FGM/C; opinion of gender equality, women's role in community **Intentions**: intention not to cut The time interval between end of intervention and post-intervention data collection is not reported.
Diop, 2004	**Prevalence**: self-reported prevalence of FGM/C **Public declaration to abandon FGM/C** **Knowledge/Awareness**: awareness of harmful consequences of FGM/C **Beliefs/Attitudes**: belief it is necessary to practise FGM/C; belief FGM/C is supported by religion; approval of

	FGM/C; perceived partner approval of FGM/C; belief FGM/C is against women's rights; regrets having had daughter cut; preferral of cut woman; perceived benefits from cutting **Intentions**: intention to cut daughter; intend to encourage someone not to perform FGM/C; intention to support women who wish discontinuation of FGM/C The time interval between end of intervention and post-intervention data collection is not reported
Ouoba, 2004	**Prevalence**: self-reported prevalence of FGM/C **Public declaration to abandon FGM/C** **Behaviours**: discusses FGM/C with others **Knowledge/Awareness**: awareness of harmful consequences of FGM/C; knowledge of law against FGM/C **Beliefs/Attitudes**: belief it is necessary to practise FGM/C; approval of FGM/C; perceived partner approval of FGM/C; belief FGM/C is against women's rights; regrets having had daughter cut **Intentions**: intention to cut daughter Post-intervention data collection was done at the end of the programme.
Babalola, 2006	**Behaviours**: encouragement of others not to perform FGM/C **Beliefs/Attitudes**: belief there are benefits to FGM/C; personal approval of FGM/C; perceived self-efficacy to resist pressure from spouse to perform FGM/C on daughters; belief that most men and women in the community favour discontinuation of FGM/C **Intentions**: intention not to perform FGM/C on daughters The time interval between end of intervention and post-intervention data collection is not reported.

QUALITY ASSESSMENT

Study quality

We arrived upon a final decision of 'weak' study quality for all six included studies, based on the A – F component ratings of the McMaster University Quality Assessment Tool for Quantitative Studies. Detailed results of the quality assessment are presented in Appendix 3: Quality assessment, Table A2.

Quality of the evidence

The risk of bias assessment comprised five domains: similarity of prognostic factors at baseline, blinding of outcome assessor, incomplete outcome data, selective reporting, and other potential threats to validity. The results showed that there was 'high' or 'unclear' risk of bias in one or more of the assessed domains in all studies (see Appendix 3 Table A3).

There were potentially important differences between the intervention and comparison groups in three studies regarding similarity of prognostic factors at baseline (36;38;39). The level of education significantly different in two studies (36;39), and in both studies participants in the intervention groups had a higher level of education. We were not able to identify differences in ethnicity in any of the studies, mainly due to insufficient information. The prevalence of FGM/C differed significantly in one study (39), with higher prevalence in the comparison group than the intervention group. Religious affiliation differed significantly in two studies (38;39). In the study by Ouoba and colleagues (38), a majority of participants in the intervention group was Muslim while the majority in the comparison group was Catholic.

With respect to blinding of outcome assessor, there was insufficient information in all studies to judge whether outcome assessors were blinded to the participants' group affiliation.

There may be a risk of bias regarding incomplete outcome data. The four studies that evaluated community-based interventions apparently used cross-sectional and independent samples at baseline and follow-up. Exposure to the intervention was measured in all of the four community-based studies (36-39) and varied between 63% and 100%.

With respect to selective reporting, two studies failed to report comparisons between the intervention and control groups regarding one of their stated objectives (42;44) and were judged as not free of selective reporting. In Diop and colleagues' study (42), results regarding how likely they were to oppose the practice of FGM/C were not reported separately for the intervention and comparison groups. Likewise, Mounir and colleagues (44) did not report between-groups comparisons regarding attitudes to FGM/C.

The last risk of bias domain assessed was other potential threats to validity. Because the studies were not randomized we did not assess risk of bias concerning sequence generation and allocation concealment. These dimensions were recorded as bias related to the specific study design under the heading 'Other potential threats to validity'. Aside from the lack of randomization in all studies, the design was contaminated by population movements into and out of the intervention and comparison areas in one study (36).

