

“A bit more truthful”: the validity of adolescent sexual behaviour data collected in rural northern Tanzania using five methods

M L Plummer, D A Ross, D Wight, J Changalucha, G Mshana, J Wamoyi, J Todd, A Anemona, F F Moshia, A I N Obasi, R J Hayes

Sex Transm Infect 2004;**80**(Suppl II):ii49–ii56. doi: 10.1136/sti.2004.011924

See end of article for authors' affiliations

Correspondence to:
Dr M Plummer,
Department of Infectious
and Tropical Diseases,
London School of Hygiene
and Tropical Medicine,
Keppel Street, London
WC1E 7HT, UK;
mary.plummer@
lshtm.ac.uk

Accepted for publication
19 July 2004

Objective: To assess the validity of sexual behaviour data collected from African adolescents using five methods.

Methods: 9280 Tanzanian adolescents participated in a biological marker and face to face questionnaire survey and 6079 in an assisted self-completion questionnaire survey; 74 participated in in-depth interviews and 56 person weeks of participant observation were conducted.

Results: 38% of males and 59% of females reporting sexual activity did so in only one of the two 1998 questionnaires. Only 58% of males and 29% of females with biological markers consistently reported sexual activity in both questionnaires. Nine of 11 (82%) in-depth interview respondents who had had biological markers provided an invalid series of responses about sex in the survey and in-depth interview series. Only one of six female in-depth interview respondents with an STI reported sex in any of the four surveys, but five reported it in the in-depth interviews.

Conclusion: In this low prevalence population, biological markers on their own revealed that a few adolescents had had sex, but in combination with in-depth interviews they may be useful in identifying risk factors for STIs. Self-reported sexual behaviour data were fraught with inconsistencies. In-depth interviews seem to be more effective than assisted self-completion questionnaires and face to face questionnaires in promoting honest responses among females with STIs. Participant observation was the most useful method for understanding the nature, complexity, and extent of sexual behaviour.

We need to improve our understanding of the sexual behaviour of young people in sub-Saharan Africa in order to develop and improve HIV prevention programmes for them. Assessing sexual behaviour is challenging, however, because of its sensitive and private nature.^{1–2} Data collected using quantitative methods, such as self-completion questionnaires (SCQs) or face to face questionnaires (FFQs), are generally considered to be more representative, but have less depth, than those collected via qualitative methods, such as in-depth interview or participant observation.^{3–4}

Methods that rely on self-reported sexual behaviour (such as SCQ, FFQ, and in-depth interview) may have problems of misunderstanding, poor recall, and social desirability bias.^{5–8} However, methods that attempt to evaluate sexual behaviour “externally” can also be problematic. For example, measurement of biological markers such as STIs or pregnancy may not always be ethical, practical, or affordable.^{9–10} Similarly, direct observation of sexual behaviour by participant observation may be both difficult and inappropriate.¹¹ In practice, participant observation researchers often rely on informal first or third person reports, or observation of behaviour that precedes or follows sexual activity, rather than observation of the acts themselves.¹²

The general strengths and weaknesses of these research methods are widely recognised, but few studies have attempted to assess the effectiveness of two such methods within one population, let alone all five. This paper compares the reliability and validity of sexual behaviour data collected using biological markers, FFQ, SCQ, in-depth interview, and participant observation in the same rural African adolescent population. Both biological marker and participant observation data are explored as external means of validating sexual

behaviour reported in formal interview settings. The research took place as part of the MEMA kwa Vijana (MkV) project, a community randomised adolescent sexual and reproductive health intervention trial in Mwanza, Tanzania. Twenty rural communities participated in the MkV trial, 10 of which received the intervention; each research method was used equally in intervention and comparison arms. Detailed descriptions of the process and results of each research method are published elsewhere.^{13–19}

METHODS

From September to December 1998, 9283 primary school pupils in years 4–6 participated in a school based combined biological marker/FFQ survey; only 18 pupils (<1%) refused to participate. All participants were born before 1985; the mean age was 15.5 years for males and 14.8 years for females. Young, same sex research assistants conducted 15–20 minute FFQ interviews, consisting of 50 questions focused on sexual and reproductive health knowledge, attitudes, and behaviour.

After the FFQ interview, participants were asked to provide a urine specimen that was later tested for HIV antibodies via GACPAT (Central Public Health Laboratory, Virus Reference Division, Colindale, London, UK) with positive specimens confirmed by GACELISA (Murex Biotech Ltd, Dartford, UK), *Chlamydia trachomatis* and *Neisseria gonorrhoeae* via pooled polymerase chain reaction (Roche Diagnostics Systems, Branchburg, New Jersey, USA),²⁰ and pregnancy in females via an HCG test (Ipas, Carrboro, North Carolina, USA).

Abbreviations: ASCQ, assisted self-completion questionnaire; FFQ, face to face questionnaire; MkV, MEMA kwa Vijana; SCQs, self-completion questionnaire.

Participants were then seen by a clinician, who provided syndromic treatment of STI and schistosomiasis, and offered confidential HIV counselling and testing. Participants who wished to know their HIV status then provided a blood specimen for that test.

