Psychological, social and sexual consequences of female genital mutilation/cutting (FGM/C): a systematic review of quantitative studies

kunnskapssenteret
Norwegian Knowledge Centre for the Health Services

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 conduct a systematic review about the consequences of female genital mutilation/cutting (FGM/C). The review would answer the question: What are the psychological, social and sexual consequences of FGM/C? **Methods:** We searched systematically for relevant literature in international scientific databases, in databases of international organisations that are engaged in aspects related to FGM/C, and in reference lists of relevant reviews and included studies. We also communicated with professionals working with FGM/C related issues. We selected studies according to pre-specified criteria, appraised the methodological quality using checklists, and summarized the study level results using tables and calculated effect estimates (risk ratio and mean difference) in addition to performing meta-analyses to estimate effect. We applied the instrument GRADE to assess the extent to which we could have confidence in the effect estimates. **Results:** We included and summarized results from 17 comparative *(continued)*

kunnskapssenteret

studies with a total of 12,755 participants from communities where FGM/C is practiced. All studies compared women with FGM/C to women without FGM/C. The evidence base was insufficient to draw conclusions about the psychological and social consequences of FGM/C. The effect estimates show that compared to women without FGM/C women with FGM/C are more likely to experience 1) pain during intercourse, 2) reduced sexual satisfaction, and 3) reduced sexual desire. However, the quality of the evidence was too low to draw conclusions about a causal relationship between FGM/C and psychological, social and sexual consequences. **Conclusion:** There is a paucity of high quality evidence regarding the consequences of FGM/C. While the evidence base is insufficient to draw causal conclusions about the consequences of FGM/C, our results show that women with FGM/C experience pain and reduction in sexual satisfaction and desire compared to women without FGM/C.

List of abbreviations

DHS	Demographic and Health Survey.
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th ed.
FGM/C	Female Genital Mutilation/Cutting.
FSFI	Female Sexual Function Index.
NKVTS	Nasjonalt kunnskapssenter om vold og traumatisk stress / Norwegian Centre for Violence and Traumatic Stress Studies.
NOKC	Nasjonalt kunnskapssenter for helsetjenesten / Norwegian Knowledge Centre for the Health Services.
PRB	Population Reference Bureau.
PTSD	Post traumatic stress disorder.
UNFPA	United Nations Population Fund.
UNICEF	United Nations Children's Fund.
WHO	World Health Organization.

Key messages

Psychological, social and sexual consequences of female genital mutilation/cutting (FGM/C): a systematic review of quantitative studies

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 conduct a systematic review about the consequences of female genital mutilation/cutting (FGM/C). The review would answer the question: What are the psychological, social and sexual consequences of FGM/C?

Methods: We searched systematically for relevant literature in international scientific databases, in databases of international organisations that are engaged in aspects related to FGM/C, and in reference lists of relevant reviews and included studies. We also communicated with professionals working with FGM/C related issues. We selected studies according to pre-specified criteria, appraised the methodological quality using checklists, and summarized the study level results using tables and calculated effect estimates (risk ratio and mean difference) in addition to performing meta-analyses to estimate effect. We applied the instrument GRADE to assess the extent to which we could have confidence in the effect estimates.

Results: We included and summarized results from 17 comparative studies with a total of 12,755 participants from communities where FGM/C is practiced. All studies compared women with FGM/C to women without FGM/C. The evidence base was insufficient to draw conclusions about the psychological and social consequences of FGM/C. The effect estimates show that compared to women without FGM/C women with FGM/C are more likely to experience 1) pain during intercourse, 2) reduced sexual satisfaction, and 3) reduced sexual desire. However, the quality of the evidence was too low to draw conclusions about a causal relationship between FGM/C and psychological, social and sexual consequences.

Conclusion: There is a paucity of high quality evidence regarding the consequences of FGM/C. While the evidence base is insufficient to draw causal conclusions about the consequences of FGM/C, our results show that women with FGM/C experience pain and reduction in sexual satisfaction and desire compared to women without FGM/C.

Executive summary

Female genital mutilation/cutting (FGM/C) is a traditional practice that involves "the partial or total removal of the female external genitalia 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 some countries in the Middle East and Asia. Although limited data exist, it is speculated that FGM/C is practised by immigrant communities in a number of other countries, including Australia, Canada, France, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The practice of FGM/C is rooted in social conventions 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 abrogates human rights. It is prohibited by law in several African and Western countries. The current WHO classification describes four types of FGM/C: Type I, *clitoridectomy*, involves partial or total removal of the clitoris and/or the prepuce. Type II, *excision*, involves partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora. Type III, *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 IV, *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 sociodemographic factors. 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, shock, infections, and difficulty in passing urine and faeces. Caesarean section, blood loss and increased perinatal mortality are associated birth risks. Non-medical consequences from FGM/C are less clear.

We asked the following question: What are the psychological, social and sexual consequences of FGM/C?

METHODS

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 research concerning FGM/C, manually in reference lists of relevant reviews and studies included in this systematic review, as well as communicated with experts engaged in FGM/C related work. We searched for studies that used the following study designs: systematic reviews, cohort studies, case-control studies, and cross-sectional studies.

Two of the authors independently assessed studies for inclusion according to pre-specified criteria and considered the methodological quality of the studies using checklists. We summarized the study level results in text and tables and calculated effect estimates (relative risk and mean difference). We also performed meta-analyses to estimate effect, using Mantel-Haenszel random effects meta-analyses for dichotomous outcomes and inverse variance random effects meta-analyses for continuous outcomes. We applied the instrument GRADE to assess the extent to which we could have confidence in the effect estimates.

RESULTS

We identified 3,669 publications and after having assessed titles, abstracts, and articles in full text we included 17 studies that fulfilled the inclusion criteria. All included studies were observational comparative studies (15 cross-sectional studies and 2 case-control studies) that compared women who had been subjected to FGM/C with women who had not been subjected to FGM/C. We failed to obtain two potentially relevant records, despite extensive retrieval efforts.

We arrived upon a final decision of low study quality for ten of the 17 studies, moderate quality for five and high quality for two. In our assessment, using the GRADE instrument, the quality of the evidence was very low with regards to documenting a causal relationship between FGM/C and psychological, social and sexual consequences. Collectively, the studies involved a total of 12,755 participants from nine different countries. One study was from Israel, one was from Saudi Arabia, while the remaining fifteen studies were from countries in Africa: Central African Republic, Egypt, Gambia, Ghana, Nigeria, Senegal, and Sudan.

Four studies reported on psychological consequences. Study level results suggested that women with FGM/C may be more likely than women without FGM/C to experience psychological disturbances (have a psychiatric diagnosis, suffer from anxiety, somatisation, phobia, and low self-esteem). However, our meta-analyses for anxiety, somatisation, depression, and hostility failed to reach significance and were marred by high heterogeneity. We were unable to draw solid conclusions concerning psycho-

logical consequences. Only two studies, both of low study quality, included some measure of social consequences of FGM/C and we were unable to draw any conclusions. Concerning sexual consequences, several studies were sufficiently similar to warrant pooling of effect sizes in meta-analysis for the outcomes pain during intercourse, satisfaction, desire, initiation of sex, orgasm, reporting clitoris as the most sensitive area of the body, and reporting the breasts as the most sensitive areas of the body. Compared to women without FGM/C, women with FGM/C were 1.5 times more likely to experience pain during intercourse (RR= 1.52, 95%CI= 1.15, 2.0). The pooled effect estimate from two studies suggested that women with FGM/C experience significantly less sexual satisfaction (St.mean diff= -0.34, 95%CI= -0.56, -0.13). Women with FGM/C were twice as likely to report that they did not experience sexual desire (RR= 2.15, 95%CI= 1.37, 3.36). These meta-analysis results were supported by other study level findings. The results of the remaining meta-analyses were marred by high heterogeneity and the results were inconclusive. Collectively, the results provide evidence that women with FGM/C are more likely to experience pain during intercourse, reduced sexual satisfaction and reduced sexual desire than women without FGM/C, but the low quality of the body of evidence precludes us from drawing conclusions regarding causality.

CONCLUSION

The psychological, social and sexual consequences of FGM/C is an under-researched and neglected issue. The low quality of the body of evidence precludes us from drawing conclusions regarding causality, and the evidence base is insufficient to draw solid conclusions about the psychological and social consequences of FGM/C. However, our results substantiate the proposition that a woman whose genital tissues have been partly removed is more likely to experience increased pain and reduction in sexual satisfaction and desire. Future studies investigating the consequences of FGM/C should compare clearly defined groups that differ by the extent of FGM/C, whereby classification is based on gynaecological examination. Multi-centre, comparative studies which apply a methodology that increases the likelihood of equivalency of exposed and unexposed groups of women and standardized data collection, are preferable.

The Norwegian Knowledge Centre for the Health Services 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 of Health, 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

1-side oppsummering (norsk)

Bakgrunn: I november 2008 gav Nasjonalt kunnskapssenter om vold og traumatisk stress (NKVTS) i oppdrag til Nasjonalt kunnskapssenter for helsetjenesten (NOKC) å utføre en systematisk kunnskapsoppsummering om konsekvensene av kjønnslemlestelse. Oppsummeringen skulle besvare spørsmålet: Hva er de psykiske, sosiale og seksuelle konsekvensene av kjønnslemlestelse?

Metode: Vi søkte systematisk etter relevant litteratur i internasjonale databaser, i databaser til internasjonale organisasjoner som driver prosjekter om kjønnslemlestelse, i referanselistene til relevante kunnskapsoversikter og de inkluderte studiene, og kommuniserte med eksperter som arbeider med kjønnslemlestelse. Vi valgte ut studier som oppfylte våre predefinerte inklusjonskriterier. Vi brukte sjekklister for å vurdere den metodiske kvaliteten til studiene og vurderte den samlede dokumentasjonen for endepunktene ved hjelp av GRADE. Vi oppsummerte resultater på studienivå i tabeller og beregnet effektestimat (relativ risiko og gjennomsnittsforskjell) samt utførte meta-analyser hvor vi mente dette var metodologisk forsvarlig.

Resultater: Vi inkluderte og oppsummerte resultatene for 17 observasjonsstudier som sammenlignet kvinner utsatt for kjønnslemlestelse med kvinner uten kjønnslemlestelse. Til sammen deltok 12,755 kvinner. Av de 17 inkluderte studiene ble 10 vurdert til å ha lav metodologisk kvalitet. Vi vurderte den samlede dokumentasjonen for endepunktene ved hjelp av GRADE som viste at dokumentasjonen hadde svært lav kvalitet, som innebærer at effektestimatet er for usikkert til at vi kan trekke noen kausale slutninger. Resultatene viser at kvinner med kjønnslemlestelse har større risiko for å oppleve 1) smerte under samleie, 2) nedsatt seksuell tilfredshet, og 3) nedsatt seksuell lystfølelse, sammenlignet med kvinner uten kjønnslemlestelse.

Konklusjon: Det mangler dokumentasjon av høy kvalitet når det gjelder konsekvenser av kjønnslemlestelse. Kunnskapsgrunnlaget er utilstrekkelig for å dra sikre konklusjoner angående psykiske og sosiale konsekvenser. Resultatene viser at kvinner med kjønnslemlestelse i større grad opplever seksuelle problemer enn kvinner uten kjønnslemlestelse, men også her er kunnskapsgrunnlaget utilstrekkelig og av for lav kvalitet til at vi kan konkludere at det er en direkte årsakssammenheng.

Sammendrag (norsk)

BAKGRUNN

Kjønnslemlestelse er en tradisjonell praksis som innebærer at hele eller deler av de eksterne kvinnelige kjønnsorganene fjernes eller skades av ikke-terapeutiske grunner. Kjønnslemlestelse praktiseres i mer enn 28 land i Afrika, i noen land i Midtøsten og Asia, og muligens i immigrantsamfunn i vestlige land som Australia, Canada, Frankrike, Norge, New Zealand, Storbritannia, Sveits, Sverige, og USA. Kjønnslemlestelse er grunnet i kulturelle 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 menneskelige rettigheter og er uttrykkelig forbudt i mange afrikanske og vestlige land. Kjønnslemlestelse klassifiseres i fire kategorier: Type I, *klitoridektomi*, delvis eller total fjerning av klitoris og/eller forhuden; Type II, *eksisjon*, delvis eller total fjerning av klitoris og de små kjønnsleppene; Type III: *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 IV: *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, sjokk, vanskeligheter med urinering og avføring, samt infeksjoner. Keisersnitt, blodtap og økt perinatal dødelighet er risikofaktorer i sammenheng med fødsel. Det fins mindre forskning angående andre konsekvenser av kjønnslemlestelse.

Vi stilte følgende spørsmål: Hva er de psykiske, sosiale og seksuelle konsekvensene av kjønnslemlestelse?

METODE

Vi søkte systematisk etter relevant 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, i referanselistene til relevante kunnskapsoversikter og de inkluderte studiene, og kommuniserte med eksperter som arbeider med kjønnslemlestelse. Vi søkte etter litteratur med følgende studiedesign: systematiske kunnskapsoversikter, kohortestudier, kasuskontrollstudier, og tverrsnittsstudier.

Vi valgte ut studier som oppfylte våre predefinerte inklusjonskriterier. Deretter brukte vi sjekklister for å vurdere den metodiske kvaliteten til studiene. Vi oppsummerte resultater på studienivå i tabeller og beregnet effektestimat (relativ risiko og gjennomsnittsforskjell) samt utførte meta-analyser hvor vi mente dette var metodologisk forsvarlig. Vi vurderte den samlede dokumentasjonen for endepunktene ved hjelp av GRADE.

RESULTAT

Vi identifiserte 3669 publikasjoner og etter å ha vurdert titler, sammendrag og artikler i fulltekst fant vi 17 studier som oppfylte inklusjonskriteriene. Alle studiene var observasjonsstudier (15 tverrsnittsstudier og 2 kasuskontrollstudier) som sammenlignet kvinner utsatt for kjønnslemlestelse med kvinner uten kjønnslemlestelse. To mulig relevante publikasjoner fikk vi ikke tak i, på tross av omfattende forsøk.

Av de 17 inkluderte studiene ble 10 vurdert til å ha lav metodologisk kvalitet. Vi vurderte kvaliteten på den samlede dokumentasjonen for endepunktene ved hjelp av GRADE til svært lav. Det betyr at effektestimatet er for usikkert til å kunne dokumentere en kausal sammenheng mellom kjønnslemlestelse og psykiske, sosiale og seksuelle konsekvenser. Til sammen deltok 12 755 kvinner i de 17 studiene, som var utført i ni ulike land. Én studie var fra Israel, én var fra Saudi Arabia, og de andre 15 studiene var fra land i Afrika: Den Sentralafrikanske republikk, Egypt, Gambia, Ghana, Nigeria, Senegal, og Sudan.

Fire studier inkluderte psykiske utfallsmål. Resultater på studienivå tydet på at kvinner med kjønnslemlestelse kan ha større risiko for å oppleve psykiske problemer (inneha psykiatrisk diagnose, lide av angst, somatisering, fobi, lav selvfølelse). Men meta-analyse resultatene for angst, somatisering, depresjon, og fiendtlighet viste ikke signifikante forskjeller og høy heterogenitet tilsa uforenlighet mellom studiene. Vi kunne ikke trekke solide konklusjoner for psykiske konsekvenser. Kun to studier, med lav studiekvalitet, inkluderte sosiale utfallsmål og vi kunne ikke trekke noen konklusjoner. I forhold til seksuelle konsekvenser var flere studier tilstrekkelig like til at vi kunne utføre meta-analyser. Vi utførte meta-analyser for syv utfallsmål: smerte under samleie, seksuell tilfredshet, lystfølelse, ta initiativet til sex, orgasme, mene at klitoris er den mest følsomme delen av kroppen, mene at brystene er de mest følsomme delene av kroppen. Sammenlignet med kvinner uten kjønnslemlestelse var kvinner med kjønnslemlestelse 1.5 ganger mer utsatt for smerte under samleie (RR= 1,52, 95 % CI= 1,15 til 2,0). Resultatet fra to studier viste at kvinner med kjønnslemlestelse opplever mindre seksuell tilfredshet (St.mean diff= -0,34, 95 % CI

-0,56 til -0,13) enn kvinner uten kjønnslemlestelse. Kvinner med kjønnslemlestelse var mer enn to ganger så utsatt for ikke å oppleve seksuell lystfølelse (RR= 2,15, 95 % CI= 1,37 til 3,36). Lignende resultater på studienivå støttet disse funnene. Det var høy heterogenitet i de andre meta-analysene og resultatene var ikke entydige. Til sammen viser disse resultatene at kvinner med kjønnslemlestelse i større grad opplever seksuelle problemer enn kvinner uten kjønnslemlestelse, men kunnskapsgrunnlaget er for svakt til at vi kan slå fast at dette er en direkte følge av kjønnslemlestelse.

KONKLUSJON

Det fins lite forskning om de psykiske, sosiale og seksuelle konsekvensene av kjønnslemlestelse. Kunnskapsgrunnlaget er utilstrekkelig for å dra sikre konklusjoner angående psykiske og sosiale konsekvenser, men resultatene fra denne studien viser at kvinner som har vært utsatt for kjønnslemlestelse er mer utsatt for seksuelle problemer, slik som smerte under samleie, og redusert seksuell tilfredshet og lystfølelse. Kunnskapsgrunnlaget er utilstrekkelig og av for lav kvalitet til å dra kausale slutninger. For å få mer sikker kunnskap er det behov for ytterligere studier. Slike studier bør sikre at gruppene av kvinner med og uten kjønnslemlestelse er sammenlignbare, at forekomst og grad av kjønnslemlestelse er målt ved fysisk undersøkelse, og at standardiserte datainnsamlingsmetoder blir brukt.

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 psychological, social and sexual consequences of FGM/C. One systematic review, about the effectiveness of interventions designed to reduce the prevalence of FGM/C, is published (1). Another systematic review, about the factors promoting and hindering the practice, is underway, and a third, delineating extant guidelines regarding FGM/C will likely be completed by the end of 2010.

The project team was composed of:

- Project coordinator, researcher, Eva Denison, NOKC
- Researcher, Rigmor C Berg, NOKC
- Research director, Atle Fretheim, NOKC

Additionally, the literature search was conducted by research librarian Sari Ormstad, and Jan Odgaard-Jensen provided statistical guidance. Both are with the NOKC.

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

- Susan Munabi-Babigumira, researcher, NOKC
- Tove Ringerike, researcher, NOKC
- R. Elise Johansen, Technical officer at World Health Organization, Geneva
- Anonymous reviewer

The aim of this report is to support well-informed decisions in health promotion that inform the work to reduce the prevalence of FGM/C and improve 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*	*Project coordinator*

Objective

This systematic review summarizes published, empirical quantitative data describing the psychological, social and sexual consequences of FGM/C.

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 inform work to reduce the prevalence of FGM/C and improve quality of services related to FGM/C.

The main research question was:

* What are the psychological, social and sexual consequences of FGM/C?

Background

The terminology used for the cutting of external female genital tissues varies. It has been referred to as "female circumcision", "female genital mutilation", "female genital cutting" and "female genital mutilation/cutting" (2). 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 of the female external genitalia or other injury to the female genital organs for cultural or other non-therapeutic reasons" (3).

Although FGM/C transcends geography, it is primarily practised among various ethnic groups in more than 28 countries in Africa, usually on girls under the age of 15 years. The practise is also found in some countries in the Middle East and Asia (4;5), for example among some Bedouin tribes in the western part of Saudi Arabia (6). Although limited data exist, it is speculated that FGM/C is practised by immigrant communities in a number of other countries, including Australia, Canada, France, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States (4). Unlike male circumcision, which provides some protection from certain infections, such as urinary tract infections and human immunodeficiency virus (7;8), there are no known health benefits to FGM/C (2).

CLASSIFICATIONS

To clarify understanding of the prevalence and consequences of FGM/C, WHO has classified the procedure into four categories: Type I, *clitoridectomy*, involves partial or total removal of the clitoris and/or the prepuce. Type II, *excision*, involves partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora. Type III, *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 IV, *other*, involves all other harmful procedures to the female genitalia for non-medical purposes, for example: pricking, piercing, incising, scraping, and cauterizing (2). In type IV, no genital tissue is excised.

Within these classifications there is a wide range of variation, not yet systematically studied and documented. For example, the technical variation of cutting differs within the same practitioner over time and instrumentation used, resulting in variation in degrees of FGM/C (2;9). Common to all operations, except type IV procedures, is some degree of excision of the external genitalia, from excision of minor skin parts around the clitoris to clitoridectomy and removal of the labia. Each community use the cutting of their own choice for their own reasons and beliefs. Similarly, various instruments are used to perform the procedure, including razor blades, glass, knives, and scissors (10).

PREVALENCE

Globally, it has been proposed that FGM/C type II is the most frequently practiced form, representing an estimated 80% of all procedures of FGM/C (11). While type III is thought to represent about 10% of FGM/C in Africa, it is probably the most frequently used type of FGM/C in some countries, including Djibouti, Somalia and northern Sudan (12).