We judged that there was 'high' or 'unclear' risk of bias for one or more of the assessed domains in all six included studies (Figure 3). Detailed results of the risk of bias assessment are presented in Appendix 3: Quality assessment, Table A3.

Figure 3: Results of the risk of bias assessment

Legend: Judgements of risk of bias was as follows: + = 'Yes' low risk of bias , - = 'No' high risk of bias , ? = 'Unclear' unknown risk of bias.

We planned to assess the quality of the evidence through GRADE. However, all the studies were dissimilar, non-randomized studies, thus we could not combine the outcomes, and there were important baseline differences between the groups in all studies. Taken together with the weak methodological quality of all studies and 'high' or 'unclear' risk of bias, the evidence was deemed to be of low quality. Therefore, we considered it not resource wise to conduct a formal GRADE assessment for the outcomes of included studies.

RESULTS OF TRAINING HEALTH PERSONNEL

One study with 108 participants, conducted in Mali, evaluated effects of training health personnel (42).

Table 7 shows the results for each outcome for which we could calculate effect estimates. The groups were judged to be different at baseline in three out of four outcomes, and there was inconsistency between adjusted absolute risk difference and relative risk in these outcomes. The 95%CI for each outcome indicated no difference in relative risk. The two groups were reasonably similar at baseline with respect to the outcome variable 'Wished to play a role in educating the population

who visited their health facilities about FGM/C', but the effect estimates show that there was no difference between the groups for this outcome.

Table 7: Study outcomes and effect estimates, in Diop (1998)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Knowledge								
Could name any type of FGM/C	76	95	19	47	81	34	-15	1.12 (0.97, 1.30)
Could name at least three types of long-term FGM/C complications	50	72	22	61	73	12	10	0.99 (0.79, 1.26)
Beliefs/Attitudes								
Believed that FGM/C pose no health risks if done in a hygienic environment	29	44	15	9	29	20	-5	1.54 (0.91, 2.60)
Wished to play a role in educating the population who visited their health facilities about FGM/C	93	91	-2	89	86	-3	-1	1.07 (0.93, 1.23)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI).

Prevalence was not measured in this study. Data regarding behaviours and intentions regarding FGM/C were not presented in such a way that calculation of effect estimates was possible.

The results do not provide convincing evidence that training of health personnel in Mali is likely to have an effect on knowledge or belief/attitudes regarding FGM/C.

RESULTS OF EDUCATING FEMALE STUDENTS

One study with 684 participants evaluated the effectiveness of health education about reproductive health on knowledge and attitudes of female university students in Alexandria, Egypt (44).

Table 8 shows that the mean knowledge score (maximum mean score = 3 points) increased by 0.47 points in the intervention group and by 0.02 points in the comparison group.

Table 8: Study outcomes and effect estimates, in Mounir (2003)

Outcome	Intervention		Comparison		Mean difference	95%CI
	Pre	Post	Pre	Post		
Mean knowledge score	1.00	1.47	0.70	0.72	0.75	0.63, 0.87

Note: Pre-, and post scores are in points and reproduced from the study publication. We calculated mean difference and 95% confidence interval (CI).

Prevalence, behaviours, and intentions regarding FGM/C were not measured in this study. Data regarding beliefs/attitudes were not presented in such a way that calculation of effect estimates was possible.

The results show that two educational sessions about reproductive health may to a small extent increase female university students' knowledge about FGM/C.

RESULTS OF MULTIFACETED COMMUNITY ACTIVITIES

Two studies with a total of 3,216 participants, carried out in Ethiopia/Kenya (36) and Nigeria (39) evaluated the effectiveness of multifaceted community activities.

Table 9 shows the results for each outcome in the Ethiopia study for which we could calculate effect estimates. The groups were judged to be similar at baseline in two outcomes ('Had knowledge of harmful consequences of FGM/C' and 'Believed that FGM/C compromised the human rights of women'). The effect estimates for these two outcomes indicated the intervention may have had a positive effect in the intervention group.