During the same time period, an overlapping cohort of 6079 young people participated in another school based survey, using an innovative assisted self-completion questionnaire (ASCQ).¹⁵ ASCQ participants were in school years 5 and 6 only, with no age restrictions; mean age was 15.6 years for males and 14.6 years for females. Eleven pupils refused to participate (<1%), and the results of another 28 (<1%) were excluded from analysis as interviewers noted that respondents were illiterate. The format of the questionnaire was simple, all responses were closed, and conceptually complex questions such as those involving ranking or multiple answers were avoided. The 1–2 hour ASCQ interview was administered to same sex groups of 20 by a young, same sex research assistant who read the 87 questions and related answer options aloud, so pupils could tick or write responses independently.

Respondents generally had the option of using Swahili (the national language) or Sukuma (the main local language) for both survey interviews, except that some FFQ interviewers did not speak Sukuma so could not read those standardised translations. Only 18 questions were asked in the same or very similar ways in both interviews. The ASCQ had more questions on sociodemographic characteristics, future expectations, and sexual behaviour, while the FFQ had more knowledge and attitude questions. Of the 6709 adolescents who participated in the ASCQ survey, 4958 also participated in the FFQ (the ASCQ-FFQ cohort).

From September 1999 to May 2000, two Swahili speaking graduate researchers and two Swahili/Sukuma speaking secondary school graduates conducted in-depth interviews with 74 participants in the biological marker/FFQ survey: 51 randomly selected (including one who tested positively for *N gonorrhoeae*), three pregnant, and 20 who had had positive HIV results in preliminary tests. Interviewers made multiple visits if necessary to locate in-depth interview respondents; only one (1%) of these cohort members refused to participate in an in-depth interview. Interviewers did not know the STI or pregnancy status of in-depth interview respondents. They spent 1–2 hours chatting, strolling, playing games, and sharing snacks with respondents to build rapport before the semistructured in-depth interviews. Respondents chose Sukuma or Swahili for their interviews, which were conducted away from both home and school settings, were recorded, and took 1–2 hours to complete. During early interviews, a technique was developed to clarify sexual histories using sketches of sexual partners with key identifying characteristics. For example, if a participant mentioned a new sex partner, the researcher would sketch a simple drawing of that individual with identifying characteristics—such as a hoe if he was a farmer, a chicken if he sold poultry, a book if he was a student—and the colour of his favourite shirt. This process enabled clarification of the sequence of partners, for example if the participant corrected herself to say that there was a partner prior to the first one mentioned, the sequence of sketches could be rearranged.

From July 1999 to June 2000, the in-depth interview researchers also carried out 56 person weeks of participant observation, mainly in six MkV trial villages. Each of these individuals lived with the family of an MkV trial participant for a seven week period, usually sharing a room (and often bedding) with other young people of the same sex. They accompanied young people in their daily activities, such as fetching water, preparing meals, going to market, farming, fishing, burning charcoal, socialising, and attending special

events. Participant observation researchers wrote daily field notes in English or Swahili.

Follow up biological marker/FFQ^{13 16} and ASCQ¹⁵ surveys and participant observation and in-depth interview rounds, took place from 2000–2002. As quantitative and qualitative data were not collected contemporaneously, a series of responses to questions about an individual's sexual experience are considered reliable if they were consistently affirmative, consistently negative, or were first consistently negative and then changed to become consistently positive during the 1998–2000 formal interview series (that is, 1998 ASCQ, 1998 FFQ, 1999–2000 in-depth interview, 2000 ASCQ, and 2000 FFQ). A series of responses are considered to be invalid if they were unreliable, and/or were contradicted by a positive biological marker or a participant observation report of sexual activity.

The research process was assessed by regular supervisory field visits, staff check-in meetings, field staff reports, and data quality checks. All quantitative data were double entered, verified, and cleaned using Dbase Version 4 (Borland Corporation, California, USA), and analysed using Stata Version 7 (Stata Corporation, College Station, TX, USA). McNemar's test was used to compare proportions for paired data, unless there were too few reports, in which case a binomial test was used.²¹ All qualitative data (approximately 2.5 million words of in-depth interview recordings and 750 000 words of participant observation field notes) were transcribed, translated into English, and coded using 32 broad codes in NUD*IST Version 4 (QSR International Pty Ltd, Melbourne, Australia).

Before beginning the research, the study proposal was reviewed and approved by the London School of Hygiene and Tropical Medicine Ethics Committee, the Tanzanian Medical Research Coordinating Committee, the Tanzanian Ministry of Health, and the Tanzanian Ministry of Education and Culture. School committee chairs, head teachers, and pupils gave written consent for trial participation, and parents were given the opportunity to withdraw their children. Respondents gave verbal consent prior to in-depth interviews, and village authorities and host households gave verbal consent for participant observation research.

RESULTS

Problems with data collection

The research generally proceeded smoothly for all methods, but difficulties will be focused on here. For the biological marker method, two problems involved the limits of tests and logistical delays in receiving results. For example, 23 people initially tested positive for HIV using the urine HIV antibody tests, and these data were the basis for selection of 20 HIV+ individuals to participate in in-depth interviews in 1999–2000. However, a series of more rigorous blood tests in the 2000 and 2001–2002 follow up surveys found that in fact only 8/23 (and 7/20 of those selected for in-depth interviews) were truly HIV+. In an earlier publication¹⁴ from this trial the authors did not at the time have access to results of all subsequent biological tests, and true positives were overestimated (14 rather than eight out of 23).