Recent national 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 (12). 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 (5). UNICEF (5) has proposed that countries be categorized in three groups according to prevalence rates. 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. In Table 1, we present prevalence data of FGM/C among women aged 15-49 from countries which will be discussed in this systematic review and for which comparable national data exist.

Table 1: Prevalence of FGM/C in selected countries

Country	Total prevalence	Year data collected	Group
Central African Republic	35.6	2000 [1]	2
Egypt	91.0	2008 [2]	1
Gambia	55.0	Not stated [1]	2
Ghana	5.4	2003 [1]	3
Nigeria	19.0	2003 [2]	3
Senegal	28.2	2005 [1]	2
Sudan	89.2	1990 [2]	1

Legend: 1= data from Yoder & Kahn (12). 2= data from most recent DHS survey in the country (13-15).

No national data exist for Israel and Saudi Arabia, but studies document the practice among some ethnic groups in these areas. In Israel, the practice has been reported

among particularly two ethnic groups, Ethiopian Jews and the Bedouin of southern Israel (16). However, while FGM/C was considered normative among Jews in Ethiopia the custom appeared to be abandoned upon arrival in Israel (17). Similarly, while Bedouin women previously declared female genital surgery as an important part of their cultural identity (18), a recent study concluded that the practice had virtually disappeared among Israel's Bedouin population (19). Although indigenous to Saudi Arabia, FGM/C is currently only practiced among a few immigrant groups, Bedouin tribes and residents of the Hejaz (6).

INTERVENTIONS TO REDUCE THE PREVALENCE OF FGM/C

Consistent with international condemnation of FGM/C, there has been an increasing amount of initiatives geared towards the eradication of the practice among practising communities. As Western governments have become more aware of FGM/C among the immigrant communities, legislation as the main intervention tool has been instituted, and European Union institutions and Member States have taken active steps towards ending FGM/C (20;21). There are now laws prohibiting FGM/C in several Western countries, including Australia, Canada, New Zealand, USA, and at least 13 countries in Western Europe (4;5). However, the implementation of anti-FGM/C laws and their impact on eliminating the practice has so far not been extensively studied (22).

Efforts to abandon the practice of FGM/C in Africa have used several different approaches, including those based on legal mechanisms, human rights frameworks, health risks, alternative rites, positive deviance, training health workers as change agents, training and converting circumcisers, and the use of comprehensive social development approaches. Interventions based on these approaches have targeted stakeholders at individual, interpersonal, community, and national levels (23). Recently, the Population Reference Bureau (PRB) carried out an extensive survey of current intervention projects taking place in African countries (24). 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.

Although such surveys 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 (23;24), rigorous appraisal of the evidence was until recently lacking. A 2009 systematic review (1) took stock of progress to date, using rigorous methods which allowed valid assessment of intervention effects. The authors identified and included six controlled, before-and-after studies. Each study was set in a different country in Africa: Burkina Faso, Egypt, Ethiopia / Kenya, Mali, Nigeria, and Senegal. Two of the studies were directed at the individual level

and four at the community level. The results suggested 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 reduced prevalence of FGM/C, increased participants' knowledge about the consequences of FGM/C, and increased regrets about having had daughter cut. However, the authors warned that the low quality of the body of evidence affected the interpretation of results and raised doubts about the strength of the findings. The authors concluded that there was a paucity of high quality evidence regarding the effectiveness of interventions to reduce the prevalence of FGM/C and called for second-generation studies, which at a minimum should be of high methodological quality, focus on prevalence, and take into account regional, ethnic and sociodemographic variation in the practice of FGM/C. A protocol has been submitted to the Campbell Collaboration, in which we propose to assess not only effectiveness of interventions, but also the extent to which intervention programs have heeded and built upon factors promoting and hindering FGM/C, the extent to which interventions have been provided to the most appropriate stakeholder groups, and which forces may have been overlooked as critical program elements.

CONSEQUENCES

FGM/C causes permanent, irreparable changes in the external female genitalia. The external female organs encompass the mons pubis, clitoris, labia minora, labia majora, and vaginal opening (Figure 1). All of these organs, collectively known as the vulva, serve important sexual functions. The clitoris, with its rich supply of nerve endings, and prepuce form the most consistently erotic area of the female body (25). Sexual organs in females and males arise from the same embryologic origin. For example, the clitoris is equivalent to parts of the male penis (corpora cavernosa) (25). Further, anatomical, histological and MRI-based research has identified five sexually responsive, vascular tissues of the female external genitalia: clitoris, clitoral bulbs, labia minora, urethra, and vestibule/vagina. All were found to engorge during sexual arousal, with the erectile tissue compartments (corpus cavernosum of the clitoris and corpus spongiosum of the clitoral bulbs) having the greatest change in blood volume during sexual arousal (26). Researchers have concluded that vascular tissue is important in the context of female sexual response. Operations that disrupt or potentially disrupt the female external organs, such as FGM/C, "can potentially affect sexual functioning by ablating some or all of the genital organs, or their innervation" (26) as well as damage neural innervation (27). In effect, with FGM/C, some fundamental structures for sexual stimulation and orgasm have been excised, although not all. In contrast to male circumcision where the foreskin is cut off from the tip of the penis without damaging the organ itself, the degree of cutting and

likely harm in FGM/C is anatomically much more extensive (28) (with the possible exception of variations of type IV, such as pricking and stretching).

Figure 1: Female external genitalia
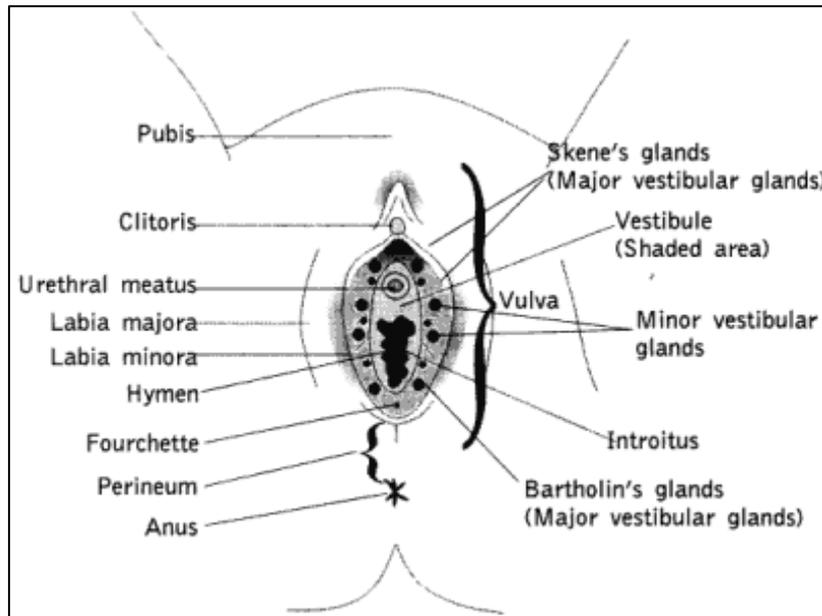
Physical consequences

Girls exposed to FGM/C are at risk of immediate physical consequences such as severe pain, bleeding, shock, difficulty in passing urine and faeces, and infections. Long term consequences can include chronic pain and infections. 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 (2).

A systematic review (29) of the health complications of FGM/C 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 (30) investigating 28,393 women attending obstetric centres in African countries concluded that women with FGM/C were significantly more likely than those without to have adverse obstetric outcomes such as a caesarean section, infant resuscitation, and inpatient perinatal death. The authors also concluded that the risks seemed to be greater with more extensive FGM/C. The associations were modest, however, and two studies from Sweden did not confirm a link between FGM/C and prolonged labour or perinatal death (31;32). Also the literature regarding infertility is inconclusive. While a study using DHS data from the Central African Republic, Côte d'Ivoire and Tanzania failed to confirm a statistical association between FGM/C and infertility (33), a case-control study from Sudan concluded there was a statistically as well as clinically significant association between FGM/C and primary infertility (34).

It is possible that FGM/C performed by medical personnel in health clinics may reduce some short term complications regularly seen when it is performed by traditional practitioners. However, conditions are not necessarily sanitary or complications less severe. Further, there is no evidence that medicalization reduces obstetric or other long term complications associated with FGM/C (2). Henceforth, the medical profession, led by the WHO and the World Medical Association, has condemned medicalization of FGM/C (35).

Other consequences

Similarly to the general physical consequences of FGM/C, the impact of FGM/C on other areas of women's health, particularly psychological, social and sexual health, has not been sufficiently investigated.

Psychological consequences

For many girls and women, undergoing FGM/C is a traumatic experience that may leave a lasting psychological mark and adversely affect their mental health. As an example, some researchers assert that FGM/C, representing a violation of women's physical intactness, can be classified as a "psychological trauma according to DSM-IV and a potential cause of posttraumatic stress disorder" (36). Karim (in (37)) stated that FGM/C leads to psychological disturbances and Toubia (28) explained that many infibulated women in Sudan suffer chronic anxiety and depression as a result of worry over their cut status. A qualitative study of the psychosocial impact of FGM/C among Bedouin-Arabs in Israel found that women with FGM/C expressed various emotional difficulties and psychosocial problems, including loss of trust within the mother-daughter relationship. Women reported feelings of fear, helplessness, and anger related to FGM/C. One said: "I perceive it as abuse" (38). Recently, Vloeberghs and colleagues (39) interviewed 66 women with various types of FGM/C. Most of the women, who lived in the Netherlands, reported a number of psychological effects from FGM/C, including anxiety, bad memories and stress. On the other hand, many women with FGM/C report feeling proud (40) and some argue FGM/C made them a better person: "I was very happy to go through it since I had been looking forward to it" (41).

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. FGM/C can be a source of personal and collective identity, as well as power in their daily affairs, as illustrated by one Sudanese woman: "It gives women a lot of power in the household" (42). When girls and families conform to the practice they acquire social position and respect. Conversely, failure to conform leads to difficulty in finding a husband for the girl, shame, stigmatization, as well as loss of social position, honour and protection, resulting in the family's social exclusion in the community (35;43;44). Women from five African countries reported that FGM/C influenced their relations with their partner, children and relatives in their country of origin (39).

Sexual consequences

As stated earlier, with FGM/C, parts of the women's erogenous genital areas as well as sexually responsive vascular tissue are removed. From the above concepts of the female sexual response, it is easy to deduce that excision of women's genital parts, coupled with damaged nerve-endings and the development of scar tissue and adhesions around the excised parts, reduce a women's capacity for sexual enjoyment (45). It has been found that vascular tissue is important for sexual response (26). Some researchers state that the sensitivity and integrity of the clitoris and labia minora are essential for experiencing sexual satisfaction (45). However, women's sexuality, as men's, is a complex interaction of neurophysiological and biochemical mechanisms, and influenced by relationship dynamics and family and sociocultural issues (46). It is also likely that a woman's sexuality is affected according to the extent of excision and the degree to which other social messages that inhibit sexual expression are internalized (28). For example, although interviews with Bedouin women in southern Israel (18) and Eritrean women (47) found that women reported pain during intercourse in the months after marriage, none felt that is was related to having been cut. On the other hand, Mukoro (48) concluded that sexual satisfaction diminished after FGM/C was performed. In this study of 46 women with FGM/C type I or II carried out in young adulthood or adolescence, the majority of the women (63%) had satisfactory sexual relationships before being cut, but this dropped to 8.7% after being cut.

A few other studies report on sexual arousal and experiences. Reyners (49) referring to a study among 2000 Egyptian women, summarized that 25% of women with FGM/C did not feel any arousal during intercourse, 50% experienced pain during intercourse, and 56% never experienced orgasm. Among participants in Ghana (50) 46% reported they did not enjoy sex and 41% said sex was associated with pain. In contrast, among 195 immigrated women in northern Europe, 90% of adult women with FGM/C reported that sex gave them pleasure and 69% always experienced orgasm (51). Also results from qualitative studies show a spectrum of experiences. Lightfoot-Klein (52) found that the majority of her infibulated Sudanese interviewees experienced sexual desire, pleasure and orgasm. Johnsdotter and Essen (53) reported that most of their respondents stated they did not have sexual problems and enjoyed sexual relations. Johansen (54) found that a little more than half of the female Somali respondents complained of reduced sexual feelings from FGM/C. This variation was evident also in Abusharaf's study (42): For example, one woman stated "My infibulations did not eliminate my desire to have sex even at this age" while another said "I have to tell the truth: infibulating does not allow women to want sex." Among Eritrean interviewees, sexual satisfaction came from being in tune with their partner and having a loving relationship (47).

In the mid 1990s, Toubia (28) concluded that little scientific research was available on the sexual consequences of FGM/C, and Obermeyer's review (55) a few years later confirmed that only a handful of studies existed. Until now, the researcher's

updated review of the consequences of FGM/C for health and sexuality may have represented the best available evidence regarding the sexual sequelae of FGM/C. The review included 35 sources and the results with respect to sexuality were summarized: "while one study ... reports that circumcised women are significantly more likely to suffer adverse consequences for sexual enjoyment, other studies that measure sexual activity and pleasure find no significant difference between circumcised and uncircumcised women" (56). Obermeyer concluded that most of the existing studies suffered from conceptual and methodological shortcomings, and the available evidence did not support the hypotheses that FGM/C destroys sexual functioning or precludes enjoyment of sexual relations.

Presently, there are no systematic reviews of the impact of FGM/C on women's psychological, social and sexual health. Thus, the aim of this systematic review was to summarize published, empirical data describing the psychological, social and sexual consequences of FGM/C. We believe a systematic review on this topic is important for several reasons. A systematic review provides a more precise estimate of consequences than is possible from individual, primary studies. It can form the basis of evaluating current practice as well as highlighting gaps and uncertainties in current research knowledge, and thus the research questions that need to be addressed in subsequent studies.

Method

We conducted a systematic review of the psychological, social and sexual consequences of FGM/C in accordance with the NOKC handbook for summarizing evidence (57) and most of the guidelines in the Cochrane Handbook for Systematic Reviews of Interventions (58).

LITERATURE SEARCH

The database search strategy was designed and executed February 4-9 2009, and repeated in February 2010, by research librarian Sari Ormstad at 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 strategy incorporated subject headings (for example MeSH terms in MEDLINE) and text words (in title and abstract) relating to FGM/C and the four classifications thereof, such as mutilation, circumcision, excision. No method filters were applied as we were more concerned about sensitivity than specificity. The complete 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 searches in reference lists of relevant

reviews and included studies. Additionally, we communicated with experts engaged in FGM/C related work. 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. cohort studies,
3. case-control studies,
4. cross-sectional studies.

Population: Girls and women who had been subjected to any type of FGM/C, as classified by the WHO (2).

Interest: FGM/C classified as type I to type IV according to the WHO modified typology (2).

Comparison: Girls and women who had not been subjected to any type of FGM/C, as defined by the WHO (2).

Outcomes: Psychological, social and sexual consequences of FGM/C, including but not limited to anxiety, post traumatic stress syndrome, school absenteeism, sexual satisfaction, libido and dyspareunia (pain during intercourse). The focus was behavioural and experiential consequences in the three interest areas psychological, social and sexual consequences.

We enforced no limitations on age, race/ethnicity, nationality or other participant characteristics. The women needed to be part of a community in which FGM/C was a customary practice. We included all publication years and languages. When considered likely to meet the inclusion criteria, studies were translated to English.

In sum, studies eligible for the review were those that as a minimum compared two groups of females (girls/women with FGM/C vs girls/women without FGM/C) with respect to a quantitatively measured psychological, social or sexual outcome.

EXCLUSION CRITERIA

We excluded all studies not meeting our pre-specified inclusion criteria. Specifically, we excluded non-systematic reviews, qualitative studies, quantitative studies that did not compare women with FGM/C to women without FGM/C. With respect to outcomes, we excluded studies without a quantitative measure of a psychological, social or sexual outcome.

SELECTION OF STUDIES

Two authors (Berg and Denison) independently read all titles/and or abstracts resulting from the search process. We compared our judgements and eliminated any obviously irrelevant publications. Next, we obtained full text copies of the potentially relevant studies (two records could not be obtained in full text). The same pair of authors, acting independently, classified the studies read in full text as relevant, that is, met all inclusion criteria and therefore to be included, or irrelevant and therefore to be excluded. We then compared our judgements and excluded studies that did not meet all inclusion criteria. Pre-designed inclusion/exclusion forms were used for each screening level.

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, comparison groups, and outcome measures used. Differences in opinion in the screening process were few and were resolved through consensus. Studies formally considered in full text 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 by re-examination of the publication and subsequent discussion.

With respect to quality of included studies, we used an adapted version of the NOKC quality assessment tool for cross-sectional studies. Given our focus on consequences of exposure to FGM/C, the assessment tool was modified by the addition of five questions from the NOKC quality assessment tool for cohort studies in order to capture whether a) the compared groups (women with FGM/C and women without FGM/C) were selected from the same population; b) the groups were comparable with respect to important backgrounds factors; c) exposure and outcome were measured in the same way in the two groups; d) the person who assessed the outcome was blind to whether participants were exposed or not; and e) known, potentially important confounders had been considered in the study design and/or analyses. The adapted check list with its 12 questions is listed in Appendix 3, including our assessment of each question. Berg and Denison agreed upon a final decision of strong, moderate or weak methodological quality for each study after discussing whether there was a discrepancy between the two reviewers with respect to the questions. We applied this assessment tool for all included studies (cross-sectional and case-control studies).

We applied the instrument Grading of Recommendations Assessment, Development and Evaluation (GRADE) with GRADE-Profiler version 3.2.2 to assess the extent to which we could have confidence in the effect estimates (58). We applied six criteria: type of study, methodological quality of study, consistency (were results consistent across studies?), directness (did the evidence directly answer the health care question?), precision (were the results precise enough?), and publication bias. For more details about the GRADE system we refer to publications by the GRADE Working Group (59). We used the standard definitions in grading the quality of the evidence: High = further research is very unlikely to change our confidence in the estimate of effect; Moderate = further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; Low = further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; Very low = any estimate of effect is very uncertain.

We extracted and grouped dichotomous and continuous data for all outcomes meeting the inclusion criteria. We grouped the descriptive data according to outcomes across the studies and presented data for women with FGM/C and without FGM/C separately in the results tables to allow for comparison. We estimated effect on dichotomous variables by the relative risk (RR) and 95% confidence interval (95%CI) and effect on continuous variables by mean difference (mean diff) and 95%CI. We also decided, *a priori*, if studies were sufficiently similar, to pool those that could be grouped together and use the statistical techniques of meta-analyses to estimate effect. To be pooled, studies had to assess the same outcome in a similar manner in similar populations. As is standard, we used Mantel-Haenszel random effects meta-analyses for dichotomous outcomes and inverse-variance random effects meta-analyses for continuous outcomes. Given that under the fixed effects model it is assumed that all studies in the meta-analysis are replications of each other, we deemed it appropriate to use the random effects model. In this model the more relaxed assumption is made that the included studies can be seen as a sample drawn from a population of studies, and each primary study is allowed to introduce its own amount of heterogeneity into the meta-analysis. This is reflected in the broader 95%CIs usually observed under the random effects model and its more conservative test results. We also examined between-study heterogeneity, with the Chi-square (Chi²) and I-square (I²) tests. A high value shows that most of the variability across studies is due to heterogeneity rather than to chance. We used RevMan 5, Cochrane Collaboration's meta-analysis software (58).

Results

DESCRIPTION OF INCLUDED LITERATURE

Results of the search

The electronic search resulted in 3,666 individual records and the manual search in three relevant records (Figure 2). (The 2010 electronic update search returned 755 relevant records, of which none were eligible for inclusion.) After removal of duplicates, Berg and Denison screened the records by reviewing titles and abstracts. We eliminated obviously irrelevant records based on titles and where available, abstracts.

We excluded 3,632 records, leaving 38 potentially relevant records. Two records (60;61) could not be obtained in full text, despite extensive retrieval efforts through both national and international libraries and attempts at contacting the authors. We read the full text for 36 publications. We excluded 19 publications (characteristics of excluded studies are presented in Appendix 2) and included 17 studies.