Table 9: Study outcomes and effect estimates for Ethiopia, in Chege (2004)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Knowledge/Awareness								
Had knowledge of harmful consequences of FGM/C [$]	33	87	54	36	71	35	19	1.37 (1.26, 1.49)
Beliefs/Attitudes								
Believed that FGM/C compromised the human rights of women [$]	8	42	34	10	19	9	25	2.21 (1.75, 2.79)
Supported the abandonment of FGM/C in their community [£]	22	54	32	15	25	10	22	2.16 (1.78, 2.62)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [$]	8	34	26	14	13	-1	27	2.62 (1.96, 3,49)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups.

Table 10 shows the results for each outcome in the Kenya study for which we could calculate effect estimates. The intervention and comparison groups were judged to be different at baseline for three out of four outcomes. There was inconsistency between the effect estimates for these three variables, and the 95%CIs showed the effects failed to reach significance. For one outcome, 'Believed that FGM/C compromised the human rights of women', the baseline difference between the groups was small and the effect estimates were consistent, indicating an effect in favour of the comparison group.

Table 10: Study outcomes and effect estimates for Kenya, in Chege (2004)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Knowledge/Awareness								
Had knowledge of harmful consequences of FGM/C [$]	57	91	34	71	89	18	16	1.02 (0.99, 1.06)
Beliefs/Attitudes								
Believed that FGM/C compromised the human rights of women [£]	25	31	6	27	40	13	-7	0.77 (0.67, 0.89)

Supported the abandonment of FGM/C in their community [#]	23	23	0	11	19	9	-9	1.21 (0.99, 1.48)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [#]	14	17	3	10	18	8	-5	0.94 (0.75, 1.17)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups, #= significantly greater change in comparison group.

Table 11 shows the results for women in Nigeria for each outcome for which we could calculate effect estimates. The intervention and comparison groups were judged to differ at baseline in all outcomes; an exception possibly being 'Did not intend to perform FGM/C on their daughters'. For this outcome, the effect estimates indicated an effect in favour of the intervention group.

Table 11: Study outcomes and effect estimates for women, in Babalola (2006)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Behaviours								
Had encouraged someone not to perform FGM/C on their daughters [$]	16	24	8	11	9	-2	10	2.68 (1.76, 4.08)
Beliefs/Attitudes								
Did not believe there were benefits to FGM/C[1] [$]	58	75	17	66	72	6	11	1.04 (0.95, 1.15)
Disapproved of FGM/C[1] [$]	63	88	15	71	73	2	13	1.21 (1.11, 1.13)
Believed that most men and women in the community favoured discontinuation of FGM/C [£]	36	49	13	21	14	-7	20	3.50 (2.58, 4.76)
Perceived they had self-efficacy to resist pressure from their spouse to perform FGM/C on their daughters [$]	57	72	15	40	42	2	13	1.71 (1.47, 1.99)
Intentions								
Did not intend to perform FGM/C on their daughters [$]	59	76	17	64	67	3	14	1.13 (1.02, 1.26)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups. 1= item reversed by us.

Table 12 shows the results for men in Nigeria for each outcome for which we could calculate effect estimates. The groups were judged to be similar at baseline in three out of six outcomes. In two of these ('Did not believe there were benefits to FGM/C' and 'Believed that most men and women in the community favoured discontinuation of FGM/C'), the effect estimates indicated an effect in favour of the intervention group.

Table 12: Study outcomes and effect estimates for men, in Babalola (2006)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Behaviours								
Had encouraged someone not to perform FGM/C on their daughters ˣ	12	14	2	10	12	2	0	1.19 (0.71, 2.01)
Beliefs/Attitudes								
Did not believe there were benefits to FGM/C[1] £	53	76	23	58	65	7	16	1.17 (1.02, 1.33)
Disapproved of FGM/C[1] $	67	77	10	30	30	0	10	2.57 (2.06, 3.20)
Believed that most men and women in the community favoured discontinuation of FGM/C $	24	35	11	25	20	-5	16	1.76 (1.25, 2.47)
Perceived they had self-efficacy to resist pressure from their spouse to perform FGM/C on their daughters $	72	88	16	51	40	-11	27	2.20 (1.85, 2.61)
Intentions								
Did not intend to perform FGM/C on their daughters $	53	73	20	64	66	2	18	1.11 (0.97, 1.27)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups, x= no significant change. 1= item reversed by us.