In FFQ interviews the quick pace of questions and answers may have meant that some respondents had insufficient time to reflect before answering. In addition, the use of Sukuma varied greatly depending on each FFQ interviewer's language fluency. The vast majority of ASCQ respondents seemed to complete their questionnaires without difficulty, and generally provided plausible results. However, 7% of respondents were subsequently found to have selected all first or all last answer options in sections for which this was illogical, so their responses were partially or entirely excluded from

analysis.¹⁵ In addition, some respondents were frustrated by the length of the interviews.

Efforts to build rapport with respondents before in-depth interviews were partially effective, and many respondents seemed to find the sketches of sexual partners helpful. Nonetheless, interviewees tended to be awkward, responding with only a few words and rarely making expansive contributions. Much of the interview was devoted to establishing basic information, such as whether a respondent had ever had sex, and then to understanding the circumstances surrounding his or her first sexual experience. Attempts to ascertain broader sexual behaviour patterns were often superficial, and reports about frequency of sex or total number of partners sometimes changed within the course of the interview.

The participant observation researchers integrated quickly and effectively into households and villages. The Sukuma speaking participant observation researchers were particularly successful at this because they were so similar in background to their participant observation household that they were not treated very differently from others of their age and sex. They could also more readily participate in and document informal discussions and exchanges, as these often took place in Sukuma. Participant observation researchers sometimes found that young people hid sexual relationships from them in the early days of a visit, but this would change very quickly as they shared activities and developed friendships. However, it was difficult for researchers to spend much time with MkV trial participants other than those in the researcher's household, because most of them were in school. Even when pupils were out of school, it was not always easy to establish close relationships with them. Although young people of the same sex sometimes shared chores and socialised in large groups of varying ages, it would have been unusual for people in their 20s, like the participant observation researchers, to specifically seek out and socialise at length with same sex individuals 5–10 years younger than them. During participant observation, MkV trial participants were not systematically questioned about their experience of sex, and participant observation researchers did not systematically document their impressions of each MkV trial participants' sexual experience based on their observations and informal discussions, so such data were only collected opportunistically.

ASCQ-FFQ cohort members with biological markers

In the ASCQ-FFQ cohort, 40% of respondents (56% of males and 22% of females) reported sex in the ASCQ, while 38% (52% of males and 23% of females) reported it in the FFQ (table 1). However, only 62% of males and 41% of females who reported sex did so in both questionnaires. Sensitive sexual behaviours, such as having more than one sex partner in the past four weeks, or one's first partner being a teacher or a stranger, were reported by a higher proportion of respondents in the ASCQ than in the FFQ. A detailed comparison of results for the ASCQ-FFQ cohort is published elsewhere.¹⁵ This section focuses on data for respondents who tested positively for biological markers.

Twelve male and 49 female ASCQ-FFQ cohort members (1%) tested positively for one or more biological markers (table 1). Among males with biological markers, 75% reported sex in the ASCQ and 58% in the FFQ, compared with 31% and 45% for females, respectively. In both surveys, higher proportions of respondents with biological markers reported sex than those without biological markers, but this trend was only significant for females in the FFQ survey ($p < 0.001$). Only 58% of males and 29% of females with biological markers reported ever having had sex in both interviews; most of those remaining denied it in both

Table 1 ASCQ-FFQ cohort: proportions reporting sex in 1998 assisted self-completed questionnaires (ASCQs) and in 1998 face to face questionnaires (FFQs), according to the presence of biological markers

ASCQ-FFQ cohort	n	n (%) reporting vaginal intercourse		p Value*
		1998 ASCQ	1998 FFQ	
Males	2573	1430 (56)	1327 (52)	<0.001
Male; biological marker				
HIV+	2	1 (50)	1 (50)	NA
CT and/or NG	10	8 (80)	6 (60)	NA
Subtotal	12	9 (75)	7 (58)	NA
Male; no biological marker	2561	1421 (56)	1320 (52)	<0.001
Females	2166	480 (22)	489 (23)	0.655
Female; biological marker				
HIV+	2	0 (0)	0 (0)	NA
CT and/or NG	28	6 (21)	9 (32)	NA
Pregnant	20	9 (47)	13 (68)	0.109
Subtotal	49	15 (31)	22 (45)	0.020
Female; no biological marker	2117	465 (22)	467 (22)	0.920
Total	4739	1910 (40)	1816 (38)	0.004

*The McNemar's and binomial tests assessed the difference between (a) the number of people who had reported sex in the ASCQ but not the FFQ, and (b) the number of people who reported sex in the FFQ but not the ASCQ. The tests were not applicable when the value for (a) or (b) was zero; for example, none of the females with CT and/or NG reported sex in the ASCQ but not in the FFQ.
CT, *Chlamydia trachomatis*; NG, *Neisseria gonorrhoeae*; NA, not applicable.

surveys. Females with biological markers who reported sex in only one interview (most of whom were pregnant) were significantly more likely to do so in the FFQ than in the ASCQ ($p = 0.020$). Of those who reported sex at all, females with biological markers were significantly more likely than other females to report sex in one interview only ($p = 0.047$).