Figure 2: Flowchart of the reviewing process

4,434 records 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	4

→ 765 duplicates removed

3,670 titles/abstracts read

→ 3,632 irrelevant records excluded

→ 2 relevant records have not been obtained in full text

36 publications read in full text

→ 19 publications excluded due to:
- women with FGM/C not compared to women without FGM/C (16)
- study did not include psychological, social or sexual outcome (3)

17 studies included

Description of included studies

A total of 17 primary studies were included in this systematic review:

- Adinma (1997). Current status of female circumcision among Nigerian Igbos. (62)

- Alsibiani (2010). Sexual functioning in women with female genital mutilation. (63)

- Applebaum (2008). Symptoms of posttraumatic stress disorder after ritual female genital surgery among Bedouin in Israel: Myth or reality? (16)

- Badawi (1989). Epidemiology of female sexual castration in Cairo, Egypt. (9)

- Behrendt (2005). Posttraumatic stress disorder and memory problems after female genital mutilation. (64)

- El-defrawi (2001). Female genital mutilation and its psychosexual impact. (65)

- Elnashar (2007a). The impact of female genital cutting on health of newly married women. (66)

- Elnashar (2007b). Female sexual dysfunction in Lower Egypt. (67)

- Megafu (1983). Female ritual circumcision in Africa: An investigation of the presumed benefits among Ibos of Nigeria. (68)

- Morison (2001). The long-term reproductive health consequences of female genital cutting in rural Gambia: A community-based survey. (69)

- Nwajei (2003). Female genital mutilation: Implications for female sexuality. (70)

- Odoi (1997). Female genital mutilation in rural Ghana, West Africa. (71)

- Okonofua (2002). The association between female genital cutting and correlates of sexual and gynaecological morbidity in Edo State, Nigeria. (72)

- Osinowo (2003). Impact of female genital mutilation on sexual functioning, self-esteem and marital instability of women in Ajegunle. (73)

- Shandall (1967). Circumcision and infibulations of females. A general consideration of the problem and a clinical study of the complications in Sudanese women. (74)

- Stewart (2002). Determinants of coital frequency among married women in Central African Republic: The role of female genital cutting. (75)

- Thabet (2003). Defective sexuality and female circumcision: The cause and the possible management. (45)

All of the included studies were published in peer-reviewed journals and with the clear exception of Shandall's 1967 study (74), most can be considered relatively new publications, the great majority being published within the last decade (Table 2). Two studies were judged to be of high methodological quality (69;75), five were judged to be of moderate quality (16;63-65;72) and the remaining ten studies were of low methodological quality. Collectively, the studies involved a total of 12,755 participants from nine different countries. One study was from Israel (16), one was from Saudi Arabia (63), while the remaining 15 studies were from countries in Africa: Central African Republic, Egypt, Gambia, Ghana, Nigeria, Senegal, and Sudan.

The most frequently reported outcome was sexual functioning, such as desire or arousal, and psychological outcomes, such as anxiety, were also regularly reported. Only two studies (66;73) reported results for social outcomes.

Table 2: Included studies (N= 17)

Author, year	Study quality	Population	Study design	Results
Adinma, 1997 (62)	Low	N=256, Nigeria (southeast)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual outcome. There were significant differences in sexual satisfaction during intercourse (favouring non-cut).

Alsibiani, 2010 (63)	Moderate	N=260, Saudi Arabia	Prospective case-control (cut vs non-cut). Clinical-/hospital-based	Sexual outcomes. There were significant differences in scores for arousal, lubrication, orgasm, satisfaction, overall female sexual function index (favouring non-cut).
Applebaum, 2008 (16)	Moderate	N=37, Israel (south)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Psychological outcomes. There were no significant differences.
Badawi, 1989 (9)	Low	N=159, Egypt (Cairo)	Cross-sectional comparative (cut vs non-cut). Purposive	Sexual outcomes. There were significant differences in orgasm, sexual excitement, masturbation (favouring non-cut).
Behrendt, 2005 (64)	Moderate	N=47, Senegal (Dakar)	Cross-sectional comparative (cut vs non-cut). Purposive	Psychological outcomes. There were significant differences in post traumatic stress disorder, other anxiety disorders, affective disorders, psychiatric diagnosis (favouring non-cut).
El-defrawi, 2001 (65)	Moderate	N=250, Egypt (Ismailia)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual outcomes. There were significant differences in sexual desire, enjoyment of sex, orgasm, dryness during intercourse, dyspareunia (favouring non-cut).
Elnashar, 2007a (66)	Low	N=264, Egypt (Benha city)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Psychological, social, sexual outcomes. There were significant differences in somatisation, anxiety, phobia; marital satisfaction; pain, loss of libido, orgasm (favouring non-cut).
Elnashar, 2007b (67)	Low	N=936, Egypt (Dakahlia governorate)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual. There was a significant difference in sexual problems (favouring non-cut).
Megafu, 1983 (68)	Low	N=500, Nigeria	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual. There were significant differences in experience of orgasm and most erotic organ of the body (favouring non-cut).
Morison, 2001 (69)	High	N=1,157, Gambia (Farafenni area)	Cross-sectional comparative (cut vs non-cut). Community-based	Sexual. There was no significant difference in experiencing painful sex.
Nwajei, 2003 (70)	Low	N=400, Nigeria (Abraka)	Cross-sectional comparative (cut vs non-cut). Cluster	Sexual. There were differences in frequency of sexual intercourse.*
Odoi, 1997 (71)	Low	N=195, Ghana (north)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual. There were differences in dyspareunia, and orgasm.*
Okonofua, 2002 (72)	Moderate	N=1,836, Nigeria (Edo state)	Cross-sectional comparative (cut vs non-cut). Clinical-/hospital-based	Sexual. There were significant differences in initiating sex and most sensitive bodily part (favouring non-cut).
Osinowo, 2003 (73)	Low	N=99, Nigeria (Lagos)	Cross-sectional comparative (cut vs non-cut). Purposive	Psychological, social, sexual outcomes. There were significant differences in self-esteem; marital instability; sexual functioning (favouring non-cut).
Shandall, 1967 (74)	Low	N=4,024, Sudan (Khartoum area)	Cross-sectional comparative (FGM/C type I vs type III vs non-cut). Clinical-/hospital-based	Sexual outcomes. There were differences in frequency of orgasm.*
Stewart, 2002 (75)	High	N=2,188, Central African Republic	Cross-sectional comparative (cut vs non-cut). Subsample of DHS 1994-95	Sexual. There was a significant difference in frequency of sexual intercourse (favouring cut).
Thabet, 2003 (45)	Low	N=147, Egypt (Cairo)	Case-control (FGM/C type I vs type II&III vs any type with cysts vs non-cut). Clinical-/hospital-based	Sexual outcomes. There were significant differences in desire/arousal, satisfaction (favouring non-cut).

Legend: Cut= Have FGM/C; non-cut= Do not have FGM/C. *= study-level descriptive data showed differences but no statistical analyses were conducted to show whether the differences were statistically significant.

Study design and sample recruitment

According to the study descriptions, 15 included studies employed a cross-sectional comparative study design in which two or more groups of women were compared. Two studies were described as case-control (45;63). No systematic reviews or cohort studies were identified for inclusion. All studies were based on a non-random sample, with the possible exception of the study by Stewart and colleagues (75). This study was based on a subsample (37.2% of original) of women (currently married or in a union where both partners were residing together) from the 1994-95 Demographics and Health Survey (DHS), which is a random household sample of the larger population. This study design appeared to meet the characteristics for being representative. The other studies, on the other hand, were not. The majority of the studies were clinical-/hospital-based. Five studies recruited women attending either a general- or a specialist hospital (62;63;66;67;74), and another six studies recruited from family practitioners, family planning centres, antenatal clinics, and gynaecological clinics (16;45;65;68;71;72). A few of these studies attempted to minimize sampling bias, for example by selecting every third woman (patient or woman accompanying patient) visiting sites providing services for women (67). Shandall (74) included all female patients attending the Referred Outpatient Clinic in Khartoum General Hospital (gynaecological and obstetric) between 1962 and 1966. Two studies were community-based: Morison and colleagues (69) randomly selected 20 villages (among 40 in the Farafenni area) for inclusion in the study, and all women aged 15-54 were eligible. Nwajei and Otiono (70), who aimed to study sexuality among female students at Delta State University, created four sampling clusters from which four hundred women were selected. Lastly, three studies used what we term 'purposive sampling': they recruited participants through communities and socioeconomic sectors (9), nurses, private contacts, and study participants once enrolment started (64); significant others, contact persons, and snowballing techniques (73).

Fifteen of the studies compared women with FGM/C versus women without FGM/C (9;16;62-69;71;72;75). That is, different types of genital cutting were combined to make a binary variable of women with FGM/C versus women without FGM/C. Shandall (74) compared three groups of women: women with FGM/C type I, women with type III, and women without any genital cutting. Lastly, Thabet and Thabet (45) compared four groups of women: women with FGM/C type I ("minorly circumcised"), women with type II or III ("circumcised mutilated"), women with any type of genital cutting who developed clitoral cysts as late complications of circumcision, and women without any genital cutting ("uncircumcised").

Population

Two studies included women residing in the Middle East: One study enrolled 37 Bedouin women from south Israel (16) and the other study 260 women residing in Saudi Arabia (63). The remaining 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). These study participants were spread across seven

different African countries: Central African Republic, Egypt, Gambia, Ghana, Nigeria, Senegal, and Sudan. Five of the studies were from different cities in Egypt, and five studies were from Nigeria.

The sample sizes ranged from 37 to 4,024 women (mean sample size was 750). The studies compared women with FGM/C to women without FGM/C, and there were more women in the former group (n= 8,934, 70%). Two samples consisted of pregnant women (62;72). Across the studies, the women's ages ranged from 15-77 with a mean age of 33. Ten studies stated the educational level of the participants, which was generally low. The majority of women had no education beyond high school and illiteracy was common. In terms of employment, this was reported in five studies, and on average 35% of the women were employed (range 0% - 56.1%). The remaining women were generally considered housewives. The participants' religious affiliation was only reported in four studies (9;16;73;75), but included Catholics, Muslims, Protestants, as well as women who practiced traditional religions. Most studies reported some information about the women's civil status, which in the overwhelming majority of studies was married, and in six studies all women were married (45;62;65-67;73). The subsample from the DHS study (75) was restricted to those women currently married or in a union where the partners resided together.

With respect to type of FGM/C, almost all the studies (except four: (9;16;66;73)) provided some information about type of FGM/C, age at which FGM/C had been performed, and/or complications. The great majority of the women had been subjected to FGM/C type I or II. However, three quarters of the Sudanese women (74) and 42.3% of the residents of Saudi Arabia (63) had FGM/C type III. With the exception of 11 women in one study (72), no studies included women with FGM/C type IV. In the majority of cases, the women had been subjected to FGM/C in early childhood (mean age 8.5), but some as late as adulthood during their first pregnancy. Six studies reported physical complications among women with FGM/C, the most common of which were scarring and infections.

In nine of the 17 studies, the women were examined gynaecologically (45;62;65;67-69;71;72;74). This was done to confirm whether or not they had been genitally cut, and in some studies, to which type of FGM/C the women had been subjected. Physical examinations were not undertaken to verify the cutting statements in the remaining eight studies (9;16;63;64;66;70;73;75) and instead relied on women's self-report.

Table 3: Description of the population in included studies

Author, year	Population
Adinma, 1997	N=256 (124 cut, 132 non-cut) married, pregnant Igbo women from Nigeria. Age 16-40 (91.4% 21-35 years). Cutting: 21.8% type I, 78.2% type II; done in childhood (96.8%) or as teenager (3.2%). Complications: 37.9% moderate scarring, 23.4% severe scarring.
Alsibiani, 2010	N=260 (130 cut, 130 non-cut), sexually active women. 41.5% were Saudi, 58.5% residents of Saudi Arabia from neighbouring countries. Mean age 30.5 (16-40). 91% high school. Cutting: 40.8% type I /II, 42.3% type III, 16.9%

unknown.

Applebaum, 2008	N=37 (19 cut, 18 non-cut) Bedouin women from south Israel. Mean age 49.5 (SD=12.5) (31-77). 64% married, 5.3% single. 59.5% no education, 27.0% primary education, 8.1% college/university.100% Muslim. 0% employed. Cutting: "circumcised." [1]
Badawi, 1989	N=159 (26 cut, 133 non-cut) women from Egypt. 76% were 20-39 years old. 37% married, 59% single. 2% illiterate, 8% elementary school, 52% high school, 27% college. 94% Muslim, 6% Coptic. 46% employed, 27% housewife, 27% student. Cutting: "Genitally mutilated women."
Behrendt, 2005	N=47 (23 cut, 24 non-cut) women from Senegal. Mean age 22.9 (15-40). 21% married, 79% single. Mean years of education 11.5. Cutting: "circumcised" done at mean age 8.2 (5-14), without local analgesics or narcotics.
El-defrawi, 2001	N=250 (200 cut, 50 non-cut) married women from Egypt (Ismailia), from lower and middle social classes. Cutting: 50% injury to clitoris, 36.5% excision of clitoris, 13% excision of clitoris and surrounding parts, done at mean age 10.8 (8-12). Operator: 48.5% daya, 36.5% physician, 8.5% barber. Complications: 53.5% (pain, bleeding, infection, swelling immediately after procedure).
Elnashar, 2007a	N=264 (200 cut, 64 non-cut) newly married (<5 years) women from Egypt (Benha city). 70.5% (semi)urban residency. 6.8% illiterate or elementary school. 56.1% working, 43.9% housewife. Cutting: "circumcised".
Elnashar, 2007b	N=936 (845 cut, 91 non-cut) women from Lower Egypt (Dakahlia governorate). Mean age 29.9 (16-49). 11.5% polygamous marriage, 88.4% monogamous marriage. 28.6% illiterate, 14.4% primary education, 48% secondary education. 70.4% housewife, 18.7% working professional. 40.6% urban residency. Cutting: 62% "first- or second-degree circumcision".
Megafu, 1983	N=500 (340 cut, 160 non-cut) Ibo (aka Igbo) women from Nigeria. 16-45 years old. 76% married with children. Cutting: 78.8% type I, 21.2% type II. Done within first 8 days of life or near puberty. Performed outside medical institution by unqualified personnel (elderly man or woman in the vicinity) with sharp knife.
Morison, 2001	N=1,157 (668 cut, 489 non-cut) women from Gambia (Farafenni) of ethnic groups Mandinka, Wollof, Fula. Age 15-54. 51.7% in polygamous marriage, 6.7% single. <6% primary or secondary education and 95% said farming or working in household was main occupation. Cutting: 0.5% type I, 98% type II, 1.5% type III, done at mean age 6.1 by traditional operator. Complications: women with FGM/C significantly higher prevalence of bacterial vaginosis and herpes simplex virus 2 than women without FGM/C.
Nwajei, 2003	N=400 (120 cut, 280 non-cut) women students from Delta State University, Abraka, Nigeria. Ethnicity mainly Urhobo. 80% 16-21 years. Cutting: "circumcision (type I) done in this part of Nigeria. Only the clitoral hood is touched, largely leaving the clitoris intact." Age at cutting: "some girls undergo circumcision during their first pregnancy after marriage as required by the Urhobo culture." Genitourinary complaints were similar for both groups of women (itching, offensive odours or discharge).
Odoi , 1997	N=195 (76 cut, 119 non-cut) women from three different tribes in northern Ghana who attended an antenatal clinic. Cutting: 100% type I or II, age of cutting from early childhood to 18 years. Complications: 34% (haemorrhage, lacerations during first vaginal delivery).
Okonofua, 2002	N=1,836 (827 cut, 1,009 non-cut) pregnant women from Nigeria (Edo state). 93.2% married. Cutting: 71.4% type I, 24.4% type II, 2.9% type III, 1.3% type IV. Complications: 43.1% (scar formation, cysts, disfigurement, narrowed introitus).
Osinowo, 2003	N=99 (53 cut, 46 non-cut) married women from Nigeria (Lagos). Mean age 34.1 (18-60). 41.1% primary educ., 30.3% secondary educ. 20.2% housewife, 18.2% trader, 24.2% business, 11.1% civil servant. 48.5% Muslim, 42.4% Christian, 9.1% traditional religion. 55% of cut women polygamous marriage. Cutting: "circumcised".
Shandall, 1967	N=4,024 (807 type I, 3013 type III, 204 non-cut) women from Sudan (Khartoum area). Complications: ~20% (haemorrhage, shock, infection, scarring, Keloid formation, cysts, bacteriuria, chronic pelvic infection)
Stewart, 2002	N=2,188 (1346 cut, 842 non-cut) women from Central African Republic. Subsample of 1994-95 DHS survey. Age 15-49 (68.6% 20-39 years). Ethnicity: 37% Gbaya, 39% Banda, 10% Mandjia. All women either married or cohabitating with partner, 27.6% in polygamous marriage. 55.2% protestant, 33.1% Catholic, 11.7% Muslim. 62.2% no education, 30.6% primary education, 7.2% secondary education. Residency: 66% rural. Cutting: 100% type II ("clitoris and labia minora are partially or completely removed"), 54.6% cut >5 years prior to marriage, 45.4% cut within 5 years of marriage.
Thabet, 2003	N=147 (30 type I, 30 type II&III, 57 any type with clitoral cysts, 30 non-cut) married women from Egypt (Cairo). Age 18-28. Cutting: done at mean age 9.0 (7-12). Operator: 66.7% non-professional and paramedical, 26.5% trained gynaecologist.

Legend: Cut= Have FGM/C; non-cut= Do not have FGM/C; DHS= Demographic and Health Survey; [1]= "circumcision" here likely refers to FGM/C type II, because the most common form of FGM/C among Bedouin-Arabs involves the removal of part or all of the clitoris plus about 1 centimeter of the labia minora (10).

Outcomes

All outcomes regarding psychological, social and sexual functioning were measured through self-report. To collect outcome information about psychological, social and sexual functioning, the researchers either interviewed the women face-to-face (structured interview) or gave them a paper-and-pencil questionnaire to self-complete (self-report questionnaire). The most frequently measured outcome was sexual functioning (Table 4). All of the studies, except two (16;64), measured sexual functioning, four of the studies measured a psychological outcome (16;64;66;73), and two studies included a measure about social functioning (66;73). With respect to time-line, the majority of participating women had been cut in early childhood (mean age 8.5), and were presently in their mid thirties (mean age 33) when they were asked about current psychological, social and sexual functioning. Henceforth, for the great majority of women with FGM/C the event occurred several decades in the past and all consequences must be considered long term.

Table 4: Description of the outcomes in included studies

Author, year	Outcomes
Adinma, 1997	Sexual: Sexual satisfaction. Measurement: Single item.
Alsibiani, 2010	Sexual: Frequency of sexual intercourse, Female Sexual Function Index (FSFI) which includes 6 domains (libido, arousal, lubrication, orgasm, satisfaction, pain). Instrument: Female Sexual Functioning Index.
Applebaum, 2008	Psychological: Post traumatic stress syndrome, general psychiatric symptom severity (SCL-90), impact of event, quality of life. Instrument: PTSD Inventory, Symptom Checklist-90.
Badawi, 1989	Sexual: Sexual excitement from clitoral area stimulation, orgasm from manual stimulation of clitoris. Measurement: Single items.
Behrendt, 2005	Psychological: Post traumatic stress syndrome, other anxiety disorders, affective disorders, psychiatric diagnosis. Instrument: Mini International Neuropsychiatric Interview.
El-defrawi, 2001	Sexual: Initiation of sex, sexual desire, enjoyment of sex, orgasm, foreplay, dryness during intercourse, dyspareunia. Instrument: Sexual Behavior Assessment Schedule-Adult.
Elnashar, 2007a	Psychological: Somatisation, anxiety, phobia, depression, hostility. Instrument: Symptom Checklist-90. Social: Marital satisfaction. Measurement: Single item. Sexual: Dyspareunia, libido. Measurement: Single items.
Elnashar, 2007b	Sexual: Sexual problems. Measurement: Single items.
Megafu, 1983	Sexual: Orgasm, most erotic organ of the body. Measurement: Single items.
Morison, 2001	Sexual: Painful sex. Measurement: Single item.
Nwajei, 2003	Sexual: Frequency of sexual intercourse, most sensitive part of the body. Measurement: Single items.
Odoi, 1997	Sexual: Orgasm, dyspareunia, post-coital bleeding. Measurement: Single items.
Okonofua, 2002	Sexual: Sexual activity, arousal, initiation of sex, orgasm, dyspareunia, most sensitive bodily part. Measurement: Single items.
Osinowo, 2003	Psychological: Self-esteem. Measurement: Single item. Social: Marital instability. Measurement: Single item. Sexual: Sexual functioning. Instrument: Golombok-Rust Inventory of Sexual Satisfaction.
Shandall, 1967	Sexual: Orgasm. Measurement: Single item.
Stewart, 2002	Sexual: Frequency of sexual intercourse. Instrument: Single item from Demographic and Health Survey.
Thabet, 2003	Sexual: Sexual desire/arousal, orgasm, sexual satisfaction. Instrument: Kasr El Aini sexual assessment questionnaire sheet.

Legend: Dyspareunia= pain during intercourse; PTSD= Post traumatic stress disorder.

Measurement

All studies, except one (71), measured exposure and outcome the same way in the two groups. The four studies that measured a psychological outcome (16;64;66;73) focused on Post Traumatic Stress Disorder (PTSD) and general psychiatric symptoms, such as depression, anxiety and phobia. Applebaum and colleagues (16) measured PTSD with the PTSD Inventory, while Behrendt and Moritz (64) applied the Mini International Neuropsychiatric Interview, which is a short, structured diagnostic interview that includes measurement of PTSD in addition to a host of other mental and addictive disorders. Two studies (16;66), used the frequently-used and validated Symptom Checklist-90 (SCL-90), which comprises 90 items measuring several clinical subscales, to assess general psychiatric symptom severity. They computed the data differently, however, and presented the results on different scales.