Prevalence was not measured in these studies. Behaviours were not measured by Chege and colleagues (36) and knowledge/awareness was not measured by Babalola and colleagues (39).

The results showed that multifaceted community activities may:
- increase knowledge/awareness about consequences of FGM/C
- increase the number of men who do not believe that there are benefits to FGM/C
- increase the number of men who believe that most men and women in the community favour discontinuation of FGM/C
- increase the number of women who do not intend to perform FGM/C on their daughters.

RESULTS OF COMMUNITY EMPOWERMENT THROUGH EDUCATION

Two studies with a total of 2,797 participants evaluated community empowerment through education (37;38).

Table 13 shows the results for women in Senegal for each outcome for which we could calculate effect estimates. The groups were judged to be different at baseline in five out of seven outcomes, consistently with higher scores in the intervention group. In two outcomes ('Were aware of at least two consequences of FGM/C' and self-

reported prevalence of 0-10 year old girls who had been cut), the groups were reasonably similar at baseline. The effect estimates for these two outcomes were in favour of the intervention group. The 95%CI was wide around the relative risk in the outcome 'Were aware of at least two consequences of FGM/C'.

Table 13: Study outcomes and effect estimates for women, in Diop (2004)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Self-reported prevalence								
0-10 year old girls who had been cut [$]	54	40	-14	52	52	0	-14	0.77 (0.64, 0.93)
Knowledge/Awareness								
Were aware of at least two consequences of FGM/C[1] [$]	7	73	66	3	25	22	44	2.92 (2.28, 3.74)
Beliefs/Attitudes								
Believed that FGM/C was unnecessary[1] [$]	30	85	55	12	39	27	28	2.18 (1.82, 2,61)
Disapproved of FGM/C[1] [£]	28	84	56	11	40	29	27	2.10 (1.76, 2.51)
Perceived that their husband disapproved of FGM/C[1] [$]	35	86	51	14	46	32	19	1.87 (1.60, 2.18)
Regretted having had their daughter cut [$]	66	90	24	52	58	6	18	1.55 (1.37, 1.76)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [£]	29	88	59	11	46	35	24	1.91 (1.64, 2.23)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups. 1= item reversed by us.

Table 14 shows the results for men in Senegal for each outcome for which we could calculate effect estimates. The groups were judged to be different at baseline in two out of three outcomes, the exception being 'Were aware of at least two consequences of FGM/C'. The effect estimates for this outcome were to the advantage of the intervention group, albeit with a wide 95%CI around the relative risk.

Table 14: Study outcomes and effect estimates for men, in Diop (2004)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Knowledge/awareness								
Were aware of at least two consequences of FGM/C [$]	11	66	55	14	21	7	48	3.10 (2.28, 4.23)
Beliefs/attitudes								
Believed that FGM/C was supported by religion [£]	14	34	20	25	48	23	-3	0.71 (0.51, 0.99)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [$]	34	87	53	22	44	22	31	1.97 (1.65, 2.36)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups.

Table 15 shows the results for women in Burkina Faso for each outcome for which we could calculate effect estimates. The groups were judged to be similar at baseline in seven out of eight outcomes, the exception being 'Discussed FGM/C with others'. The effect estimates were close to one in all outcomes, except 'Were aware of at least two consequences of FGM/C' and 'Regretted having had daughters cut'. In both of these, the effect estimates showed an effect in favour of the intervention group.