In-depth interview respondents

In 1998–2000, reported sex data were collected from the 74 in-depth interview respondents as many as five times. Results were excluded for one randomly selected male who was discovered to have taken the name of a classmate as an alias. In the 1998 ASCQ and FFQ surveys, 45% and 55% of male respondents, respectively, reported sex; this proportion increased to 58% in the in-depth interviews and 70% and 57% in the 2000 ASCQ and FFQ, respectively (table 2). The proportion of females without biological markers reporting sex increased substantially in all three of the 1999–2000 interviews. In contrast, of the six females with HIV or *N gonorrhoeae*, only one reported sex in any of the surveys, but five reported it in an in-depth interview.

If biological markers are used to validate this interview series externally, 32% of respondents provided unreliable responses, while an additional 8% provided reliable but invalid responses (a variable that could only be assessed for those who tested positively for biological markers) (table 3). Nine (82%) of the 11 respondents with biological markers provided an invalid series of responses; however, if no biological marker data had been available, only three (27%) would have had inconsistent and thus clearly invalid reports, similar to the 32% inconsistent reports for respondents without biological markers.

Participant observation informants

Participant observation data revealed a great deal about the complex nature of young people's sexual behaviour in rural Mwanza. The data suggest that rural youth experience many

Table 2 In-depth interview respondents: aggregate comparison of reported sex in 1998–2000 in-depth interviews, assisted self-completion questionnaires (ASCQs), and face to face questionnaires (FFQs)

In-depth interview respondents	n	Number (%) reporting vaginal intercourse				
		1998 ASCQ n = 43	1998 FFQ n = 73	1999–2000 in-depth interview n = 73	2000 ASCQ n = 56	2000 FFQ n = 62
Males	31	11/20 (55)	14/31 (45)	18/31 (58)	13/23 (57)	19/27 (70)
Male; biological marker	2	1/2 (50)	1/2 (50)	2/2 (100)	2/2 (100)	2/2 (100)
Male; no biological marker	29	10/18 (56)	13/29 (45)	16/29 (55)	11/21 (52)	17/25 (68)
Females	42	3/23 (13)	5/42 (12)	30/42 (71)	17/33 (52)	20/35 (57)
Female; biological marker:						
HIV+	5	0/2 (0)	0/5 (0)	4/5 (80)	0/3 (0)	1/3 (33)
NG+	1	0/1 (0)	0/1 (0)	1/1 (100)	0/0 (NA)	0/0 (NA)
Pregnant	3	0/2 (0)	2/3 (67)	3/3 (100)	2/2 (100)	2/2 (100)
Subtotal	9	0/5 (0)	2/9 (22)	8/9 (89)	2/6 (33)	3/6 (50)
Female; no biological marker	33	3/18 (17)	3/33 (9)	22/33 (67)	15/27 (56)	17/29 (59)
Total	73	14/43 (33)	19/73 (26)	48/73 (66)	30/56 (54)	39/62 (63)

NA, not applicable.

contradictory social norms (for example, school pupils should be abstinent) and expectations (for example, normal restrictions on sex are relaxed at festivals), and that young people manage such contradictions by carefully concealing their sexual relationships.¹⁷ Although participant observation provided only indicative evidence of age at first sexual intercourse, it was widely reported that young people start to have sex at very early ages, and most were estimated to have had sex by the age of 15 years (Pratt *C et al*, unpublished data). Participant observation research found that men almost always provide a gift or money to young women in exchange for sex,¹⁸ and female participant observation researchers estimated that a quarter of young women had been pressured into sex after accepting a gift, and half had been pressured into it by intermediaries.¹⁹

Such general participant observation findings suggest that the vast majority of in-school and out-of-school 14–17 year olds in rural Mwanza have already had sex. However, the

participant observation research was not originally intended to collect systematic, quantitative data, and the opportunistic participant observation reported sex data are limited. Nonetheless, for the purpose of this multimethod comparison they are considered here. In 1999–2000, the four participant observation researchers identified a total of 402 informants of various ages, at least 45 (11%) of whom participated in the MkV trial. “Informant” here means anyone for whom information was documented; the amount of information varied from very brief (for example, someone was seen playing cards) to detailed and lengthy. Specific experience of sex, or lack thereof, was recorded for 24 (53%) of the MkV trial participants who were participant observation informants: two (4%) told a researcher that they had never had sex, 10 (22%) reported that they had had sex, and 12 (27%) were reported to have had sex by other informants. When names were used in third person participant observation reports, it usually involved a widely reported, scandalous story about a young woman becoming pregnant and/or having an abortion (four), having more than one boyfriend (two), having a teacher as a sexual partner (two), or being caught and punished with a lover (two). The exceptions were stories about an MkV peer educator reported to have a boyfriend, and a young woman who unhappily married a man of her parents choosing during the participant observation.

None of these 24 individuals tested positive for a biological marker in the 1998 and 2000 surveys. Of those individuals who participated in the 2001–2002 biological marker/FFQ survey, eight out of nine (89%) with third person participant observation reports of sex tested positive for one or more biological markers (seven had two, one had one). This was the case for only two of the 10 (20%) who directly reported sex in the 1999–2000 participant observation, and one of the two (50%) who denied it.