With respect to measurement of sexual functioning, few studies applied a validated instrument, and none of the studies applied the same instrument. One study (63) used the Female Sexual Functioning Index (FSFI), which is a validated, multidimensional tool for assessment of six domains of sexual functioning (libido, arousal, lubrication, orgasm, satisfaction, dyspareunia). El-dafrawi and colleagues (65) applied the validated Sexual Behavior Assessment Schedule-Adult (SEBAS-A), which measures many of the same domains as the FSFI. The Golombok-Rust Inventory of Sexual Satisfaction (GRISS), which was applied in one study (73), is a widely-used inventory that measures overall sexual functioning in adults. The authors of one of the studies from Egypt (45) created what they referred to as the 'Kasr El Aini sexual assessment questionnaire sheet', which provided a sexual functioning score. The assessment sheet had good construct, discriminate and concurrent validity, as well as test-retest reliability. The remaining studies that measured sexual functioning (9;62;66-72;74;75) generally asked the women to respond 'yes' or 'no' to several questions regarding their current sexual functioning and satisfaction. It appears that these studies asked rather similarly worded questions, such as "Do you reach orgasm during sexual intercourse?" Three studies (69;72;75) controlled for confounders, such as age, parity and ethnic group.

QUALITY ASSESSMENT

We arrived upon a final decision of high study quality for two (69;75) of the 17 studies, five studies were of moderate quality (16;63-65;72), and the remaining ten studies were of low study quality (9;45;62;64;66-68;70;71;73;74), as judged by the 12 quality assessment questions of the NOKC modified checklist (Appendix 3). It was a strength that in all studies, except one where it was unclear (9), the authors explained that the non-exposed group was selected from the same population as the exposed group, and that many described that the groups were comparable with respect to important background factors. However, all of the studies, except two (69;75) failed to explain whether and how the participants who agreed to participate

with respect to important background factors. However, all of the studies, except two (69;75), failed to explain whether and how the participants who agreed to participate were different from those who refused to participate. All of the studies, except two (65;73), failed to show that the measures used were reliable and valid, and all of the studies, except two (65;72), failed to describe whether the person who assessed the outcome was blind to whether participants were exposed or not. Lastly, several of the studies also failed clearly to show that 1) the sample was representative of the population, 2) standardized data collection methods were used, and 3) statistical methods were appropriate.

With respect to grading of the evidence, evidence based on observational studies will generally be appreciably weaker than evidence from experimental studies with regards to establishing a causal relationship between exposure to an intervention and an outcome. Because all the studies were necessarily non-randomized and the majority also had methodological shortcomings, the evidence was deemed to be of low quality. For resource reasons we decided to assess the quality of the evidence through GRADE only for outcomes which were eligible for meta-analysis. We considered it not resource wise to conduct a formal GRADE assessment for the remaining outcomes of included studies. For each outcome eligible for meta-analysis, we applied six criteria and the final grading for each outcome was "very low", defined as "any estimate of effect is very uncertain" (Appendix 4).

CONSEQUENCES OF FGM/C

We extracted data about psychological, social and sexual consequences of FGM/C and structured the results according to these three functioning areas.

Psychological consequences of FGM/C

Four studies (16;64;66;73) of moderate and low study quality reported on psychological consequences. Two studies (16;64) provided prevalence data about PTSD (Table 5). The proportion of women who suffered from PTSD was not significantly different between the groups in either study. Notably, while one of them (16) found that about the same proportion of women in both groups (~10%) suffered from PTSD, the second study (64) found that about a third of women with FGM/C and none of the women without FGM/C suffered from PTSD. The latter study, which used the Mini International Neuropsychiatric Interview, also reported that a larger proportion of women who had been subjected to FGM/C than women without FGM/C had a psychiatric diagnosis and suffered from affective disorders and anxiety. Only the outcome 'psychiatric diagnosis' was significant, however, the number of events was generally low and confidence intervals wide.

Table 5: Study outcomes (dichotomous) and effect estimates for psychological consequences

Author, year	Study quality	Outcome	FGM/C group	Non-FGM/C group	Results RR (95%CI)
Applebaum, 2008	Moderate	PTSD	2/19 (10.5%)	2/18 (11.1%)	0.95 (0.15, 6.03)
Behrendt, 2005	Moderate	PTSD	7/23 (30.4%)	0/24 (0%)	15.63 (0.94, 258.85)
		psychiatric diagnosis	18/23 (78.3%)	1/24 (4.2%)	18.78 (2.72, 129.5)
		affective disorders	5/23 (21.7%)	1/24 (4.2%)	5.22 (0.66, 41.32)
		anxiety disorders	6/23 (26.2%)	0/24 (0%)	13.54 (0.81, 227.5)

Legend: PTSD= post traumatic stress disorder; RR= relative risk with 95% confidence interval (CI).

Three studies (16;66;73) reported continuous outcomes for psychological consequences (Table 6). At study level, of eight outcomes measured, only one significant difference between the two groups of women was found in the Applebaum and colleagues' study (16). There was a significant difference with respect to PTSD. Elnashar and colleagues (66) on the other hand, found that levels of anxiety (Mean diff= 9.2), somatization (Mean diff= 11.5) and phobia (Mean diff= 7.2) were significantly higher among women who had been subjected to FGM/C than among women who had not, and Osinowo and Taiwo (73) found that women with FGM/C had significantly lower self-esteem than women without FGM/C (Mean diff= -2.8).

Table 6: Study outcomes (continuous) and effect estimates for psychological consequences

Author, year	Study quality	Outcome	FGM/C group	Non-FGM/C group	Results Mean diff (95%CI)
Applebaum, 2008	Moderate	PTSD	57.3 (SD=9.5)	48.1 (SD=8.2)	9.2 (3.49, 14.91)
		anxiety	1.3 (SD=0.3)	1.4 (SD=0.3)	-0.1 (-0.29, 0.09)
		somatization	1.6 (SD=0.3)	1.6 (SD=0.3)	0.0 (-0.19, 0.19)
		phobic anxiety	0.8 (SD=0.3)	0.8 (SD=0.2)	0.0 (-0.16, 0.16)
		depression	1.6 (SD=0.4)	1.8 (SD=0.4)	-0.2 (-0.46, 0.06)
		hostility	0.7 (SD=0.1)	0.8 (SD=0.2)	-0.1 (-0.20, 0.00)
		impact of event	21.6 (SD=5.7)	18.2 (SD=4.9)	3.4 (-0.02, 6.82)
		quality of life	69.5 (SD=6.7)	73.3 (SD=6.6)	-3.8 (-8.09, 0.49)
Elnashar, 2007a	Low	anxiety	31.5 (SD=6.7)	22.3 (SD=3.8)	9.2 (7.31, 11.09)
		somatization	33.2 (SD=10.4)	21.7 (SD=6.9)	11.5 (9.28, 13.72)
		phobia	22.5 (SD=5.4)	15.3 (SD=4.7)	7.2 (5.45, 8.95)
		depression	30.1 (SD=7.3)	29.9 (SD=4.6)	0.2 (-1.91, 2.31)
		hostility	18.7 (SD=7.3)	16.9 (SD=6.2)	1.8 (-0.55, 4.15)
Osinowo, 2003	Low	self-esteem	27.7 (SD=5.7)	30.5 (SD=5.3)	-2.8 (-4.97, -0.63)

Legend: PTSD= Post traumatic stress disorder; Mean diff= Mean difference with 95% confidence interval (CI); SD= Standard deviation.

Results of meta-analyses for psychological consequences

Two studies (16;66) included general psychiatric symptom severity outcomes (measured by SCL-90) and were sufficiently similar to warrant pooling of effect sizes in meta-analysis. In total, we could conduct meta-analyses for four psychological outcomes: anxiety, somatisation, depression, hostility. One study (66), with women from Egypt, was of low study quality and the other (16), which was from Israel, was of moderate study quality. The quality of the evidence for these four outcomes was judged as 'very low' in GRADE. Briefly, no significant effect for any of the four psychological outcomes was found and substantial heterogeneity showed inconsistency across the two studies. We detail the results of the meta-analyses for these four outcomes below.

We conducted meta-analysis of the outcome 'anxiety'. As shown in the forest plot (Figure 3), no significant effect for anxiety was found (Std.mean diff= 0.69, 95%CI= -1.27, 2.66) and the findings were inconsistent across the two studies (I²= 96%, Chi²= 26.38, p= 0.00001).

Figure 3: Forest plot, anxiety

Study or Subgroup	Cut Mean	SD	Total	Non-cut Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Applebaum 2008	1.3	0.3	19	1.4	0.3	18	49.2%	-0.33 [-0.98, 0.32]
Elnashar 2007a	31.5	6.7	64	22.3	3.8	64	50.8%	1.68 [1.27, 2.08]
Total (95% CI)			83			82	100.0%	0.69 [-1.27, 2.66]

Heterogeneity: Tau² = 1.93; Chi² = 26.38, df = 1 (P < 0.00001); I² = 96%
Test for overall effect: Z = 0.69 (P = 0.49)
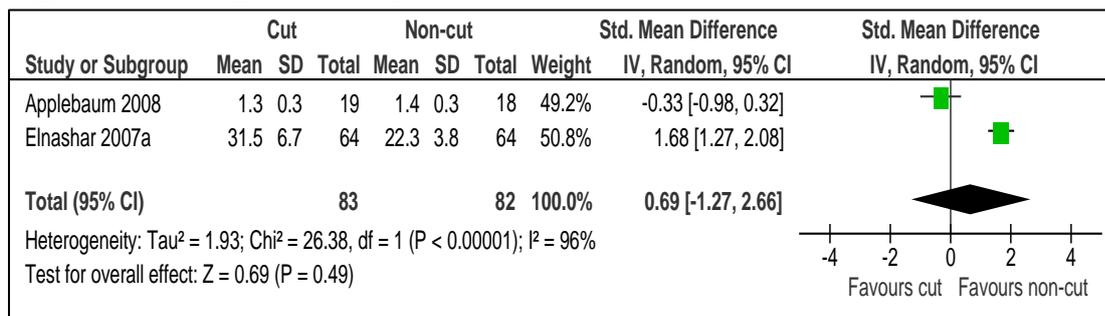
As shown in the forest plot (Figure 4), no significant effect for the outcome 'somatization' was found (Std.mean diff= 0.67, 95%CI= -0.59, 1.94). Considerable heterogeneity indicated by I² and Chi² (I²= 91%, Chi²= 11.48, p= 0.0007) showed inconsistency across studies.

Figure 4: Forest plot, somatization

Study or Subgroup	Cut Mean	SD	Total	Non-cut Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Applebaum 2008	1.6	0.3	19	1.6	0.3	18	47.9%	0.00 [-0.64, 0.64]
Elnashar 2007a	33.2	10.4	64	21.7	6.9	64	52.1%	1.30 [0.91, 1.68]
Total (95% CI)			83			82	100.0%	0.67 [-0.59, 1.94]

Heterogeneity: Tau² = 0.77; Chi² = 11.48, df = 1 (P = 0.0007); I² = 91%
Test for overall effect: Z = 1.04 (P = 0.30)
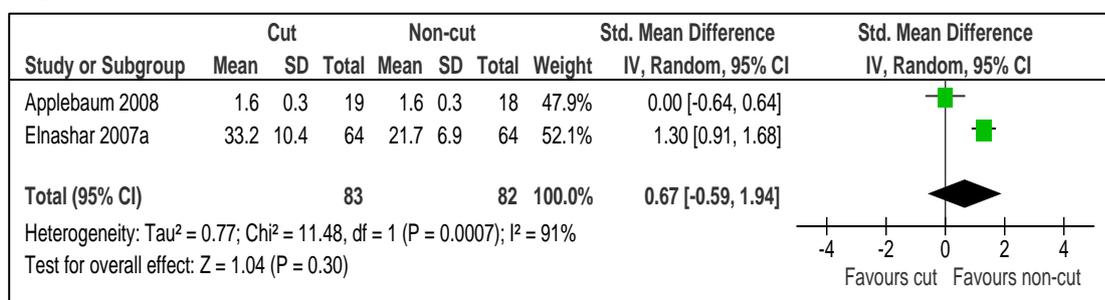
There was no significant effect for the outcome 'depression' (Std.mean diff= -0.15, 95%CI= -0.64, 0.34) (Figure 5), and the findings were inconsistent across the two studies (I²= 47% Chi²= 1.90, p= 0.17).

Figure 5: Forest plot, depression

Study or Subgroup	Cut Mean	SD	Total	Non-cut Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Applebaum 2008	1.6	0.4	19	1.8	0.4	18	35.2%	-0.49 [-1.14, 0.17]
Elnashar 2007a	30.1	7.3	64	29.9	4.6	64	64.8%	0.03 [-0.31, 0.38]
Total (95% CI)			83			82	100.0%	-0.15 [-0.64, 0.34]

Heterogeneity: Tau² = 0.06; Chi² = 1.90, df = 1 (P = 0.17); I² = 47%
Test for overall effect: Z = 0.61 (P = 0.54)

Std. Mean Difference IV, Random, 95% CI
-4 -2 0 2 4
Favours cut   Favours non-cut

Finally, there was no significant effect for the outcome 'hostility' (Std.mean diff= -0.13, 95%CI= -1.00, 0.73) (Figure 6), and considerable heterogeneity indicated by I² and Chi² (I²= 82% Chi²= 5.42, p= 0.02) showed inconsistency across studies.

Figure 6: Forest plot, hostility

Study or Subgroup	Cut Mean	SD	Total	Non-cut Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Applebaum 2008	0.7	0.1	19	0.8	0.2	18	44.8%	-0.62 [-1.29, 0.04]
Elnashar 2007a	18.7	7.3	64	16.9	6.2	64	55.2%	0.26 [-0.08, 0.61]
Total (95% CI)			83			82	100.0%	-0.13 [-1.00, 0.73]

Heterogeneity: Tau² = 0.32; Chi² = 5.42, df = 1 (P = 0.02); I² = 82%
Test for overall effect: Z = 0.30 (P = 0.76)

Std. Mean Difference IV, Random, 95% CI
-4 -2 0 2 4
Favours cut   Favours non-cut
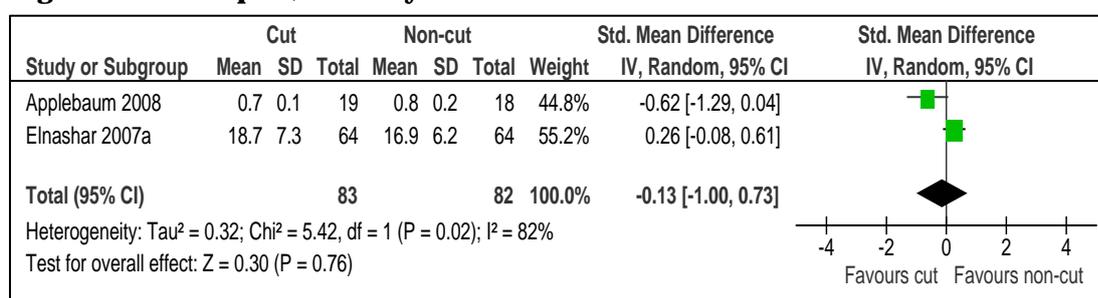
Social consequences of FGM/C

Two studies (66;73), both of low study quality, included some measure of social consequences of FGM/C (Table 7). The proportion of women who reported they were not satisfied with their marriage was significantly greater among women who had been subjected to FGM/C (43.0%) than among women who had not (10.9%) (RR=3.93, 95%CI= 1.92, 8.05). Marital instability was also significantly different between the two groups; it was higher among women with FGM/C (Mean diff= 1.7).

Table 7: Study outcomes and effect estimates for social consequences

Author, year	Study quality	Outcome	FGM/C group	Non-FGM/C group	Results (95%CI)
Elnashar, 2007a	Low	marital satisfaction ('no')	86/200 (43.0%)	7/64 (10.9%)	RR= 3.93 (1.92, 8.05)
Osinowo, 2003	Low	marital instability	2.4 (SD=1.9)	0.7 (SD=1.2)	Mean diff= 1.7 (1.08, 2.32)

Legend: RR= relative risk with 95% confidence interval (CI); Mean diff= Mean difference; SD= Standard deviation.

Sexual consequences of FGM/C

Fifteen studies (9;45;62;63;65-75), of varying study quality, reported a total of 53 dichotomous outcomes for sexual consequences of FGM/C, the majority (58.5%) of which were statistically associated with FGM/C status at study level (Table 8). In several studies, a significantly greater proportion of women who had been subjected to FGM/C than women who had not, reported negative sexual experiences, includ-

ing dryness during intercourse, pain during intercourse, not enjoying sex, and never or rarely experiencing orgasm. In opposition to these results, one study (72) reported that a significantly smaller proportion of women with FGM/C than women without FGM/C reported never initiating sexual activity, not experiencing orgasm, and not engaging in sexual intercourse in the last week or in the last month. This study included pregnant women with FGM/C type I as well as type II and found that there were no significant differences between these two groups. In summary, all 15 studies except one suggested that there were negative consequences from FGM/C on women's sexual functioning.

Table 8: Study outcomes (dichotomous) and effect estimates for sexual consequences

Author, year	Study quality	Outcome	FGM/C group	Non-FGM/C group	Results RR (95%CI)
Adinma, 1997	Low	sexually satisfied ('no')	21/124 (16.9%)	9/132 (6.8%)	2.48 (1.18, 5.21)
		satisfied ('very rarely')	7/124 (5.6%)	0/132 (0%)	15.96 (0.92, 276.54)
		satisfied ('sometimes')	42/124 (33.9%)	40/132 (30.3%)	1.12 (0.78, 1.6)
		satisfied ('yes')	11/124 (8.9%)	31/132 (23.5%)	0.38 (0.20, 0.72)
		satisfied ('very often')	43/124 (34.7%)	52/132 (39.4%)	0.88 (0.64, 1.21)
Alsibiani, 2010	Moderate	intercourse ≤3 times per wk	117/130 (90%)	120/130 (92.3%)	0.97 (0.90, 1.05)
Badawi, 1989	Low	no orgasm [1]	20/26 (75.0%)	67/133 (50.0%)	1.53 (1.17, 2.00)
El-dafrawi, 2001	Moderate	no sexual desire	83/200 (41.5%)	8/50 (16.0%)	2.59 (1.35, 5.00)
		no enjoyment of sex	88/200 (44.0%)	7/50 (14.0%)	3.14 (1.55, 6.36)
		no foreplay	97/200 (48.5%)	28/50 (56.0%)	0.87 (0.65, 1.15)
		never initiate sex	76/200 (38.0%)	16/50 (32.0%)	1.19 (0.76, 1.85)
		dryness during intercourse	97/200 (48.5%)	15/50 (30.0%)	1.62 (1.03, 2.53)
		dyspareunia	92/200 (46.0%)	16/50 (32.0%)	1.44 (0.93, 2.21)
		no orgasm	86/200 (43.0%)	9/50 (18.0%)	2.39 (1.29, 4.41)
Elnashar, 2007a	Low	dyspareunia	81/200 (40.5%)	12/64 (18.8%)	2.16 (1.26, 3.70)
		no libido	57/200 (28.5%)	10/64 (15.6%)	1.38 (0.99, 3.36)
Elnashar, 2007b	Low	sexual problems	587/845 (69.5%)	58/91 (63.7%)	1.09 (0.93, 1.28)
Megafu 1983	Low	no orgasm	140/340 (41.2%)	50/160 (31.2%)	1.32 (1.01, 1.71)
		clitoris most erotic organ	80/340 (23.5%)	70/160 (43.8%)	0.54 (0.41, 0.70)
		breasts most erotic organ	80/340 (23.5%)	30/160 (18.8%)	1.25 (0.86, 1.83)
		lips most erotic organ	20/340 (5.9%)	10/160 (6.3%)	0.94 (0.45, 1.96)
		labia minora most erotic organ	120/340 (35.3%)	30/160 (18.8%)	1.88 (1.32, 2.68)
Morison, 2001	High	painful sex	62/394 (15.7%)	47/329 (14.3%)	1.10 (0.78, 1.56)
Nwajei, 2003	Low	no intercourse ever	15/120 (12.0%)	6/280 (2.0%)	5.83 (2.32, 14.67)
		intercourse once in 6 mo	9/120 (8.0%)	0/280 (0%)	44.12 (2.59, 752.06)

		intercourse once in 3 mo	15/120 (12.0%)	0/280 (0%)	71.99 (4.34, 1193.49)
		intercourse once a month	19/120 (16.0%)	64/280 (23.0%)	1.59 (0.93, 2.71)
		intercourse 1-2 times per wk	24/120 (20.0%)	120/280 (43.0%)	0.47 (0.32, 0.68)
		intercourse ≥3 times per wk	38/120 (32.0%)	90/280 (32.0%)	0.99 (0.72, 1.35)
		clitoris most sensitive organ	49/120 (41%)	143/280 (51%)	0.80 (0.63, 1.02)
		vagina most sensitive organ	25/120 (21%)	11/280 (4%)	5.30 (2.70, 10.43)
		tongue most sensitive organ	21/120 (17%)	25/280 (9%)	1.96 (1.14, 3.36)
		breasts most sensitive organ	25/120 (21%)	101/280 (36%)	0.58 (0.39, 0.85)
Odoi, 1997	Low	no orgasm	9/76 (11.8%)	0/119 (0%)	29.61 (1.75, 501.45)
		dyspareunia	10/76 (13.2%)	6/119 (5.0%)	2.61 (0.99, 6.89)
		post-coital bleeding	4/76 (5.3%)	0/119 (0%)	14.03 (0.77, 256.88)
Okonofua, 2002	Moderate	no sexual activity in last wk	364/827 (44.0%)	535/1009 (53.0%)	0.83 (0.75, 0.91)
		no sexual activity last mo	158/827 (19.1%)	289/1009 (28.6%)	0.67 (0.56, 0.79)
		not easily aroused	558/827 (67.5%)	656/1009 (65.0%)	1.04 (0.97, 1.11)
		never initiate sex	345/827 (41.7%)	477/1009 (47.3%)	0.88 (0.80, 0.98)
		no orgasm	280/827 (33.9%)	410/1009 (40.6%)	0.83 (0.74, 0.94)
		pain during intercourse	30/825 (3.6%)	23/1003 (2.3%)	1.59 (0.93, 2.71)
		clitoris most sensitive part	87/827 (10.5%)	276/1009 (27.4%)	0.38 (0.31, 0.48)
		breasts most sensitive part	523/827 (63.2%)	439/1009 (43.5%)	1.45 (1.33, 1.59)
		other part most sensitive	217/827 (26.2%)	294/1009 (29.1%)	0.90 (0.78, 1.05)
Shandall, 1967	Low	never had orgasm	TI 98/807 (12.1%) TIII 2520/3013 (83.6%)	14/204 (6.9%)	1.77 (1.03, 3.03) 12.19 (7.35, 20.21)
		orgasm <50% of occasions	TI 139/807 (17.2%) TIII 162/3013 (5.4%)	29/204 (14.2%)	1.21 (0.84, 1.75) 0.38 (0.26, 0.55)
		orgasm 50-75% of occasions	TI 300/807 (37.2%) TIII 180/3013 (6.0%)	61/204 (29.9%)	1.24 (0.99, 1.56) 0.20 (0.16, 0.26)
		orgasm >75% of occasions	TI 270/807(33.5%) TIII 151/3013 (5.0%)	100/204 (49.0%)	0.68 (0.58, 0.81) 0.10 (0.08, 0.13)
Stewart, 2002	High	intercourse ≥5 times last mo	633/1346 (47%)	328 /842 (39%)	1.21 (1.09, 1.34)