Table 15: Study outcomes and effect estimates for women, in Ouoba (2004)

Outcome	Intervention			Comparison			Adjusted	RR
	Pre	Post	Change	Pre	Post	Change	ARD	(95%CI)
Self-reported prevalence								
0-10 year old girls who had been cut [$]	6	3	-3	4	4	0	-3	0.74 (0.33, 1.66)
Behaviours								
Discussed FGM/C with others [$]	54	91	37	64	65	1	36	1.40 (1.27, 1.55)
Knowledge/Awareness								
Were aware of at least two consequences of FGM/C [$]	52	86	34	57	73	16	18	1.18 (1.08, 1.29)
Beliefs/Attitudes								
Believed that FGM/C was unnecessary[1] [£]	93	99	6	92	97	5	1	1.02 (1.00, 1.05)
Disapproved of FGM/C[1] [$]	89	98	9	90	94	4	5	1.04 (1.01, 1.08)
Perceived that their husband disapproved of FGM/C[1] [$]	97	99	2	97	96	-1	3	1.03 (1.00, 1.06)
Regretted having had their daughter cut [$]	53	81	28	50	64	14	14	1.26 (1.14, 1.40)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [$]	97	99	2	96	98	2	0	1.01 (0.99, 1.03)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups. 1= item reversed by us.

Table 16 shows the results for men in Burkina Faso for each outcome for which we could calculate effect estimates. The groups were judged to be similar at baseline in four out of five outcomes, the exception being 'Discussed FGM/C with others'. The effect estimates were low, except in the outcome 'Were aware of at least two consequences of FGM/C'. The effect estimates indicated an effect in favour of the intervention group.

Table 16: Study outcomes and effect estimates for men, in Ouoba (2004)

Outcome	Intervention			Comparison			Adjusted ARD	RR (95%CI)
	Pre	Post	Change	Pre	Post	Change		
Behaviours								
Discussed FGM/C with others [$]	62	92	30	79	75	-4	34	1.22 (1.13, 1.33)
Knowledge/awareness								
Were aware of at least two consequences of FGM/C [$]	64	88	24	66	60	-6	30	1.47 (1.31, 1.64)
Beliefs/attitudes								
Believed that FGM/C was unnecessary[1] [#]	97	98	1	97	92	-5	6	1.06 (1.02, 1.11)
Disapproved of FGM/C[1] [$]	90	98	8	93	89	-4	12	1.10 (1.05, 1.15)
Intentions								
Did not intend to continue the practice of FGM/C with their daughter [#]	97	98	1	98	93	-5	6	1.05 (1.01, 1.09)

Note: Pre-, and post scores are in percent and reproduced from the study publication. We calculated change scores in percentage points, adjusted absolute risk difference (ARD) and relative risk (RR) based on post-intervention data with 95% confidence interval (CI). Legend: Results of significance tests reported in the study publication are indicated $= significantly greater change in intervention group, £= significant change in both groups, #= significantly greater change in comparison group. 1= item reversed by us.

Behaviours were not measured by Diop and colleagues (37).

The results showed that community empowerment through education may:
- decrease the number of women who report having had 0-10 year old girls cut
- increase knowledge/awareness about consequences of FGM/C
- increase the number of women who regret having had their daughter cut.

Both studies employed a public declaration to abandon FGM/C as an outcome. Such an event was held at the end of the intervention period. Diop (37) reported that 26% of women and 28% of men who participated in the study attended a public declaration and Ouoba (38) reported that 66% of women and 74% of men attended a public statement.

Results of the meta-analyses

Only the two studies that evaluated effectiveness of community empowerment through education (37;38) were sufficiently similar to warrant pooling of effect sizes in meta-analyses. We assessed nine dichotomous outcomes (one separately for men and women). However, while we considered the studies similar enough to warrant meta-analysis, results revealed considerable heterogeneity for five of the nine outcomes, and for all but one of the outcomes, one study (37) had a higher number of events than the other study (38) and therefore assumed most weight. As a consequence, only one study contributed to the pooled effect size. The considerable heterogeneity and unequal weight affect the interpretation of effect sizes. Given the doubt about the validity of the meta-analyses results we chose to present the results from the individual studies. The meta-analysis results are included in Appendix 4 for transparency reasons (Figure A1 – A9).

Discussion

This systematic review aimed to summarize and assess the effectiveness of interventions designed to reduce the prevalence of FGM/C. We included six controlled before-and-after studies that were all set in Africa, where the practice of FGM/C is most prevalent. The studies were characterized by low methodological quality and we judged that there was 'high' or 'unclear' risk of bias for one or more of the assessed domains in all six included studies. It is disconcerting that only six studies were identified that met our inclusion criteria, and further that only two of the studies measured prevalence.