Of those individuals for whom there were participant observation reports of sex, and who participated in one or both subsequent 2000 surveys, 53% denied having had sex. The two who denied sex in the participant observation also did so in all four 1998–2000 surveys. If participant observation reports are used to validate sexual behaviour reported in

Table 3 In-depth interview respondents: individual reliability of reported sex in 1998–2000 in-depth interviews, assisted self-completion questionnaires (ASCQs) and face to face questionnaires (FFQs), with validation using biological markers

Reported vaginal intercourse	Number (%) of males	Number (%) of females	Total (%)
Biological marker—positive	n = 2	n = 9	n = 11
Consistently reported	1 (50)	1 (11)	2 (18)
Denied earlier and reported later	1 (50)*	4 (44)	5 (46)
Consistently denied		1 (11)	1 (9)
Inconsistently reported		3 (33)	3 (27)
Biological marker—negative	n = 29	n = 33	n = 62
Consistently reported	7 (24)	1 (3)	8 (13)
Denied earlier and reported later	5 (17)	1 (3)	19 (31)
Consistently denied	7 (24)	8 (24)	15 (24)
Inconsistently reported	10 (35)	10 (30)	20 (32)

*Italicised figures represent results that are inconsistent and/or invalid given biological marker results.

the four 1998–2000 surveys (table 4), 27% of participant observation informants had unreliable responses, and another 18% had reliable but invalid responses (table 5).

Example involving all five methods

In the entire MkV trial, there was only one person (code name KG) for whom sexual behaviour data were collected via all five methods. KG did not test positive for any biological markers, and denied having had sex in all surveys. Her 2000 FFQ interview took place one month after her in-depth interview and early in the participant observation visit. The same researcher conducted KG’s Sukuma in-depth interview and the participant observation in her village, and KG reported sex in both methods.

In her in-depth interview, KG first denied having ever had sex, then reported having had only one sex partner, then only two, and finally only three; she also said that she had had sex only once or twice with each partner. During the participant observation, the researcher and KG shared chores and attended activities such as video shows, *ngoma* (drumming/dancing events), and Easter celebrations together. During those times, KG discussed three of her current sexual partners, none of whom were the same as those mentioned in her in-depth interview. An excerpt from the researcher’s field notes on this follows:

“KG said that during the rainy season she farms with her younger brothers and sister. ... She said she also gets money by making love with boys, who are K and O. ... She said she and her younger siblings are all given body and cooking oils free from her lover, Z. ... She said she had made love with Z eight times right in the house where she lives with her grandmother.”

Comments to the researcher by K, O, and Z, and the researcher’s observations of KG’s interactions with them, supported KG’s participant observation reports. In addition to these relationships, the participant observation researcher repeatedly observed KG engaging in negotiations which young women often have with men to obtain something in exchange for sex—for example, payment to enter a video show, or “Hallelujah” money given on Christian holidays.¹⁸

Table 4 Participant observation informants: aggregate comparison of reported sex in 1998–2000 participant observation, assisted self-completion questionnaires (ASCQs) and face to face questionnaires (FFQs)

Participant observation informants*	n	Number (%) reporting vaginal intercourse				
		1998 ASCQ	1998 FFQ	1999–2000 participant observation	2000 ASCQ	2000 FFQ
Male: participant observation self report	6	2/4 (50)	2/6 (33)	5/6 (83)	4/4 (100)	4/6 (67)
Female: participant observation self report	6	0/5 (0)	1/6 (17)	5/6 (83)	3/4 (75)	2/5 (40)
Female: participant observation third person report	12	3/8 (38)	2/10 (20)	12/12 (100)	3/8 (38)	5/8 (63)
Total	24	5/17 (29)	5/22 (23)	22/24 (92)	10/16 (63)	11/19 (58)

*MkV trial participants for whom there were specific reports of sex during the 1999–2000 participant observation.

Table 5 Participant observation informants: individual reliability of reported sex in 1998–2000 assisted self-completion questionnaires (ASCQs) and face to face questionnaires (FFQs), with validation using participant observation reports

Reported vaginal intercourse in 1998–2000 ASCQ and FFQ	Number (%) of males		Number (%) of females	
	Self-report* (n = 5)	Self-report* (n = 5)	Third person report (n = 12)	Total (%) (n = 22)
Consistently reported	1 (20)		1 (8)	2 (9)
Denied earlier and reported later	2 (60)	3 (60)	5 (42)	10 (46)
Consistently denied	1 (20)	1 (20)	2 (17)	4 (18)
Inconsistently reported	1 (20)	1 (20)	4 (33)	6 (27)

*Data for the one male and one female trial participants who denied sex are not included here. Italicised figures represent results that are inconsistent and/or invalid given participant observation reports.

DISCUSSION

As a methodological comparison of the validity of sexual behaviour data, this study had a number of strengths and limitations. Strengths include the extraordinary diversity and the number of methods employed, as well as their comparative use within one population. However, a key limitation is that the five methods were not employed contemporaneously. Data collection took place within a period of time spanning almost two years. Differences in results may thus relate to greater respondent maturity or experience, rather than the relative effectiveness of research methods. However, if that were the case, one would expect that the methods employed last would have found the greatest prevalences of sexual behaviour, reflecting greater sexual experience and social independence of the respondents. In fact, sexual behaviour estimates were equal or highest in the 1999–2000 participant observation and in-depth interview research, and almost all of that research took place prior to or contemporaneously with the 2000 ASCQ and FFQ surveys, suggesting that the methods themselves played a role in the different findings.