Legend: Cyst= Any FGM/C + cyst; FSFI= Female Sexual Functioning Index; mo= month; TI= FGM/C Type I; TIII= FGM/C Type III; TII&III= FGM/C Type II&III; wk= week, 1= orgasm from manual stimulation of the clitoris/clitoral area.

Four studies (45;63;65;73), of moderate and low study quality, reported a total of 12 continuous outcomes for sexual consequences of FGM/C, the majority of which (74%) were statistically significant at study level (Table 9). In these studies, women with FGM/C reported experiencing significant less desire and arousal, less satisfaction, as well as experiencing orgasm less frequently.

Table 9: Study outcomes (continuous) and effect estimates for sexual conse-quences

Author, year	Study quality	Outcome	FGM/C group	Non-FGM/C group	Results Mean diff (95%CI)
Alsibiani, 2010	Moderate	desire	3.6 (SD=1.1)	3.7 (SD=1.2)	-0.1 (-0.38, 0.18)
		arousal	3.6 (SD=1.2)	4.2 (SD=1.4)	-0.6 (-0.92, -0.28)
		lubrication	3.4 (SD=1.0)	3.9 (SD=1.3)	-0.5 (-0.78, -0.22)
		orgasm	3.7 (SD=1.2)	4.2 (SD=1.4)	-0.5 (-0.82, -0.18)
		satisfaction	4.5 (SD=1.2)	5.0 (SD=1.4)	-0.8 (-1.12, -0.48)
		pain	3.5 (SD=1.0)	3.8 (SD=1.1)	-0.3 (-0.56, -0.04)
		FSFI score	21.4 (SD=4.4)	23.5 (SD=5.0)	-2.1 (-3.24, -0.96)
El-dafrawi, 2001	Moderate	sexual desire	2.8 (SD=1.7)	3.7 (SD=1.9)	-0.9 (-1.48, -0.32)
Osinowo, 2003	Low	sexual functioning	78.9 (SD=18.3)	96.4 (SD=10.1)	-17.5 (-23.23, -11.77)
Thabet, 2003	Low	desire & arousal	TI 9.4 (SD= 0.8) TII&III 7.2 (SD=0.8) Cyst 11.2 (SD=1.0)	10.4 (SD=0.8)	-1.0 (-1.4, -0.6) -3.2 (-3.6, -2.8) 0.8 (0.41, 1.19)
		satisfaction	TI 8.3 (SD=1.1) TII&III 7.3 (SD=0.8) TI-III 7.8 (SD=0.96)[1] Cysts 8.0 (SD=0.9)	8.0 (SD=0.9)	0.3 (-0.21, 0.81) -0.7 (-1.13, -0.27) -0.2 (-0.6, 0.2)[1] 0.0 (-0.4, 0.4)
		orgasm	TI 12.4 (SD=1.7) TII&III 6.0 (SD=0.9) Cysts 10.9 (SD=1.2)	12.5 (SD=1.1)	-0.1 (-0.82, 0.62) -6.5 (-7.01, -5.99) -1.6 (-2.1, -1.1)

Legend: Cyst= Any FGM/C + cyst; FSFI= Female Sexual Functioning Index; Std.mean diff= Standardized mean difference; TI= FGM/C type I; TIII= FGM/C type III; TII&III= FGM/C type II&III; TI-III= FGM/C type I&II&III combined as one group; 1= We combined the two groups FGM/C type I and type II&III.

Results of meta-analyses for sexual consequences

Some studies included the same outcome and were sufficiently similar to warrant pooling of effect sizes in meta-analyses. Altogether we could conduct meta-analyses for seven outcomes regarding sexual functioning: dyspareunia, satisfaction, desire, initiation of sex, orgasm, reporting clitoris as the most sensitive area of the body, and reporting the breasts as the most sensitive areas of the body.

Dyspareunia

We conducted meta-analysis of the outcome 'dyspareunia', which five studies (65;66;69;71;72) measured in a similar manner. One study was judged to have high study quality, two moderate, and two studies were of low methodological quality. The quality of the total body of evidence was 'very low', defined as "any estimate of effect is very uncertain" (Appendix 4). The women were from four different African countries and almost all had FGM/C type I or II. As shown in the forest plot (Figure 7), a significant effect for dyspareunia was found (RR= 1.52, 95%CI= 1.15, 2.0). Women who had been subjected to FGM/C were 1.5 times more likely to experience pain during intercourse than women who had not been subjected to FGM/C. To ex-

amine the possible influence of including pregnant women in the model, we removed the study with pregnant women (72), finding no change in effect size.

Figure 7: Forest plot, dyspareunia

Study or Subgroup	Cut Events	Total	Non-cut Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
El-defrawi 2001	92	200	16	50	24.6%	1.44 [0.93, 2.21]	
Elnashar 2007a	81	200	12	64	18.6%	2.16 [1.26, 3.70]	
Morison 2001	62	394	47	329	30.9%	1.10 [0.78, 1.56]	
Odoi 1997	10	76	6	119	7.2%	2.61 [0.99, 6.89]	
Okonofua 2002	30	825	23	1003	18.6%	1.59 [0.93, 2.71]	
Total (95% CI)		1695		1565	100.0%	1.52 [1.15, 2.00]	
Total events	275		104				
Heterogeneity: Tau² = 0.03; Chi² = 6.03, df = 4 (P = 0.20); I² = 34%							
Test for overall effect: Z = 2.95 (P = 0.003)							

0.2 0.5 1 2 5
Favours cut   Favours non-cut

Satisfaction

We conducted meta-analysis of the continuous outcome 'satisfaction'. Two studies (45;63) of moderate and low study quality measured level of satisfaction in a similar manner. Application of GRADE showed that the quality of the evidence was 'very low.' Recall that Thabet and Thabet (45) analyzed Egyptian women with FGM/C type I and type II&III separately while Alsibiani (63) treated women (resident in Saudi Arabia) of any category of cutting (I–III) as one group. In order to create comparable groups with respect to type of cutting, we combined the group of women with FGM/C type I with type II&III in the Thabet and Thabet study (see Table 9). As evident from the forest plot (Figure 8), a significant effect for satisfaction was found (St.mean diff= -0.34, 95%CI= -0.56, -0.13). Women who had been subjected to FGM/C experienced significantly less sexual satisfaction than women who had not been subjected to FGM/C. The finding in the meta-analysis was supported by results from two other studies, which were of moderate study quality. First, in response to Adinma's survey (62) more than twice as many women with FGM/C as women without FGM/C responded 'no' to the question of whether they were sexually satisfied during intercourse. Women with FGM/C were 2.48 times as likely to say they were not sexually satisfied (95%CI= 1.18, 5.21). Second, the study by El-dafrawi and colleagues (65) reported that women with FGM/C were 3.14 times as likely to say they had no enjoyment of sex (95%CI= 1.55, 6.36) compared to women without FGM/C.

Figure 8: Forest plot, satisfaction

| Study or Subgroup	Cut Mean	SD	Total	Non-cut Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
Alsibiani 2010	4.5	1.2	130	5	1.4	130	76.2%	-0.38 [-0.63, -0.14]	
Thabet 2003	7.8	0.96	60	8	0.9	30	23.8%	-0.21 [-0.65, 0.23]	
Total (95% CI)			190			160	100.0%	-0.34 [-0.56, -0.13]	
Heterogeneity: Tau² = 0.00; Chi² = 0.45, df = 1 (P = 0.50); I² = 0%									
Test for overall effect: Z = 3.13 (P = 0.002)									

-2 -1 0 1 2
Favours non-cut   Favours cut

No sexual desire

Two studies (65;66) that measured sexual desire were sufficiently similar to warrant pooling of effect sizes in meta-analysis. One study was judged to have moderate study quality and one low study quality. The quality of the evidence for 'desire' was judged as 'very low' in GRADE. The included women were from Egypt with largely FGM/C type I or II. Women with FGM/C were twice as likely to report that they experienced no desire compared to women without FGM/C (RR= 2.15, 95%CI= 1.37, 3.36) (Figure 9). The finding that women with FGM/C were less likely to say that they experienced sexual desire was also supported by continuous outcome results in other studies. Alsibiani and Rouzi (63) as well as Thabet and Thabet (45) found that women who had been subjected to FGM/C reported lower sexual desire and arousal scores (Mean diff= -0.6, 95%CI= -0.92, -0.28 and Mean diff= -3.2, 95%CI= -3.6, -2.8, respectively), and El-dafrawi and colleagues (65) similarly found that women with FGM/C self-reported a lower frequency of sexual desire than women without FGM/C (Mean diff=-0.9, 95%CI= -1.48, -0.32).

Figure 9: Forest plot, no sexual desire

Study or Subgroup	Cut Events	Cut Total	Non-cut Events	Non-cut Total	Weight	Risk Ratio M-H, Random, 95% CI
El-defrawi 2001	83	200	8	50	46.4%	2.59 [1.35, 5.00]
Elnashar 2007a	57	200	10	64	53.6%	1.82 [0.99, 3.36]
Total (95% CI)		400		114	100.0%	2.15 [1.37, 3.36]
Total events	140		18			

Heterogeneity: Tau² = 0.00; Chi² = 0.60, df = 1 (P = 0.44); I² = 0%
Test for overall effect: Z = 3.35 (P = 0.0008)
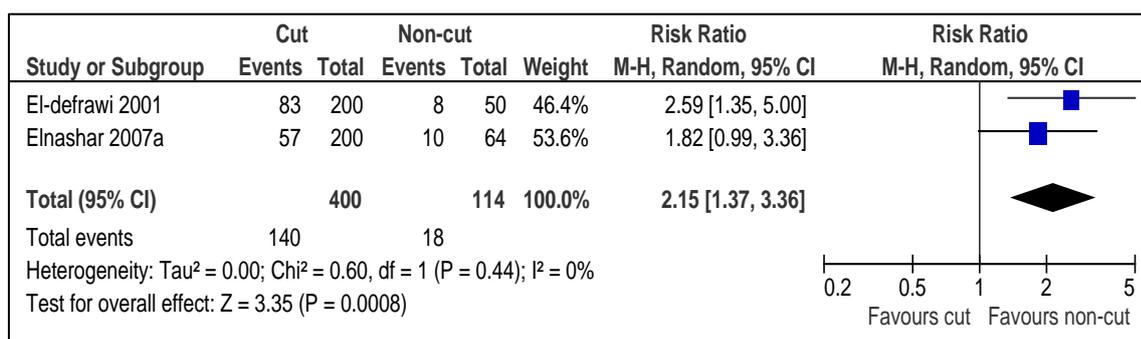
Never initiate sex

Two studies (65;72) that measured differences between women with FGM/C and women without FGM/C with respect to whether they initiated sex were sufficiently similar to warrant pooling of effect sizes in meta-analysis. Both were of moderate study quality and the quality of the evidence for this outcome was judged as 'very low' in GRADE. The women were from Egypt and Nigeria, and FGM/C type included I and II. No significant difference for the outcome 'never initiate sex' was found (RR= 0.94, 95%CI= 0.74, 1.21) (Figure 10).

Figure 10: Forest plot, never initiate sex

Study or Subgroup	Cut Events	Cut Total	Non-cut Events	Non-cut Total	Weight	Risk Ratio M-H, Random, 95% CI
El-defrawi 2001	76	200	16	50	22.9%	1.19 [0.76, 1.85]
Okonofua 2002	345	827	477	1009	77.1%	0.88 [0.80, 0.98]
Total (95% CI)		1027		1059	100.0%	0.94 [0.74, 1.21]
Total events	421		493			

Heterogeneity: Tau² = 0.02; Chi² = 1.65, df = 1 (P = 0.20); I² = 40%
Test for overall effect: Z = 0.46 (P = 0.65)
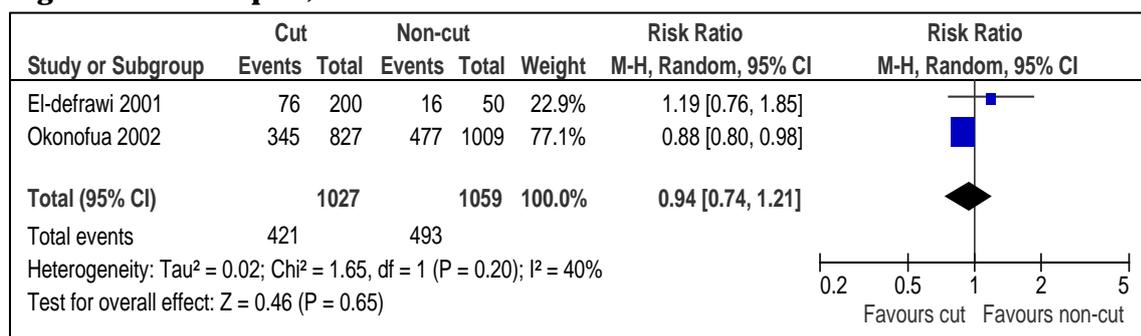
No orgasm

Five studies (65;68;71;72;74) measured whether women experienced orgasm and were sufficiently similar to warrant pooling of effect sizes in meta-analysis. Two studies were of moderate study quality and three of low study quality. In GRADE, the quality of the evidence for this outcome was judged as 'very low.' The women were from four different countries in Africa and FGM/C type included I-III. No significant effect for orgasm was found (RR= 1.5, 95%CI= 0.93, 2.44) (Figure 11). Considerable heterogeneity indicated by I^2 and Chi^2 (I^2= 87%, Chi^2= 31.2, p= 0.00001) showed inconsistency across studies. To examine the possible influence of including pregnant women in the model, we removed the study with pregnant women (72) from the analysis. The effect became significant (RR= 1.85, 95%CI= 1.12, 3.06), but heterogeneity remained high (I^2= 64%, Chi^2= 8.41, p= 0.04). Support for the finding that women with FGM/C may be more likely not to experience orgasm came from three studies with continuous estimates for orgasm (45;63;74). These three studies found that women with FGM/C experienced orgasm significantly less frequent than women who had not been subjected to FGM/C.

Figure 11: Forest plot, no orgasm
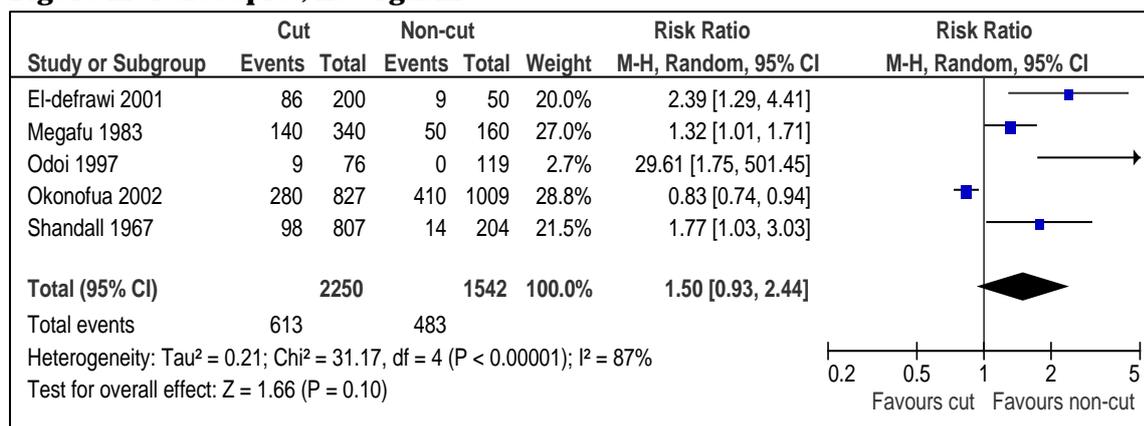
Study or Subgroup	Cut Events	Cut Total	Non-cut Events	Non-cut Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
El-defrawi 2001	86	200	9	50	20.0%	2.39 [1.29, 4.41]	
Megafu 1983	140	340	50	160	27.0%	1.32 [1.01, 1.71]	
Odoi 1997	9	76	0	119	2.7%	29.61 [1.75, 501.45]	
Okonofua 2002	280	827	410	1009	28.8%	0.83 [0.74, 0.94]	
Shandall 1967	98	807	14	204	21.5%	1.77 [1.03, 3.03]	
Total (95% CI)		2250		1542	100.0%	1.50 [0.93, 2.44]	
Total events	613		483				

Heterogeneity: Tau² = 0.21; Chi² = 31.17, df = 4 (P < 0.00001); I² = 87%
Test for overall effect: Z = 1.66 (P = 0.10)

0.2 0.5 1 2 5
Favours cut Favours non-cut

Clitoris most sensitive area of the body

Three studies (68;70;72) asked what women thought was the most sensitive part of the body, and included the answer alternative 'clitoris'. These studies were sufficiently similar to warrant pooling of effect sizes in meta-analysis. Two studies were of low study quality and one of moderate quality. The quality of the evidence for this outcome was judged as 'very low' in GRADE. The women were from Nigeria and had FGM/C type I or II. There was a significant effect (RR= 0.55, 95%CI= 0.35, 0.85) (Figure 12), suggesting that women with FGM/C were less likely to report that the clitoris was the most sensitive area of their body, but the findings were inconsistent across studies (statistical tests for heterogeneity: I^2= 90%, Chi^2= 19.78, p= 0.0001).