When pre- and post scores for both intervention and comparison groups were reported by study authors, we estimated effects of interventions by adjusted absolute risk difference and relative risk with corresponding 95% confidence intervals. Our calculated effect sizes for prevalence of FGM/C, knowledge, beliefs, and intentions about FGM/C suggested that there may be positive developments as a result of interventions. However, the low quality of the body of evidence affects the interpretation of results and raises doubts about the validity of the findings.

DISCUSSION OF MAIN RESULTS

Effectiveness of training health personnel

Our results do not provide convincing evidence that training health personnel is likely to have an effect on knowledge or belief/attitudes regarding FGM/C. This could be due to the time span provided for training (three sessions over two months). The study authors write that "training was apparently too short and should be expanded" (42). The WHO position is that "health care providers can play a key role in preventing female genital mutilation and in supporting and informing patients and communities about the benefits of eliminating it" through actions such as education and outreach (1). Unfortunately, results from the one study that evaluated the effects of an intervention programme for medical and health personnel (42), suggested that after the intervention, fewer Malian health personnel wished to play a role in educating others about the practice. Sense of advocacy among health personnel appeared to be low. Because this is a group that could play a role in halting the prevalence of FGM/C, it would be important to encourage advocacy in an effort to gain their active contribution.

Because medical personnel are important caretakers of girls and women who have been subjected to FGM/C and experience complications, and because such personnel are actors of power and authority in many communities, it is vital that their knowledge and skills about FGM/C are superior. But, another unfortunate

finding in Diop and colleagues' study (42), was that the training seemed to fail to significantly improve health personnel's knowledge level and beliefs about FGM/C. Their knowledge level remained low, particularly with regard to complications of FGM/C. For example, only 56% of them could name immediate complications that can result from FGM/C, and 39% believed that uncut women have "loose morals". This finding, coupled with health personnel's misperception that FGM/C is safe if performed in a hygienic environment, speaks to the importance of counteracting faulty heuristics and convictions about FGM/C among medical and health care personnel. While FGM/C that is performed by medical personnel in hospitals and health clinics may reduce some short-term complications regularly seen when it is performed by traditional practitioners, even when carried out by trained professionals, FGM/C is not necessarily less severe or conditions sanitary. Moreover, there is no evidence that medicalization reduces obstetric or other long-term complications associated with FGM/C (1). Most importantly perhaps, according to Budiharsana (48) the involvement of medical professionals in the practice may wrongly legitimize FGM/C as medically sound, and thereby contribute to the misconception that it is acceptable and medically safe. In effect, their involvement contributes to institutionalize FGM/C, because such personnel often hold power, authority, and respect. It should be noted that the study by Diop and colleagues (42) has low methodological quality, high risk of bias, and is over a decade old, thus knowledge and sense of advocacy about FGM/C among health personnel in Mali today may be at a higher level than this study indicates. Nevertheless, these findings are compelling and warrant attention by future researchers.

Effectiveness of educating female students

According to our results, two educational sessions about reproductive health, including dangers of FGM/C, seem to improve female students' knowledge/awareness about FGM/C. This is an encouraging finding and the role of education in the work to prevent FGM/C is discussed below.

Effectiveness of multifaceted community activities

Our effect size estimates suggest that knowledge about harmful consequences of FGM/C among Ethiopian, but not Somali study participants may have increased as a result of the intervention. This is an encouraging but also perplexing finding because the most commonly remembered message among the Somali participants was that infibulation was harmful to health. This result may be a reflexion of issues related study execution, which we discuss below.

One of our effect sizes indicated that multifaceted community activities may have resulted in a greater proportion of intervention participants believing FGM/C was against women's rights, another effect size indicated an effect in favour of the comparison group in this regard. These results may reflect issues related to the intervention itself as well as the cultures' norms with respect to women's position in society. Although the multi-component intervention in Ethiopia and Kenya sought