Biological markers

The biological test algorithms and markers used in this study were of limited value in assessing sexual behaviour in this very low prevalence population. The tests required a great deal of time, expense, and technological expertise to implement. Although such effort may have ensured a high test specificity and sensitivity, the low prevalences found in this young population meant that false positives outnumbered true positives. Discrepant HIV positive results made it difficult to interpret and use biological marker data during the trial, and ultimately substantial additional time and expense were needed to confirm the true positive results. In addition, even when biological marker results were confirmed, on their own they only revealed that a very small proportion of this population very probably had sex at some time in the past. Little was revealed about the broader sexual behaviour of those who tested positive, such as frequency of sex or number of partners, and nothing was revealed about the sexual behaviour of the vast majority of young people, who tested negative.

However, the biological marker data clearly demonstrated that some FFQ and ASCQ sexual behaviour reports were invalid, in a way that would have been difficult to achieve by any other method. Tests of a more prevalent biological

marker (for example, HSV-2) or use of a different algorithm (for example, immediately testing serum rather than urine) may have been much more useful in this population. However, a highly sensitive and specific urine test for HSV-2 was not available in 1998, and researchers were reluctant to take blood from trial participants at that time, because of their youth and the relatively invasive nature of the procedures involved.

Biological marker data may also prove useful in the identification of risk factors, as preliminary analysis of in-depth interview results has identified several variables that may be associated with HIV and other STIs, including not living with parents; sexual partners being mobile, older, self-employed, and having more disposable income; having intercourse without foreplay; and having less knowledge and more misconceptions about HIV/AIDS (Wamoyi J *et al*, unpublished data). The number of 1999–2000 in-depth interview respondents with STIs was low, making it difficult to draw firm conclusions about such findings, particularly given questions about the validity of self-reported information in general. Additional risk factor analysis using the 2001–2002 in-depth interview data would thus be useful, as 69 of the original 73 in-depth interview respondents were interviewed again at that time (that is, four HIV positive, 17 infected with one or more other STIs, and 48 HIV/STI-negative), while an additional 20 individuals who tested positive for HIV in the 2001–2002 follow up biological marker survey were interviewed.

Finally, although a high proportion of respondents with biological markers provided invalid responses at some point in the 1998–2000 interview series, we cannot conclude that they provided less honest reports than respondents who were biological marker negative. The presence of a biological marker may be evidence of having had sex, but lack of a biological marker is not evidence of never having had sex. When all self-reported data were compared at an individual level, it was notable that the reliability of reports was similar for biological marker negative and positive respondents.

Formal interview methods: ASCQ, FFQ, and in-depth interview

Both the ASCQ and FFQ methods also displayed limited effectiveness in collecting sexual behaviour data. Although each had individual strengths and weaknesses, neither was clearly superior to the other. Importantly, both survey methods provided plausible and even similar reports at an aggregate level, but individual level analysis revealed that a sizeable portion of the trial participants reported sex in one survey but denied it in the other. In addition, while higher proportions of respondents with biological markers reported sex in both surveys than did respondents without biological markers, this only amounted to three quarters of the males with biological markers, and less than half of the females with biological markers.

Despite their limitations, the survey methods may have value in identifying and monitoring general trends. For example, longitudinal analysis of the FFQ data found that intervention status was significantly associated with improved knowledge, as well as more desirable reported attitudes and behaviour (Ross D *et al*, unpublished data). It is not possible to definitively interpret such attitude and behaviour trends, as intervention participants were aware of programme goals and the data may thus reflect false as well as honest reports of change. However, these trends are useful in indicating intervention impact of some kind, and the related trend in improved sexual health knowledge is promising, as social desirability bias is far less of a concern in knowledge data.

The FFQ method was more effective than the ASCQ method in promoting honest reports among females with biological markers, particularly those who were pregnant. Some of these women may have been visibly pregnant, and thus may not have felt that they could deny having had sex in the FFQ interview as they could in the ASCQ. Two limitations in comparing the in-depth interview and survey interview data in this study are that not all in-depth interview respondents participated in all surveys, and the different selection criteria for the ASCQ (that is, years 5–6; any age) and the FFQ (that is, years 4–6; 14 years and older) surveys may have distorted comparisons. Nonetheless, it seems clear that the in-depth interviews were more effective in collecting data from the six females with STIs, five of whom reported sex in an in-depth interview, but only one of whom did so in any of the four surveys. The greater rapport and the lengthy, semistructured nature of the in-depth interviews provided more opportunity to clarify and correct contradictory reports, which was likely to have improved data validity. In addition, the random selection of in-depth interview respondents, intensive efforts to trace them, and the relatively large number of respondents interviewed for a qualitative method, were all likely to have maximised the representativeness of these data.³

However, even though efforts were made to encourage representativeness and openness, personal reflection and dialogue about sexual behaviour may have been too culturally alien for some respondents to speak honestly, particularly as they were school pupils who were generally accessed through school authorities. As is true in much of the world, sexual behaviour is rarely talked about openly with friends and loved ones in rural Mwanza, let alone with slightly older, more educated, and urban interviewers. School pupils are not accustomed to speaking at length to people in authority on any topic, and they are usually severely punished if they are found to have had sex.¹⁷ Thus while some reports of sex were higher in the in-depth interviews than in the ASCQ and FFQ surveys, there is good reason to believe that sensitive behaviours were also underreported in that method. In addition, although the in-depth interviews were intended to be semistructured, in practice they were highly structured. As such, they generally did not achieve one of the main goals of qualitative research—that is, to have respondents express their own world view, and new perspectives on sexual relationships that may not have been previously identified by researchers.