Figure 12: Forest plot, clitoris most sensitive area of the body
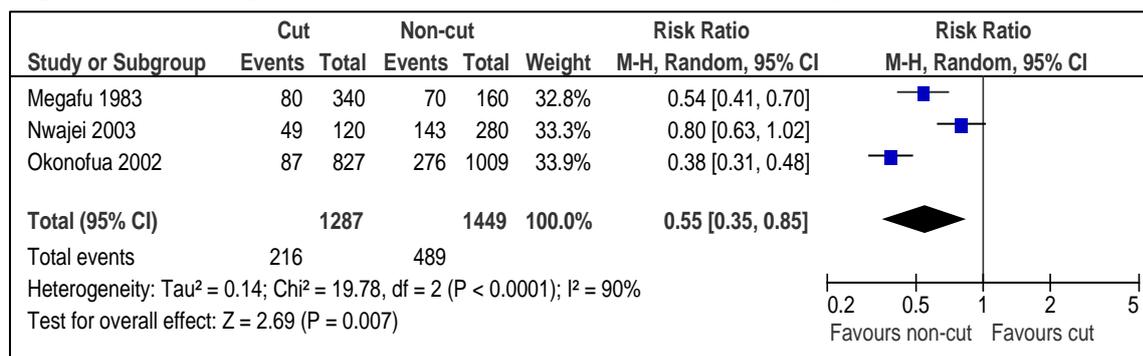
Study or Subgroup	Cut Events	Cut Total	Non-cut Events	Non-cut Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Megafu 1983	80	340	70	160	32.8%	0.54 [0.41, 0.70]	
Nwajei 2003	49	120	143	280	33.3%	0.80 [0.63, 1.02]	
Okonofua 2002	87	827	276	1009	33.9%	0.38 [0.31, 0.48]	
Total (95% CI)		1287		1449	100.0%	0.55 [0.35, 0.85]	
Total events	216		489				

Heterogeneity: Tau² = 0.14; Chi² = 19.78, df = 2 (P < 0.0001); I² = 90%
Test for overall effect: Z = 2.69 (P = 0.007)

0.2 0.5 1 2 5
Favours non-cut   Favours cut

Breasts most sensitive area of the body

Three studies (68;70;72) asked what women thought was the most sensitive part of the body, and included the answer alternative 'breasts'. These studies were sufficiently similar to warrant pooling of effect sizes in meta-analysis. Two studies were of low study quality and one of moderate quality. The quality of the evidence was judged as 'very low' in GRADE. The women were from Nigeria and had FGM/C type I or II. No significant effect was found (RR= 0.91, 95%CI= 0.62, 1.33) (Figure 13), and considerable heterogeneity indicated by I² and Chi² (I²= 97%, Chi²= 73.14, p= 0.00001) showed inconsistency across studies.

Figure 13: Forest plot, breasts most sensitive area of the body

Study or Subgroup	Cut Events	Cut Total	Non-cut Events	Non-cut Total	Weight	Risk Ratio (Non-event) M-H, Random, 95% CI	Risk Ratio (Non-event) M-H, Random, 95% CI
Megafu 1983	80	340	30	160	33.6%	0.94 [0.86, 1.03]	
Nwajei 2003	25	120	101	280	33.0%	1.24 [1.09, 1.41]	
Okonofua 2002	523	827	439	1009	33.4%	0.65 [0.59, 0.72]	
Total (95% CI)		1287		1449	100.0%	0.91 [0.62, 1.33]	
Total events	628		570				

Heterogeneity: Tau² = 0.11; Chi² = 73.14, df = 2 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.49 (P = 0.63)

0.5 0.7 1 1.5 2
Favours cut   Favours non-cut

Additional syntheses of results across studies

Three sexual outcomes – frequency of sexual activity, dryness during intercourse, overall sexual functioning – were measured differently in different studies and we could not conduct meta-analyses for these outcomes. We briefly summarize the results of these outcomes to show direction of consequences across studies (see also Tables 8-9).

Four studies included a question about frequency of sexual activity. According to the results by Alsibiani and Rouzi (63), less than 10% of both women who had and women who had not been subjected to the practice of FGM/C engaged in intercourse more than three times per week, but there was no significant difference between the groups. In contrast, Nwajei and Otiono (70) found that a significantly higher proportion of women with FGM/C had never had intercourse, had intercourse once in the last six months, and once in the last three months. Similarly, significantly fewer

women with FGM/C than women without FGM/C had intercourse one to two times per week. Another study with pregnant women (72) showed that a significantly higher proportion of women who had not been subjected to FGM/C reported no sexual activity in the last week (44% vs 53%) and in the last months (19% vs 29%). Lastly, Stewart, Morison and White (75) found that significantly more women with FGM/C had intercourse five or more times in the past four weeks. The authors state that the difference disappeared in the multivariate analysis. Across the four studies that measured frequency of sexual intercourse, no direction of results can be observed.

With respect to experiencing dryness during intercourse, two studies investigated the likelihood of experiencing dryness in women with and without FGM/C. The first study (65) found that a significantly greater proportion of women who had been subjected to FGM/C than women who had not been subjected to FGM/C reported dryness during intercourse (48.5% vs 30.0%, RR= 1.62, 95%CI=1.03, 2.53). Similarly, Alsibiani and Rouzi (63) found that women who had been exposed to the practice of FGM/C reported less lubrication (i.e., more dryness) than women who had not been subjected to the practice. Across these two studies, the results were in the same direction, showing that more women with FGM/C than without FGM/C experienced dryness during intercourse.

Lastly, three studies (63;65;73) used different scales to assess overall sexual functioning among women who had and women who had not been subjected to FGM/C, but the results were in the same direction. First, among the participants from Saudi Arabia (63), women who had been subjected to FGM/C scored significantly lower (2.1 points, possible range 2-36) on the Female Sexual Function Index. Second, according to results from the Golombok-Rust Inventory of Sexual Satisfaction (73), women with FGM/C scored significantly lower (17.5 points, possible range 56-280) on sexual functioning compared to women without FGM/C. Third, Elnashar and colleagues (67) found that sexual problems were more common among women with FGM/C than women without FGM/C (69.5% vs 63.7%), indicating, as the previous two studies, that there was a greater likelihood in women with FGM/C than without to experience reduced sexual functioning.

Discussion

This systematic review aimed to summarize published, empirical data assessing the psychological, social and sexual consequences of FGM/C. We included 17 studies, which all compared women with FGM/C type I-III to women without FGM/C. The methodological study quality was variable, with 10 of 17 of the studies judged as low quality, five as moderate quality, and two as high quality. Further, the quality of the total body of evidence for all outcomes was assessed as 'very low.' Little scientific research is available on the psychological and social consequences of FGM/C, but study level results suggest that having FGM/C may be associated with psychological disturbances. Our effect size estimates for dyspareunia, sexual satisfaction and sexual desire provide evidence in support of the argument that there is an association between FGM/C and women's reduced capacity for sexual enjoyment. However, the small- and low quality of the body of evidence preclude us from drawing conclusions regarding causality.

DISCUSSION OF MAIN RESULTS

Psychological consequences

Four studies (16;64;66;73) of moderate and low study quality reported on psychological consequences. According to study level results, it seems that women with FGM/C may be more likely than women without FGM/C to have a psychiatric diagnosis, and suffer from anxiety, somatisation, phobia, and low self-esteem. Although it appears that various psychological disturbances may be related to FGM/C, the results of the meta-analyses of anxiety, somatization, depression, and hostility failed to reach statistical significance and there was inconsistency across studies. Further, the small number of studies and low study quality preclude us from drawing firm conclusions.

The possibility that FGM/C leaves a lasting psychological mark is indicated in the literature (e.g., (39)). Behrendt and Moritz (64) concluded that more than 80% of the adult women in their study still suffered from intrusive reexperiences of being cut as girls. Comments like "I cried and screamed for help and no one helped" by women in the study by Badawi (9) was summarized: "The intensity of their recall was very strong and vivid and commonly associated with weeping and remembered pain and humiliation." At the same time, the fact that FGM/C is culturally embed-

ded may form a protective factor against the emergence of psychological stress in the aftermath of FGM/C. Future research should take the possible protective element of cultural embedment into account when examining the short term and long term psychological consequences of the practice.

Social consequences

There is a dearth of research on the social consequences of FGM/C. Of the 17 included studies in this systematic review, only two (66;73) addressed social consequences of FGM/C: marital satisfaction and stability. According to study level results, it is possible that women with FGM/C may be more likely than women without FGM/C to experience marital dissatisfaction and instability. However, both studies were of low methodological quality and we cannot draw any conclusion on whether FGM/C is linked to relationship disharmony and related problems. The interrelationship of marital- and sexual satisfaction is addressed in the literature (e.g., (47;76)) and likely relevant in the context of FGM/C, but it was neither a focus in this systematic review, nor investigated in the included studies. According to research by Van Rossem and Gage (77), age at first marriage is unrelated to the type of FGM/C a woman experiences. We concur with other researchers (e.g., (78)) that more empirical data are needed to provide a better picture of the role that FGM/C plays in the social lives of women. Many possible social consequences that merit investigation, such as (dis)honour and (loss of) social status are complex and do not easily lend themselves to controlled study. But they may be important consequences of having or not having FGM/C, both for the short term and long term.

Sexual consequences

Several studies addressing sexual consequences from FGM/C were sufficiently similar to warrant pooling of effect sizes in meta-analysis. Briefly, the results showed that FGM/C is associated with attenuation of a woman's sexual functioning, but the low quality of the body of evidence precludes us from drawing conclusions regarding causality.

Our estimate showed that women with FGM/C were more likely to report dyspareunia (pain during intercourse) and experience less sexual satisfaction than women who had not been subjected to FGM/C. We do not have sufficient evidence to say whether these associations are of a causal nature. Given the mechanisms of sexual response in women have not been clearly delineated (79) we cannot easily explain the mechanism of an association between painful sexual intercourse and reduced sexual satisfaction among women with FGM/C. However, we propose the mechanism of the relationship is likely physiological. With FGM/C there are injured clitoral nerves and related receptors, and various forms of scarring and adhesions around the excised genital parts (2;30;45). Cutting the vulva may damage neural innervations much the same way as radical vulvectomy following cancer (27). Additionally, healing from any type of cutting inevitably involves scar formation and pain in incision scars is ubiquitous, occurring through neuropathic pain mechanisms (80). Likely,

vulvae damage reduces the flexibility and sensitivity of critical genital tissue and may produce tearing during intercourse. It has been proposed that dyspareunia results from the scarring over the vulva after FGM/C and pain may be caused by friction during intercourse because of the scarring of nervous tissues (traumatic neuroma) at the site of the excision (66). All but one of the studies included in the meta-analysis for dyspareunia reported physical complications among women with FGM/C, generally scarring and infections, as the great majority of the women had FGM/C type I or II. For women with FGM/C type III, the pain may be greater, because infibulation creates a narrowed introitus (often a pinhole-sized opening) that is a mechanical obstruction. Nour (81) writes that since the infibulations must be opened up either surgically or through penetrative sex, sexual intercourse, especially in the first months of sexual intercourse, is frequently painful for the women. Such pain is reported in the literature (18;47;52).

Continuing the physiological argument, we repeat that anatomical, histological and MRI-based research identified five sexually responsive, vascular tissues of the female genitalia, which increase in blood volume during sexual arousal (26). In fact, vasocongestion of the erectile tissues, and the consequent lubrication of the vagina, form one of the initial phases of the sexual response cycle of both Master's and Johnson's model and Kaplan's Triphasic Model (25). While the limbic system in the brain and hormones influence and may also produce sexual responses, it is largely neural mechanisms that make sexual responses possible. Briefly, tactile stimulation of genital areas produces neural signals that are transmitted via the dorsal nerve of the clitoris and pudendal nerve to the sacral part of the spinal cord. This area then sends out a message to the muscles and arteries in the vulva, allowing genital vasocongestion to occur (25;79). The specialized sensory tissue of the clitoris is concentrated in a rich neurovascular area of a few centimetres (25;82). The damaged nerve-endings, scar tissues and adhesions around the excised genital parts from FGM/C are less receptive to tactile stimulation than intact genital tissues (45;66), and critical vascular tissues for sexual response may be missing in some women with FGM/C. Einstein (27) argues that FGM/C may through its potential effects on the central nervous system affect the entire body via neural rewiring. We recognize that sexual pleasure and satisfaction are multidimensional and also depend on affection, the time the partner takes to please the woman, and must be considered within the context in which a woman and her partner live (51). However, the physiological viewpoint not only helps explain dyspareunia and reduced sexual satisfaction, but also the experience of dryness during intercourse, which was found to be a problem among women with FGM/C in two of our included studies. Raina and colleagues (83) state that dyspareunia is often seen in women with decreased vaginal lubrication.

As mentioned, it is likely that the mechanism with respect to satisfaction is similarly to dyspareunia related to the physiologically and anatomically altered genital tissues, especially the clitoral tissue, following FGM/C. Our effect size estimate showed

an association between FGM/C and sexual satisfaction, but we do not have sufficient evidence to say whether this association is of a causal nature. However, women with FGM/C in Adinma's survey (62) were more than twice as likely to say they were not sexually satisfied, and in another study (65) women with FGM/C were more than three times as likely to say they had no enjoyment of sex compared to women without FGM/C. In her recent review, Obermeyer (56) suggested that the emphasis on the clitoris as the centre of sexual enjoyment in the western part of the world may be a social construction rather than a physiological reality. We cannot agree. Anatomical studies show the clitoris has an extensive and densely situated network of nerves, the majority of which innervate the top of the glans clitoris (82). Similarly, research has identified sexually responsive, vascular tissues of the female external genitalia (26). The clitoris is exquisitely sensitive, in some women it is more sensitive to touch than in others, with its only function being to produce sexual arousal (25). It is reasonable therefore, that researchers warn against disrupting the somatosensory innervation of the glans clitoris (26;82). Patently, we do not assert that capacity for sexual enjoyment depends on intact clitoris tissue, but consent it is an important part of female sexual response. Results from the study by Badawi (9) are illustrative: almost eight times as many women without FGM/C experienced sexual excitement from stimulation of the clitoral area than did women with FGM/C. While not all women who have undergone FGM/C experience sexual problems, it seems they may be physiologically less capable of becoming sexually stimulated than women without FGM/C. Our results offer some support for the hypothesis that the sensitivity and integrity of the clitoris and labia minora are essential for achieving sexual response (45).

Recognizing that humans' sexuality is a complex interaction of neurophysiological and biochemical mechanisms, and influenced by relationship dynamics and family and sociocultural issues (46), it is possible that women with FGM/C have an ability to compensate for anatomical damage to their genital tissue through enhancement of other sensory and erotic areas, or emotions and fantasy (limbic system). In fact, some researchers (e.g., (52)) suggest that sexual sensations in women with FGM/C may be maintained by a shift of the point of maximal sexual stimulation from the clitoris and/or labia to other areas of the body. Some support for this hypothesis was found in our systematic review, but considerable heterogeneity showed inconsistency across studies and our meta-analysis results were inconclusive. Women with FGM/C were less likely to say that the clitoris was their most sensitive body part and more likely to identify their breasts, tongue or vagina as their most sensitive body parts (68;70;72). An Eritrean woman in Dopico's study (47) stated: "It is not only the missing clitoris, or the stitched labia that give women pleasure; there are so many other parts of the body that increase pleasure. If you lack a little bit down there, you increase on the lips, you touch the boobs, and that compensates." The proposition regarding a possible compensatory sexual adjustment in women with FGM/C is of interest and should be investigated in future studies. It serves as a reminder that with FGM/C some essential structures for sexual stimulation and pleas-

ure have been excised, although not all – in many women an intact clitoris is buried beneath the infibulation scar (54;81) – and it indicates possibilities with respect to management, a point we address below. It is important to be reminded that sexual response is a biopsychosexual phenomenon (84), and measurement of overall sexual functioning inevitably complicated. Three studies included in this systematic review (63;65;73) used different scales to assess overall sexual functioning among women who had and women who had not been subjected to FGM/C. The study level results were in the same direction, indicating an association between sexual dysfunction and FGM/C.

Our observed greater likelihood of sexual desire among women without FGM/C than those with FGM/C is not easily explained. We do not have sufficient evidence to say whether the association is of a causal nature, and according to research, disruption of genital somatosensory tissue does not appear to affect desire (79). Sexual desire, also called libido, refers to an interest in sexual activity, leading the person to seek out such activity (25;80). It is possible that our result about sexual desire is related to conditioning. As delineated above, many women with FGM/C appear to experience pain during intercourse and impairment of sexual satisfaction. Repeated dissatisfactory sexual activity may affect women's sexual response (learned inhibition), evident as reduced desire in women. The same may be the case for initiation of sexual activity, the result for this outcome was inconclusive, and frequency of sexual intercourse, an outcome for which no direction in association could be observed. All three of these outcomes were only measured in a few studies and further research is warranted. As the experience of dyspareunia and decreased sexual satisfaction, attenuation of desire may be a detrimental sexual consequence of FGM/C that impedes a woman's ability to lead a healthy sex life. In this context, it is important to recognize that not all women experience sexual desire in the same way and that desire is difficult to measure (85), an issue we address below.

The result of our meta-analysis for orgasm, another central phase of the sexual response cycle, was inconclusive. However, it is in line with our other results that at study level, seven of the eight studies measuring orgasm found that women with FGM/C were less likely to experience orgasm. Hence orgasm may be part of a symptom constellation of the potential detrimental sexual consequences of FGM/C. The consequences of FGM/C on women's orgasm is an obvious area for future research. Based on our finding in this systematic review, we can only conclude that an association between FGM/C and orgasm seems indicated and merits further research.

A relevant question is whether there is a dose-response relationship involved in FGM/C and sexual functioning. In view of our physiological position, it is probable that a woman's sexuality is affected according to the age at which it was done, the healing process, and the extent of excision of genital tissue (rather than the suturing or closing of the vulva), the severity of which is influenced by type of operator, cutting instrument, and the use of antiseptics and anaesthetics. For years, researchers

(e.g., (2;74;86)) have proposed that the more severe the woman's excision, the more likely adverse consequences are, and there is some evidence to this effect. Almroth and colleagues (34) concluded that the anatomical extent of FGM/C, rather than whether the vulva had been sewn, was linked to infertility. A clear, positive relationship was also found between the type of FGM/C and the likelihood that a woman would have a long term complication, such as gynaecological and obstetric, or a genital infection (78). With respect to sexual functioning specifically, the dose-response question cannot be answered by this systematic review as we were only able to identify two studies which investigated differences between types of FGM/C, with insufficient similarity to allow for pooling of effect estimates. This issue can only be conclusively determined with intensified research efforts. However, since the studies included in the meta-analyses for dyspareunia, sexual satisfaction and sexual desire included mostly women with FGM/C type I or II, but also some with type III, it would appear that FGM/C of *any* type (except type IV) may be associated with sexual problems. We note that Shandall's results (74) showed that, compared to women without FGM/C, both women with FGM/C type I and those with type III were significantly more likely to report that they never experienced orgasm, and the effect size was largest in the comparison of women with FGM/C type III. Similarly, Thabet and Thabet (45) found that both women with FGM/C type I and type II&III experienced significant less frequent sexual desire than women without FGM/C, and also in this study the effect estimate was largest in the comparison of women with more anatomically extensive cutting.

We offer a few comments regarding implications. First, the low quantity- and quality of the evidence have clear bearing on arguments for increased efforts to investigate the possible consequences of FGM/C, particularly sexual harm given the direction of our results. Second, as mentioned above, it is possible some women have an ability to compensate for anatomical damage to their genital tissue through refocused development of other sensory and erotic areas, or emotions and fantasy. Sexual therapy and education, individual and couple, could be offered to women with FGM/C who want that. Other pragmatic assistance includes surgical operations for those women who want that and where resources permit it. Reports confirm women with FGM/C are seeking such assistance (87-89). Deinfibulation and clitoral restoration or reconstruction have been found to improve sexuality in women with FGM/C (88). Nour, Michels, and Bryant (81) found that defibulation resolved nearly all cases of dyspareunia and helped restore the external genitalia of women with an intact clitoris, which was buried beneath the scar in 48% of deinfibulated women. In post interviews all women and their husbands stated that sexual intercourse had improved. In another study, clitoro-labial reconstruction was followed by significant improvement in sexuality, subjectively and objectively (45).

QUALITY OF THE EVIDENCE

We rated the study quality of two studies as high, five as moderate, and the remaining ten studies as having low study quality. The assessment was based on twelve quality assessment questions of the modified NOKC checklist for cross-sectional studies (Appendix 3). Applying GRADE, the quality of the total body of evidence was assessed as 'very low,' meaning that any estimate of effect is very uncertain.

Clearly, a socially prescribed custom like FGM/C does not lend itself to a randomized controlled trial, the gold standard for drawing causal inferences. And to date, neither prospective cohort studies nor bona fide case-control studies have been conducted for FGM/C, illustrating the practical barriers to outcomes research on this issue. While acknowledging the difficulties inherent in carrying out studies of this custom, we emphasize the importance of applying the best design possible for examining the consequences of the practice. Cross-sectional studies are inherently problematic with respect to sampling. In all the included studies, 'exposed' (women with FGM/C) and 'unexposed' (women without FGM/C) groups were compared and the risk among those exposed compared to those unexposed were assessed. Recruiting sufficiently equivalent and large exposed and unexposed groups of women may be difficult. A strength of the included studies was that the non-exposed group was selected from the same population as the exposed group and the groups were largely comparable with respect to important background factors. Data on FGM/C are necessarily observational and therefore tests for the presence of confounding factors become important. Unfortunately, only three studies (69;72;75) included in this systematic review controlled for confounders. We propose that analyses such as multivariate regression analysis be undertaken in future studies, enabling isolation of the influence of known confounders and testing for unknown general confounding. Testing for confounding affords greater certainty about the usefulness of existing data and results (90).