Participant observation

The participant observation researchers neither systematically asked MkV trial participants whether they had ever had sex, nor systematically documented their own impressions of this question for each MkV trial participant, so multimethod comparison of individual level data was difficult. In later participant observation conducted in 2000–2002, researchers prioritised developing relationships with more MkV trial participants, and then asking more systematically about their experience of sex. They found it was indeed possible to collect much more data of this nature than was collected opportunistically in 1999–2000, suggesting that this approach has potential. However, such potential does not mean that participant observation should be reduced to a means of quantitative data collection, given that the strengths of anthropological research are its discursive features of description and interpretation.^{22 23}

A number of additional problems arise in comparing the participant observation reported sex data with those of other methods. Firstly, participant observation is somewhat arbitrarily taken as the “gold standard” of reported sexual behaviour against which to compare other methods, when in

fact some MkV trial participants may have reported sex in a formal interview and would have chosen to deny it if asked in the participant observation. Secondly, this is the only method which included third person reports, which are not comparable to first person reports. These third person reports of scandalous sexual behaviour are interesting in their own right, however, given that almost all of the individuals for whom there were such reports actually tested positive for two biological markers 2–3 years later. Such reports could be explored as a low cost alternative to a large scale survey in identifying high risk youth in rural areas, although the ethical implications of such an approach would first need close examination.

In KG's example, it is not surprising that someone who reported sex in an in-depth interview would speak openly about it when meeting the same researcher in her village the following month. However, participant observation data suggested that she had had sex much more often, with substantially more partners, than she had reported in her in-depth interview. Although this is only one example, it typifies findings from the 56 person weeks of participant observation research. Broader participant observation data suggest that most male and female in-school 14–17 year olds in rural Mwanza have already had sex. Participant observation data also found a powerful culture of secrecy surrounding sexual activity,¹⁷ helping to explain the many inconsistencies seen in the three formal interview methods, and suggesting that the prevalences for reported sex found using those methods are serious underestimates, particularly for females. This point is illustrated by an exchange the Sukuma male researcher witnessed while out walking with two male MkV trial participants:

“... MF interrupted KL and told him to be truthful, and not to hide his lovers like he is used to, when the reality is that he is experienced and has many lovers. KL said that at least he is a bit more truthful than MF, who, when answering MEMA kwa Vijana [survey] questions, said that he had never ever had sex, when he is actually an expert in these matters. MF in turn asked KL why even KL had said [in the survey] that he had only had sex once, when he too is an expert and has a child by now ... MF said that he answered the MEMA questions by claiming he had never had sex even once, because he thought that the teachers might follow-up on those who have lovers and expel them from school.”

Participant observation duration, geographic range, and researcher differences were all intended to maximise the representativeness of the qualitative data collected via this method. This seems to have been successful, as the wealth of data collected from different sources enabled clear distinctions between common and uncommon behaviour. Participant observation seems to have been the most effective and accurate method in documenting the nature, complexity, and extent of youth sexual behaviour in rural Mwanza.

CONCLUSION

There is no “gold standard” in sexual behaviour research. The very sensitive nature of sexual behaviour in this context makes collection of accurate data extremely challenging. Ideally, multiple, complementary research methods should be used in concert, such as biological markers and in-depth interviews to understand high risk behaviour, or a survey and participant observation to obtain both broadly representative and detailed in-depth information. Multiple approaches may be especially important among school-going adolescents,

particularly girls, for whom secrecy about sexual activity is the norm.

However, it is not always feasible to implement more than one research method. This study with young adolescents in rural Mwanza Region allowed the direct comparison of the findings from five methods (ASCQ, FFQ, biological marker, in-depth interview, and participant observation), in a setting in which the prevalence of biological markers was very low, and the research participants were very secretive about their sexual behaviour. In this context, participant observation was the most useful method, as it seems to have provided relatively subtle and valid data.

Some researchers may argue that well conducted, qualitative research is difficult to do as it requires highly trained staff. In this study, two secondary school graduates conducting qualitative research had only two weeks more training than 12 secondary school graduates conducting quantitative interviews, and yet the qualitative data they collected were arguably as valuable as those collected by graduate researchers. In addition, critics of qualitative research may argue that such data cannot be used to monitor trends, as they lack the representativeness of a large sample. However, the representativeness of survey data also comes into question when a population is very secretive about the research topic. The large amount of inaccurate reporting in our surveys suggests that, in such cases, observed trends may sometimes be widely distorted. This study suggests that by carefully selecting a moderate number of representative sites, researchers, and subjects, qualitative research can collect meaningful and representative data in a practical, effective, and economical way. If such data are collected at systematic intervals over an extended period of time, they may also be valuable in the identification and monitoring of trends.