Among the included studies in this systematic review, sample sizes ranged from 37 to 4,024 women. Recruitment concerns is likely the reason that 11 of the studies in this systematic review recruited respondents at health facilities, where the women came for a medical consultation. While the sample size in most studies was acceptable (mean sample size was 750, six studies enrolled <200 women, and five studies enrolled >900 women) and there were more women with FGM/C than without, methods of recruitment may have introduced sampling bias. Specifically, the women with FGM/C in these samples may not be representative of the general population of women with FGM/C regarding factors such as complications. Six studies reported physical complications among women with FGM/C, generally scarring and infections (62;65;69;71;72;74). Further, all of the studies, except two (69;75) failed to explain whether and how the participants who agreed to participate were different from those who refused to participate. We stress the importance of applying a methodology that not only ensures equivalency between exposed and unexposed groups

of women, but also representativeness. We believe researchers should aim to conduct well-designed comparative studies to investigate consequences of FGM/C, preferably from more than one centre or research group, similar to the WHO obstetric study (30). Such studies are comparatively simple and affordable to conduct, usually within the resource constraints of developing countries, and can be completed within a reasonable period of time.

A related issue is measurement of 'exposure' to FGM/C. Measuring exposure to FGM/C means determining the extent of genital tissue excised. In this systematic review, we applied the WHO classification system (including type I= clitoridectomy, type II= excision, type III= infibulation, and type IV= other) (2). A similar classification system was also applied by the great majority of the included studies. Further, in more than half of the included studies, classification and exposure were based on gynaecological examination. But in the remaining studies it was based on self-report. Research has shown that both validity and reliability of such self-reporting of FGM/C are variable (91). With respect to the included studies in this systematic review, Morison and colleagues (69) found 97% agreement between self-reported FGM/C status and that found on examination among Gambian women. Sensitivity in determining the presence or absence of genital cutting was 94% among Nigerian women, but it was more difficult for women to correctly identify type of cutting to which they had been subjected (72). Lastly, in our included study from Nigeria, self-reporting of FGM/C status was reliable in only 57% of women (62). Results from these and other studies ((91;92);EFCS in (69)) suggest that most women can correctly say whether or not they have been genitally cut, but are less able to correctly determine the extent of their cutting. In addition to reliability of self-report of FGM/C exposure, degree of exposure must be considered. Eight of the included studies in this systematic review reported that different types of genital cutting were combined in order to make a binary variable of women with FGM/C versus women without FGM/C. Upon finding no significant differences in the variables about sexual practices and pleasure between women with FGM/C type I and type II, Okonofua and colleagues (72) treated these women as one group of women with FGM/C. Only two studies, Shandall (74) and Thabet and Thabet (45), specifically compared different degrees of exposure. With respect to measurement of exposure, we recognize that gynaecological examination is subject to variation (interindividual and intraindividual), similar to radiology imaging interpretation there is a subjective element, but it is currently the best classification method available. Hence, future studies should base classification of FGM/C on gynaecological examination by trained personnel and compare degrees of exposure, as a dose-response relationship may be in effect where women with more severe forms of cutting experience greater difficulties.

In addition to the problems associated with study design, selection bias, and measurement of exposure to FGM/C, measurements of the psychosocial, social and sexual consequences that represent outcomes were complicated in the included studies.

By their very nature, all outcomes dealt with in this systematic review relied on women's reports of their experiences. With respect to sexual functioning, the degree to which individuals are 'at liberty' answering such questions honestly vary, thus misreporting in the included studies is possible. In Sudan and Egypt, custom places penalties on a woman's show of interest in sexual initiatives (52;93), likely disinclining certain women to participate in studies regarding sexual matters. Shyness, embarrassment and reluctance to talk about matters related to sexual functioning and behaviours may also have forced underreporting (65;67). Researchers concede "sexual experiences of mutilated women are very difficult to assess or research in our communities. Talking about sex remains a traditional taboo" (94). The core of the problem is that quality or functioning of psychological, social and sexual life is a latent concept. Perhaps therefore, but nonetheless troublesome, existed a lack of validity and reliability tested measures in all but two of the studies (65;73) in this systematic review, especially with respect to sexual functioning. It seems only Eldafrawi and colleagues (65) applied an instrument that was tested for validity and reliability with the target population. While most sexual functioning – as well as psychological and social functioning – is not amenable to direct physical measurement, many inventories and scales quantify and register sexual functioning status in a reliable and valid manner (95-97), such as the multidimensional Sexual Arousal and Desire Inventory (SADI), which is a reliable and valid tool to evaluate both state and trait aspects of subjective sexual arousal and desire (98). Such sexual inventories and tools are usually easy to administer and score, relatively unobtrusive and inexpensive (95), attesting to their usefulness. Unfortunately, most scales are as a rule validated for western populations and may not perform equally well for non-western groups.

Superimposed on the difficulty of validated measurement is the lack of a unified approach and definitions to measure sexual functioning in the included studies, which meant that we were in many cases unable to conduct meta-analyses. Few studies described how they explained their questions to the women and made sure they understood it in the same manner. Obermeyer (56) stresses the problems this causes, including non-comparable studies and widely divergent frequencies. We concur with her call for researchers to use precise definitions, appropriate communication and locally understandable terms when measuring sexual functioning. Further, research should be culturally grounded, in that issues meaningful to the target population are addressed, members of the population could be engaged in praxis (99) and resources from the culture respectfully used. Lastly, we acknowledge that several of the included studies failed to show use of standardized data collection methods. We reiterate the importance of applying a methodology that ensures quality data collection. Additionally, researchers should take care to minimize incomplete outcome data and to report on this issue; in several of the included studies in the current systematic review, information regarding response rate was lacking.

STRENGTHS AND LIMITATIONS

A strength of our systematic review is the comprehensive and systematic literature search and systematic process for identifying relevant publications. Two independent researchers at NOKC carried out the inclusion selection of publications based on pre-set inclusion criteria. A further strength is that we included controlled studies, i.e., study designs which could say something about health consequences of the exposure (FGM/C) on an outcome (e.g., dyspareunia). This approach is encouraged by researchers in the field (e.g., (100)). All included studies included a comparison group and data about the differences between the two groups, making it possible for us to estimate the likelihood of psychological, social and/or sexual consequences in women with FGM/C versus women without FGM/C, expressed as relative risk or mean difference, and in some cases meta-analyses could be performed. Randomized controlled trials were not sought as such designs are impossible in the case of socially prescribed practices (56) and ethically indefensible. *A posteriori*, we can state that we found no controlled studies that were evaluated by designs not included here. By accepting only evaluations with a comparison group we partially control for environmental stimulation, which seems important as sexuality is not culture neutral, but must be considered in the context in which a woman and her partner live.

Our systematic review may be subject to publication bias. Identifying all studies addressing the question of the systematic review is not always possible and we may have missed some studies. Studies we may have failed to identify may differ systematically from the ones we identified, the likeliest scenario being that the results of the present systematic review are biased to the positive. Further, we failed to obtain two identified records in full text (60;61), despite several attempts to obtain them. These could potentially have been included in our systematic review. One of the records (61) is a 1994 meeting paper entitled "Bad psychological effects of the procedure of girls' circumcision" and the other (60) a 1998 publication in Zigazig University Medical Journal which addresses the long term impact of circumcision on health of newly married females in Benha City, Egypt.

Caution is warranted in interpreting the results of this systematic review. We assessed the quality of the evidence as being too low to warrant conclusions about a causal relationship between FGM/C and psychological, social and sexual consequences. This was largely due to the weaknesses of the observational design of all included studies. We addressed study limitations in the 'quality of the evidence' section, but reiterate some methodological concerns. The composition characteristics of the groups, e.g., the variability of extent of cutting and age at which it was performed, may have had an influence of the results. In the included studies, the great majority of the women had been subjected to FGM/C type I or II, but most studies conflated FGM/C types I-III as having FGM/C. We were therefore unable to compare the various types of FGM/C. As discussed earlier, it is possible that with more severe forms of cutting, women's functioning is more severely affected. We also rec-

ognize the limits of self-reported FGM/C status. Of our 17 studies, slightly more than half included gynaecological examinations of the women to confirm whether or not the women had been genitally cut, and in some studies, to which type of FGM/C the women had been subjected. Reliability of self-report FGM/C status is variable.

Conclusions

The psychological, social and sexual consequences of FGM/C is an under-researched and neglected issue. The low quality of the body of evidence precludes us from drawing conclusions regarding causality, and the evidence base is insufficient to draw solid conclusions about the psychological and social consequences of FGM/C. However, our results substantiate the proposition that a woman whose genital tissues have been partly removed is more likely to experience increased pain and reduction in sexual satisfaction and desire.

NEED FOR FURTHER RESEARCH

The full range of psychological, social and sexual consequences of subjecting girls and young women to the various types of FGM/C has yet to be evaluated and further research is needed. In particular, major gaps in the literature about the psychological and social consequences of FGM/C still remain.

Additional studies investigating the consequences of FGM/C should compare clearly defined groups of women that differ by the extent of FGM/C, whereby classification is based on gynaecological examination by trained personnel. Studies should include women with FGM/C type III, compare women with varying degrees of FGM/C to shed light on the possible dose-response relationship involved in FGM/C, and recruit girls and young women so that also short term consequences of FGM/C can be better understood.

With respect to outcomes, we believe the range of psychological and social consequences are under-researched. In the sexual domain, the experience of orgasm is an obvious area for further research, as is sexual arousal. Outcomes should be identified by the study population as germane to their lived experiences.

Methodologically, research initiatives should apply a methodology that ensures equivalency of exposed and unexposed groups of women, as well as representativeness, preferably through multi-centre, well-designed comparative studies. Standardized data collection methods should be used, in which researchers use culturally appropriate communication, locally understandable terms and precise definitions. To enhance validity and value-relevance, members of the population could be actively

engaged and resources from the culture respectfully used. Instruments should be validity and reliability tested with the target populations.

We conclude that at the present time, there is a paucity of evidence regarding psychological and social consequences of FGM/C, and while evidence suggests that women with FGM/C are more likely to experience some detrimental sexual functioning than women without FGM/C, we encourage continued research in this area as well. Studies responsive to the dearth of literature regarding the psychological, social and sexual consequences of FGM/C should

- be prospective cohort or case-control studies
- examine short- and long term consequences
- be culturally grounded.

References

(1) Denison E, Berg RC, Fretheim A. Effectiveness of interventions designed to reduce the prevalence of female genital mutilation/cutting. Report from Kunnskapssenteret [25]. 2009. Oslo, Nasjonalt kunnskapssenter for helsetjenesten.

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

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

(4) WHO. Female genital mutilation – knew knowledge spurs optimism. Progress in Sexual and Reproductive Health Research 2006; 72:1-8.

(5) UNICEF. Female genital mutilation/female genital cutting: a statistical exploration. 2005. New York, United Nations Children's Fund.

(6) IRIN. Razors edge. The controversy over female genital mutilation. 2005.

(7) Benatar D, Benatar M. How not to argue about circumcision. Am J Bioeth 2003; 3(2):W1.

(8) Lannon C M. Circumcision - The debates goes on. Pediatrics 2000; 105(3):681-684.

(9) Badawi M. Epidemiology of female sexual castration in Cairo, Egypt. TruthSeeker 1989; 1(2):31-34.

(10) Al-krenawi A, Wiesel-Lev R. Attitudes toward and perceived psychosocial impact of female circumcision as practiced among the Bedouin-Arabs of the Negrev. Family Process 1999; 38(4):431-443.

(11) Dorkenoo E. Combating female genital mutilation: an agenda for the next decade. World Health Stat Q 1996; 49(2):142-147.

(12) Yoder S, Kahn S. Numbers of women circumcised in Africa: the production of a total. 39. 2008. United States Agency for International Development. DHS working papers.

(13) El-Zanaty R, Way A. Egypt demographic and health survey 2008. 2009. Cairo, Egypt, Ministries of Health, El-Zanaty and Associates, and Macro International.

(14) National population commission, Macro. Nigeria demographic and health survey 2003. 2004. Calverton, Maryland, National population commission and ORC Macro.

(15) Department of statistics, Institute for Resource Development. Sudan demographic and health survey 1989/1990. 1991.

(16) Applebaum J, Cohen H, Matar M, Abu RY, Kaplan Z. Symptoms of posttraumatic stress disorder after ritual female genital surgery among Bedouin in Israel: Myth or reality? Primary Care Companion to the Journal of Clinical Psychiatry 2008; 10(6):453-456.

(17) Grisaru N, Lezer S, Belmaker RH. Ritual female genital surgery among Ethiopian Jews. Arch Sex Behav 1997; 26(2):211-215.

(18) Asali A, Khamaysi N, Aburabia Y, Letzer S, Halihal B, Sadovsky M et al. Ritual female genital surgery among Bedouin in Israel. Arch Sex Behav 1995; 24(5):571-575.

(19) Halila S, Belmaker RH, Abu Rabia Y, Froimovici M, Applebaum J. Disappearance of female genital mutilation from the Bedouin population of Southern Israel. J Sex Med 2009; 6(1):70-73.

(20) European Parliament. Combating violence against women: European Parliament resolution on the current situation in combating violence against women and any future action. 2004/2220. 2004.

(21) Leye E, Deblonde J, Garcia-Anon J, Johnsdotter S, Kwateng-Kluvitse A, Weil-Curiel L et al. An analysis of the implementation of laws with regard to female genital mutilation in Europe. Crime 2007; 47(1):1-31.

(22) Johansen E, Bathija H, Khanna J. Work of the World Health Organization on Female Genital Mutilation: Ongoing research and policy discussions. Finnish Journal of Ethnicity and Migration 2008; 3:83-89.

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

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

(25) Hyde JS, Delamater JD. *Understanding human sexuality*. 9th ed. Boston: McGraw Hill, 2006.

(26) Yang CC, Cold CJ, Yilmaz U, Maravilla KR. Sexually responsive vascular tissue of the vulva. British Journal of Urology International 2005; 97(4):766-772.

(27) Einstein G. From body to brain: considering the neurobiological effects of female genital cutting. Perspect Biol Med 2008; 51(1):84-97.

(28) Toubia N. Female circumcision as a public health issue. N Engl J Med 1994; 331(11):712-716.

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

(30) WHO study group on female genital mutilation and obstetric outcome, Banks E, Meirik O, Farley T, Akande O, Bathija H et al. Female genital mutilation and obstetric outcome: WHO collaborative prospective study in six African countries. Lancet 2006; 367(9525):1835-1841.

(31) Essen B, Sjoberg N-O, Gudmundsson S, Ostergren P-O, Lindqvist PG. No association between female circumcision and prolonged labour: A case control study of immigrant women giving birth in Sweden. European Journal of Obstetrics Gynecology and Reproductive Biology 2005; 121(2):182-185.

(32) Essen B, Bodker B, Sjoberg N-O, Gudmundsson S, Ostergren P-O, Langhoff-Roos J. Is there an association between female circumcision and perinatal death? Bull World Health Organ 2002; 80(8):629-632.

(33) Larsen U, Yan S. Does Female Circumcision Affect Infertility and Fertility? A Study of the Central African Republic, Cote d'Ivoire, and Tanzania. Demography 2000; 37(3):313-321.

(34) Almroth L, Elmusharaf S, El Hadi N, Obeid A, El Sheikh MAA, Elfadil SM et al. Primary infertility after genital mutilation in girlhood in Sudan: a case-control study. Lancet 2005; 366(9483):385-391.

(35) UNICEF. Changing a harmful social convention: Female genital mutilation/cutting. 2005. New York, United Nations Children's Fund.

(36) Behrendt A, Moritz S. Posttraumatic stress disorder and memory problems after female genital mutilation. Am J Psychiatry 2005; 162(5):1000-1002.

(37) El Defrawi MH, Lotfy G, Dandash KF, Refaat AH, Eyada M. Female genital mutilation and its psychosexual impact. [References]. Journal of Sex & Marital Therapy 2001; 27(5):465-473.

(38) Al Krenawi A, Wiesel-Lev R. Attitudes toward and perceived psychosocial impact of female circumcision as practiced among the Bedouin-Arabs of the Negev. Fam Process 1999; 38(4):431-443.

(39) Vloeberghs E, Knipscheer J, van der Kwaak A, Naleie Z, van den Muijsenbergh M. Veiled pain. A study in the Netherlands on the psychological, social and relational consequences of female genital mutilation. 2010. Pharos.

(40) Chalmers B, Kowser OH. 432 Somali women's birth experiences in Canada after earlier female genital mutilation. Birth 2000; 27(4):227-234.

(41) Mwangi-Powell F. The SWITCH project: A FGM related study. 1999. London, FORWARD.

(42) Abusharaf RM. Virtuous Cuts: Female Genital Circumcision in an African Ontology. Differences 2001; 12(1):112-140.

(43) Almroth L, Almroth-Berggren V, Hassanein OM, El Hadi N, Al Said SS, Hasan SS et al. A community based study on the change of practice of female genital mutilation in a Sudanese village. Int J Gynaecol Obstet 2001; 74(2):179-185.

(44) Van der Kwaak A. Female circumcision and gender identity: A questionable alliance? Soc Sci Med 1992; 35(6):777-787.

(45) Thabet SMA, Thabet ASMA. Defective sexuality and female circumcision: The cause and the possible management. J Obstet Gynaecol Res 2003; 29(1):12-19.

(46) Levine SB. The nature of sexual desire: a clinician's perspective. Archives of Sexual Behavior 2003; 32(3):279-285.

(47) Dopico M. Infibulation and the orgasm puzzle: Sexual experiences of infibulated Eritrean women in rural Eritrea and Melbourne, Australia. In: Her-

nlund Y, Shell-Duncan B, editors. Transcultural Bodies: Female genital cutting in global context. London: Rutgers University Press, 2007: 224-247.

(48) Mukoro UJ. A survey of the psychosexual implications of female genital mutilation on Urhobo women of the Niger Delta communities of Nigeria. Journal of Human Ecology 2004; 16(2):147-150.

(49) Reyners M. Health consequences of female genital mutilation. Reviews in Gynaecological Practice 2004; 4(4):242-251.

(50) Abor PA. Female genital mutilation: Psychological and reproductive health consequences. The case of Kayoro traditional area in Ghana. Gender & Behaviour 2006; 4(1):659-684.

(51) Catania L, Abdulcadir O, Puppo V, Verde JB, Abdulcadir J, Abdulcadir D. Pleasure and orgasm in women with female genital mutilation/cutting (FGM/ C). J Sex Med 2007; 4(6):1666-1678.

(52) Lightfoot-Klein H. The sexual experience and marital adjustment of genitally circumcised and infibulated females in the Sudan. Journal of Sex Research 1989; 26(3):375-392.

(53) Johnsdotter S, Essen B. It Is Only a Tradition: Making Sense of Swedish Somalis' Narratives of Female Circumcision and Avoiding Submission to Hegemonic Political Discourse. Critical Social Policy 2005; 25(4):577-595.

(54) Johansen REB. Experiencing sex in exile: Can genitals change their gender? On conception and experiences related to female genital cutting (FGC) among Somalis in Norway. In: Hernlund Y, Shell-Duncan B, editors. Transcultural Bodies: Female genital cutting in global context. London: Rutgers University Press, 2007: 248-277.

(55) Obermeyer CM. Female genital surgeries: The known, the unknown, and the unknowable. Medical Anthropology Quarterly, New Series 1999; 13(1):79-106.

(56) Obermeyer CM. The consequences of female circumcision for health and sexuality: An update on the evidence. Cult Health Sex 2005; 7(5):443-461.

(57) Nasjonalt kunnskapssenter for helsetjenesten. *Slik oppsummerer vi forskning. Håndbok for Nasjonalt kunnskapssenter for helsetjenesten.* 2nd ed. Oslo: Nasjonalt kunnskapssenter for helsetjenesten., 2009.

(58) Higgins JPT. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 ed. The Cochrane Collaboration, 2008.

(59) Guyatt GH, Oxman AA, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is "quality of evidence" and why is it important to clinicians? BMJ 2008; 336(7651):995-998.

(60) Abd El-Hady RM, El-Nashar AB. Long term impact of circumcision on health of newly married females in Benha City. Zagazig University Medical Journal 1998; 6:839-851.

(61) Omara BH. Bad psychological effects of the procedure of girls' circumcision. 1994. Paper presented to the Meeting of NGO in preparation for the International Conference for Population and Development, Cairo.

(62) Adinma JI. Current status of female circumcision among Nigerian Igbos. West Afr J Med 1997; 16(4):227-231.

(63) Alsibiani S, Rouzi AA. Sexual functioning in women with female genital mutilation. Fertility & Sterility 2010; 93(3):722-724.

(64) Behrendt A, Moritz S. Posttraumatic stress disorder and memory problems after female genital mutilation. Am J Psychiatry 2005; 162(5):1000-1002.

(65) El Defrawi MH, Lotfy G, Dandash KF, Refaat AH, Eyada M. Female genital mutilation and its psychosexual impact. J Sex Marital Ther 2001; 27(5):465-473.

(66) Elnashar, A, Abdelhady, R. The impact of female genital cutting on health of newly married women. International Journal of Gynecology & Obstetrics 2007; 97(3):238-244.

(67) Elnashar AM, Dien Ibrahim M, El Desoky MM, Ali OM, Sayd Mohamed Hassan M. Female sexual dysfunction in Lower Egypt. BJOG 2007; 114(2):201-206.

(68) Megafu U. Female ritual circumcision in Africa an investigation of the presumed benefits among Ibos of Nigeria. East Afr Med J 1983; 60(11):793-800.

(69) Morison L, Scherf C, Ekpo G, Paine K, West B, Coleman R et al. The long-term reproductive health consequences of female genital cutting in rural Gambia: a community-based survey. Trop Med Int Health 2001; 6(8):643-653.

(70) Nwajei SD, Otiono AI. Female Genital Mutilation: Implications for Female Sexuality. Women's Studies International Forum 2003; 26(6):575-580.

(71) Odoi A, Brody SP, Elkins TE. Female genital mutilation in rural Ghana, west Africa. Int J Gynaecol Obstet 1997; 56(2):179-180.

(72) Okonofua FE, Larsen U, Oronsaye F, Snow RC, Slanger TE. The association between female genital cutting and correlates of sexual and gynaecological morbidity in Edo State, Nigeria. BJOG 2002; 109(10):1089-1096.

(73) Osinowo HO, Taiwo AO. Impact of Female Genital Mutilation on Sexual Functioning, Self-Esteem and Marital Instability of Women in Ajegunle. IFE Psychologia: An International Journal 2003; 11(1):123-130.

(74) Shandall AA. Circumcision and infibulation of females: a general consideration of the problem and a clinical study of the complications in Sudanese women. Sudan Med J 1967; 5(4):178-212.

(75) Stewart H, Morison L, White R. Determinants of coital frequency among married women in Central African Republic: The role of female genital cutting. J Biosoc Sci 2002; 34(4):525-539.

(76) McCabe M, Cobain MJ. The impact of individual and relationship factors on sexual dysfunction among males and females. Journal of Sex and Marital Therapy 1998; 13:131-143.

(77) Van Rossem R, Gage AJ. The effects of female genital mutilation on the onset of sexual activity and marriage in Guinea. Arch Sex Behav 2009; 38(2):178-185.

(78) Jones H, Diop N, Askew I, Kabore I. Female genital cutting practices in Burkina Faso and Mali and their negative health outcomes. Stud Fam Plann 1999; 30(3):219-230.

(79) Yang CC, Bowen JR, Kraft GH, Uchio EM, Kromm BG. Cortical evoked potentials of the dorsal nerve of the clitoris and female sexual dysfunction in multiple sclerosis. Journal of Urology 2000; 164:2010-2013.

(80) Steege J, Solnoun DA. Evaluation and treatment of dyspareunia. Obstetrics & Gynecology 2009; 113(5):1124-1136.

(81) Nour NM, Michels KB, Bryant AE. Defibulation to treat female genital cutting: Effect on symptoms and sexual function. Obstet Gynecol 2006; 108(1):55-60.

(82) Baskin LS, Erol A, Li YW, Liu WH, Kurzrock E, Cunha GR. Anatomical studies of the human clitoris. Journal of Urology 1999; 162:1015-1020.

(83) Raina R, Pahlajani G, Khan S, Gupta S, Agarwal A, Zippe CD. Female sexual dysfunction: classification, pathophysiology and management. Fertility and Sterility 2007; 88:1273-1284.

(84) Leiblum SR. Redefining female sexual response. Contemporary Ob/Gyn 2000; 45(11):120-126.

(85) Tunuguntla HSGR, Gousse AE. Female sexual dysfunction following vaginal surgery: A review. The Journal of Urology 2006; 175:439-446.

(86) Cutner LP. Female genital mutilation. Obstet Gynecol Surv 1985; 40(7):437-443.

(87) Villani M. From the "maturity" of a woman to surgery: Conditions for clitoris repair. Sexologies 2009; 18(4):259-261.

(88) Foldes P, Louis-Sylvestre C. [Results of surgical clitoral repair after ritual excision: 453 cases]. Gynecol Obstet Fertil 2006; 34(12):1137-1141.

(89) Morison LA, Dirir A, Elmi S, Warsame J, Dirir S. How Experiences and Attitudes Relating to Female Circumcision Vary According to Age on Arrival in Britain: A Study among Young Somalis in London. Ethn Health 2004; 9(1):75-100.

(90) Buck D, Sutton M. Interpreting results of observational research. BMJ 1994; 309:1439-1440.

(91) Elmusharaf S, Elhadi N, Almroth L. Reliability of self reported form of female genital mutilation and WHO classification: Cross sectional study. Br Med J 2006; 333(7559):124-127.

(92) Snow RC, Slanger TE, Okonofua FE, Oronsaye F, Wacker J. Female genital cutting in southern urban and peri-urban Nigeria: self-reported validity, social determinants and secular decline. Trop Med Int Health 2002; 7(1):91-100.

(93) El-Mouelhy M, Fahmy A, Ragab AR. Investigating women's sexuality in relation to female genital mutilation in Egypt. 2010.

(94) Daffeh J, Dumbuya S, Sosseh-Gaye A. Listening to the voice of the people: A situation analysis of female genital mutilation in the Gambia. 1995. Banjul, the Gambia, WHO, UNFPA, UNICEF.

(95) Corona G, Jannini EA, Maggi M. Inventories for male and female sexual dysfunctions. International Journal of Impotence Research 2010; 18:236-250.

(96) DeRogatis LR. Assessment of sexual function/dysfunction via patient reported outcomes. International Journal of Impotence Research 2007; 20:35-44.

(97) Kaufman M, Scarpero HM. Dyspareunia and surgery: can we measure sexual dysfunction and outcomes? Current Bladder Dysfunction Reports 2008; 3(2):85-89.

(98) Toledano R, Pfaus J. The sexual arousal and desire inventory (SADI): a multi-dimensional scale to assess subjective sexual arousal and desire. Journal of Sexual Medicine 2006; 3(5):853-877.

(99) Freire P. *Pedagogy of the oppressed.* New York: Continuum, 1995.

(100) Askew I. Methodological issues in measuring the impact of interventions against female genital cutting. Cult Health Sex 2005; 7(5):463-477.

(101) Balogun SK. Female Circumcision: Its Psychological Effects on Victims, Family and the Society. Anthropologist 2001; 3(4):261-263.

(102) Carton V, Philippe HJ. [Medical, psychological and sexual consequences of female genital mutilations]. Arch Pediatr 2008; 15(5):820-821.

(103) Gallo PG. Female circumcision in Somalia: some psychosocial aspects. Genus 1985; 41(1-2):133-147.

(104) Hosken FP. Female circumcision in Africa. Victimology 1977; 2(3-4):487-4.

(105) Larsen U, Yan S. Does female circumcision affect infertility and fertility? A study of the central African Republic, Cote d'Ivoire, and Tanzania. Demography 2000; 37(3):313-321.

(106) Lenzi E, Santuari E. [Infibulation and its consequences]. Riv Ital Ginecol 1969; 53(6):461-470.

(107) Lightfoot-Klein H. Disability in female immigrants with ritually inflicted genital mutilation. Women Ther 1993; 14(3-4):187-194.

(108) Lo Baido R, La Grutta S, Profeta E, Schiera G. Female Genital Mutilations: Echoes in the mind of scars on the body. Rivista di Psichiatria 2007; 42(3):183-188.

(109) Lo Baido R, La Grutta S, Bressi C, Mauri M, Trombini E. The Female Genital Mutilations (FGM): A clinical and psychopathological study on a group of immigrants in Sicily. Rivista di Psichiatria 2004; 39(4):229-237.

(110) Onadeko MO, Adekunle LV. Female circumcision in Nigeria: a fact or a farce? J Trop Pediatr 1985; 31(4):180-184.

(111) Tabrizi E. The price of womanhood: Female genital cutting. Curr Surg 2004; 61(6):565-566.

(112) Thabet SMA. Reality of the G-spot and its relation to female circumcision and vaginal surgery. J Obstet Gynaecol Res 2009; 35(5):967-973.

Appendices

---

1. SEARCH FOR LITERATURE

Database: African Index Medicus (AIM)
Search: Sari Ormstad
Date: 05.02.2009
Retrieval: 12
Strategy:
"CIRCUMCISION" [Descriptor] or "CIRCUMCISION, FEMALE" [Descriptor] or "INFIBULATION" [Descriptor]

Database: Anthropology Plus
Search: Hege Oswald
Date: 18.02.09
Retrieval: 200
Strategy:
((kw: female* or kw: wom#n) or kw: girl*) and ((kw: mutilation* or kw: infibulate*) or kw: cutting*) or kw: infibulate*

Database: Ovid British Nursing Index and Archive 1985 to January 2009
Search: Sari Ormstad
Date: 04.02.2009
Retrieval: 135
Strategy:
1. Circumcision/
2. ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.
3. "fgm/c".tw.
4. ((removal$ or alteration$ or excision$) adj6 female genital$).tw.
5. pharaonic circumcision$.tw.
6. sunna.tw.
7. (clitoridectom$ or clitorectom$).tw.
8. (infibulate$ or reinfibulat$ or deinfibulat$).tw.
9. or/1-8

Database: The Cochrane Library 2009, Issue 1: Cochrane Central Register of Controlled Trials (Clinical Trials)

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 3

Strategy:

#1 MeSH descriptor Circumcision, Female, this term only

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

#3 (#1 OR #2)

Database: Ovid EMBASE 1980 to 2009 Week 05

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 570

Strategy:

1.   female circumcision/ or female genital mutilation/ or female genital cutting/ or infibulatens/

2.   ((female$ or wom#n or girl$1) adj3 (mutilation$ or infibulate$ or cutting$)).tw.

3.   "fgm/c".tw.

4.   ((removal$ or alteration$ or excision$) adj6 female genital$).tw.

5.   pharaonic circumcision$.tw.

6.   sunna.tw.

7.   (clitoridectom$ or clitorectom$).tw.

8.   (infibulate$ or reinfibulat$ or deinfibulat$).tw.

9.   or/1-8

Database: EPOC Register – BiblioWeb – Advanced search

Search: Sari Ormstad

Date: 09.02.09

Retrieval: 0

Strategy:

Title or Abstract or Keyword: circumcise% or mutilation% or FGM or clitoridectom% or clitorectom% or infibulate% or deinfibulat% or de-infibulat% or reinfibulat% or re-infibulat% or female genital%

Database: Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE® 1950 to Present

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 1100

Strategy:

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

Database: CSA Illumina: PILOTS database (1871-Current)

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 14

Strategy:

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

Database : POPLINE® (POPulation information onLINE)

Search: Sari Ormstad

Date: 05.02.2009

Retrieval: 1566

Strategy:

POPLINE Advanced Search KEYWORDS:

FEMALE GENITAL CUTTING

Database: Ovid PsycINFO 1967 to February Week 1 2009

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 401

Strategy:

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

Database: CSA Illumina: Social Services Abstracts (1979-Current)

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 40

Strategy:

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

Database: CSA Illumina: Sociological Abstracts (1952-Current)

Search: Sari Ormstad

Date: 04.02.2009

Retrieval: 325

Strategy:

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

Database: WHO Library & Information Networks for Knowledge Database (WHOLIS)

Search: Sari Ormstad

Date: 05.02.2009

Retrieval: 65

Strategy:

words or phrase "((female$ or wom?n or girl or girls) near3 (mutilation$ or circum-cis$ or cutting$))"

OR

words or phrase ""fgm/c""

OR

words or phrase "((removal$ or alteration$ or excision$) near6 (female adj geni-tal$))"

OR

words or phrase "(pharaonic adj circumcision$)"

OR

words or phrase "sunna"

OR

words or phrase "(clitoridectom$ or clitorectom$)"

OR

words or phrase "(infibulate$ or reinfibulat$ or deinfibulat$)"

We searched databases of six international organizations that are engaged in projects regarding FGM/C: the Centre for Development and Population Activities (CEDPA), Population Council, Population Reference Bureau (PRB), the United Nations Children's Fund (UNICEF), the United Nations Population Fund (UNFPA), and the World Health Organization (WHO).

2. EXCLUDED STUDIES

Table A1: Excluded studies read in full text and reason for exclusion

Study	Cause for exclusion of study
Abor, 2006 (50)	Not a study where females with FGM/C are compared with females without FGM/C
al-Krenawi, 1999 (10)	Psychological, social or sexual consequences of FGM/C are not reported (quantitatively)
Balogun, 2001 (101)	Not a study where females with FGM/C are compared with females without FGM/C
Carton, 2008 (102)	Not a study where females with FGM/C are compared with females without FGM/C
Catania, 2007 (51)	Not a study where females with FGM/C are compared with females without FGM/C who were part of a community in which FGM/C is customary practiced
Gallo, 1985 (103)	Not a study where females with FGM/C are compared with females without FGM/C
Hosken, 1977 (104)	Not a study where females with FGM/C are compared with females without FGM/C
Jones, 1999 (78)	Not a study where females with FGM/C are compared with females without FGM/C
Larsen, 2000 (105)	Psychological, social or sexual consequences of FGM/C are not reported
Lenzi, 1969 (106)	Not a study where females with FGM/C are compared with females without FGM/C
Lightfoot-Klein, 1993 (107)	Not a study where females with FGM/C are compared with females without FGM/C
Lo Baido, 2007 (108)	Not a study where females with FGM/C are compared with females without FGM/C
Lo Baido, 2004 (109)	Not a study where females with FGM/C are compared with females without FGM/C
Mukoro, 2004 (48)	Not a study where females with FGM/C are compared with females without FGM/C

Nour, 2006 (81)	Not a study where females with FGM/C are compared with females without FGM/C
Onadeko, 1985 (110)	Not a study where females with FGM/C are compared with females without FGM/C
Tabrizi, 2004 (111)	Not a study where females with FGM/C are compared with females without FGM/C
Thabet, 2009 (112)	Psychological, social or sexual consequences of FGM/C are not reported
Toubia, 1994 (28)	Not a study where females with FGM/C are compared with females without FGM/C

3. QUALITY ASSESSMENT

Quality assessment questions for included studies

All questions are answered 'yes', 'unclear/somewhat', or 'no':

1. Was the population from which the sample was drawn clearly defined?
2. Was the sample representative of the population?
3. Was the non-exposed group selected from the same population as the exposed group?
4. Were the groups comparable with respect to important background factors?
5. Is it explained whether (and how) the participants who agreed to participate are different from those who refused to participate?
6. Is the response rate adequate?
7. Were standardized data collection methods used?
8. Were measures shown to be reliable and valid?
9. Were exposure and outcome measured in the same way and reliably in the two groups?
10. Was the person who assessed the outcome blind to whether participants were exposed or not.
11. Were the statistical methods appropriate?
12. Have known, potential confounders been considered in the study design and/or analyses?

Description of assessment of study quality:

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

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

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

Table A2: Results of quality assessment

Study	1	2	3	4	5	6	7	8	9	10	11	12	Assessment
Adinma 1997	Yes	Unclear	Yes	Unclear	No	Yes	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Alsibiani 2010	Yes	Unclear	Yes	Yes	No	Yes	Unclear	No	Yes	Unclear	Yes	Yes	Moderate
Applebaum 2008	Yes	Unclear	Yes	Yes	No	Yes	Yes	No	Yes	Unclear	Yes	Yes	Moderate
Badawi 1989	No	Unclear	Unclear	Unclear	No	Yes	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Behrendt 2005	No	Unclear	Yes	Yes	No	Unclear	Yes	No	Yes	Unclear	Yes	Yes	Moderate
El-dafrawi 2001	Yes	Unclear	Yes	Unclear	No	Unclear	Yes	Yes	Yes	Yes	Unclear	Unclear	Moderate
Elnashar 2007a	Yes	Yes	Yes	Unclear	No	Unclear	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Elnashar 2007b	Yes	Yes	Yes	Unclear	No	Yes	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Megafu 1983	Yes	Unclear	Yes	Unclear	No	No	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Morison 2001	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Unclear	Yes	Yes	High
Nwajei 2003	Yes	Unclear	Yes	Unclear	No	Unclear	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Odoi 1997	No	Unclear	Yes	Unclear	No	Unclear	Unclear	No	Unclear	Unclear	Unclear	Unclear	Low
Okonofua 2002	Yes	Unclear	Yes	Unclear	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Moderate
Osinowo 2003	No	Unclear	Yes	Unclear	No	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Low
Shandall 1967	Yes	Unclear	Yes	Unclear	No	Yes	Unclear	No	Yes	Unclear	Unclear	Unclear	Low
Stewart 2002	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	No	Yes	Unclear	Yes	Yes	High
Thabet 2003	No	Unclear	Yes	Yes	No	Unclear	Yes	No	Yes	Unclear	No	Yes	Low

Legend: Number refers to assessment question.

4. GRADE

Women with FGM/C compared to women without FGM/C regarding sexual and psychological functioning

Patient or population: women
Settings: Central African Republic, Egypt, Gambia, Ghana, Israel, Nigeria, Saudi Arabia, Senegal, Sudan
Condition: FGM/C
Comparison: non-FGM/C

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk non-FGM/C	Corresponding risk FGM/C				
dyspareunia self-report	**Study population**		**RR 1.52** (1.15 to 2)	3260 (5 studies)	⊕⊖⊖⊖ **very low**[1]	
	66 per 1000	100 per 1000 (76 to 132)				
	Low risk population					
	20 per 1000	30 per 1000 (23 to 40)				
	High risk population					
	320 per 1000	486 per 1000 (368 to 640)				
satisfaction self-report. Scale from: 1 to 6.	The mean satisfaction in the control groups was **6.5 points**	The mean satisfaction in the intervention groups was **0.34 lower** (0.56 to 0.13 lower)		350 (2 studies)	⊕⊖⊖⊖ **very low**[2,3]	
desire self-report	158 per 1000	340 per 1000 (216 to 531)	**RR 2.15** (1.37 to 3.36)	514 (2 studies)	⊕⊖⊖⊖ **very low**[4]	
orgasm self-report	**Study population**		**RR 1.50** (0.93 to 2.44)	3790 (5 studies)	⊕⊖⊖⊖ **very low**[5,6,7]	
	314 per 1000	471 per 1000 (292 to 766)				
	Low risk population					
	High risk population					
	400 per 1000	600 per 1000 (372 to 976)				
clitoris most sensitive area self-report	**Study population**		**RR 0.55** (0.35 to 0.85)	2736 (3 studies)	⊕⊖⊖⊖ **very low**[8,9]	
	337 per 1000	185 per 1000 (118 to 286)				
	Low risk population					
	270 per 1000	149 per 1000 (94 to 230)				
	High risk population					
	510 per 1000	281 per 1000 (178 to 434)				
breasts most sensitive area self-report	**Study population**		**RR 1.04** (0.6 to 1.78)	2736 (3 studies)	⊕⊖⊖⊖ **very low**[10,11,12]	
	393 per 1000	409 per 1000 (236 to 700)				
	Low risk population					
	180 per 1000	187 per 1000 (108 to 320)				
	High risk population					
	440 per 1000	458 per 1000 (264 to 783)				
initiate sex self-report	**Study population**		**RR 0.94** (0.74 to 1.21)	2086 (2 studies)	⊕⊖⊖⊖ **very low**[13,14,15]	
	466 per 1000	438 per 1000 (345 to 564)				
	Low risk population					
	380 per 1000	357 per 1000 (281 to 460)				

	High risk population				
	410 per 1000	**385 per 1000** (303 to 496)			
anxiety SCL-90		The mean anxiety in the intervention groups was **0.69 standard deviations higher** (-1.27 lower to 2.66 higher)		165 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,6]
somatization SCL-90		The mean somatization in the intervention groups was **0.67 standard deviations higher** (-0.59 lower to 1.94 higher)		165 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,6]
depression SCL-90		The mean depression in the intervention groups was **-0.15 standard deviations lower** (-0.64 lower to 0.34 higher)		165 (2 studies)	⊕⊖⊖⊖ **very low**[2,3]
hostility SCL-90		The mean hostility in the intervention groups was **-0.13 standard deviations lower** (-1.00 lower to 0.73 higher)		165 (2 studies)	⊕⊖⊖⊖ **very low**[2,3,9]

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Study quality moderate across the five studies
[2,4,5,8,10] Variable study quality (moderate and low)
[3] Total population size is less than 300 (for dichotomous outcome) or 400 (for continuous outcome)
[6,11] CIs do not overlap; heterogeneity
[7,12,15] Wide CI
[9] Heterogeneity
[13] Study quality moderate across the two studies
[14] Heterogeneity