ACKNOWLEDGEMENTS

This was a collaborative study involving the Tanzanian National Institute for Medical Research, the African Medical and Research Foundation, the London School of Hygiene and Tropical Medicine, the UK Medical Research Council's Social and Public Health Sciences Unit, and the Tanzanian Ministries of Health and of Education and Culture. We are grateful to Mwanza regional, ward, village, and subvillage government, school and health authorities for their assistance. Special thanks are also due to the many field research, laboratory, data entry, and administrative staff, who worked very hard to ensure the research was completed successfully. Finally, we thank the young people of Mwanza Region, whose helpful participation made this study possible.

The research reported here was funded by the UK Medical Research Council, the European Commission, Development Cooperation Ireland, and the UK Department for International Development.

CONTRIBUTORS

The principal investigators of this study include DR, DW, FM, and RH, who held the study grants. The study was designed by DR, DW, and RH, with input from MP, JC, JT, FM, and AO. The study field tools were designed by MP, DR, DW, JC, JT, and FM; the field work was supervised by MP, DR, JT, and FM; the graduate researcher qualitative field work was conducted by GM and JW; and the laboratory work was supervised by JC. Data management was done by JT and AA, and analysis was conducted by MP and AA. This paper was written by MP; all authors reviewed and commented on it before finalisation.

Authors' affiliations

M Plummer, D Ross, J Todd, A Anemona, A Obasi, R Hayes, London School of Hygiene and Tropical Medicine, London, UK

M Plummer, D Ross, J Changalucha, G Mshana, J Wamoyi, J Todd, A Anemona, F Moshia, National Institute for Medical Research, Mwanza, Tanzania

M Plummer, D Ross, J Wamoyi, J Todd, A Anemona, A Obasi, African Medical and Research Foundation, Mwanza, Tanzania

D Wight, Medical Research Council, Glasgow, UK

REFERENCES

- 1 **Fenton KA**, Johnson AM, McManus S, *et al*. Measuring sexual behaviour: methodological challenges in survey research. *Sex Transm Infect* 2001;**77**:84–92.
- 2 **Seal DW**, Bloom FR, Somlai AM. Dilemmas in conducting qualitative sex research in applied field settings. *Health Educ Behav* 2000;**27**:10–23.
- 3 **Smith PG**, Morrow RH. *Field trials of health interventions in developing countries: a toolbox*, second edn. Oxford: Macmillan, 1996.
- 4 **Parker RG**, Herdt G, Carballo M. Sexual culture, HIV transmission, and AIDS research. *J Sex Res* 1991;**28**:77–98.
- 5 **Wight D**, West P. Poor recall, misunderstandings and embarrassment: interpreting discrepancies in young men's reported heterosexual behaviour. *Cult Health Sex* 1999;**1**:55–78.
- 6 **Ross DA**, Vaughan JP. Health interview surveys in developing countries: a methodological review. *Stud Fam Plann* 1986;**17**:78–94.
- 7 **Bleek W**. Lying informants: a fieldwork experience from Ghana. *Popul Dev Rev* 1987;**13**:314–22.
- 8 **Catania JA**, Turner H, Pierce RC, *et al*. Response bias in surveys of AIDS-related sexual behavior. In: Ostrow DG, Kessler RC, eds. *Methodological issues in AIDS behavioral research*. New York: Plenum Press, 1993:133–62.
- 9 **Fishbein M**, Pequegnat W. Evaluating AIDS prevention interventions using behavioral and biological outcome measures. *Sex Transm Dis* 2000;**27**:101–10.
- 10 **Schachter J**, Chow JM. The fallibility of diagnostic tests for sexually transmitted diseases: the impact on behavioral and epidemiologic studies. *Sex Transm Dis* 1995;**22**:191–6.
- 11 **Wight D**, Barnard M. The limits of participant observation in HIV/AIDS research. *Practicing Anthropology* 1993;**15**:66–9.
- 12 **Pool R**. Anthropological research on AIDS. In: Ng'weshemi J, Boerma JT, Bennett J, *et al*, eds. *HIV prevention and AIDS care in Africa: a district level approach*. Amsterdam: Royal Tropical Institute, 1997:69–83.
- 13 **Hayes R**, Changalucha J, Grosskurth H, *et al*. MEMA kwa Vijana: a randomized controlled trial of an innovative adolescent sexual health intervention programme in rural Tanzania: 1 rationale and trial design. 15th meeting of the International Society for Sexually Transmitted Disease Research, Ottawa, Canada; 27–30 July 2004. Abstract Guide p253, no 695.
- 14 **Todd J**, Changalucha J, Ross DA, *et al*. The sexual health of pupils in years 4 to 6 of primary schools in rural Tanzania. *Sex Transm Infect* 2004;**80**:35–42.
- 15 **Plummer ML**, Wight D, Ross DA, *et al*. Asking semi-literate adolescents about sexual behaviour: The validity of assisted self-completion questionnaire (ASCQ) data in rural Tanzania. *Trop Med Int Health* 2004;**9**:737–54.
- 16 **Ross DA**, Changalucha J, Plummer ML, *et al*. MEMA kwa Vijana, a randomized controlled trial of an adolescent sexual and reproductive health programme in rural Mwanza, Tanzania: 3 results—knowledge, attitudes and behaviour. 15th Meeting of the International Society for Sexually Transmitted Disease Research, Ottawa, Canada; 27–30 July 2004. Abstract Guide p253, no 698.
- 17 **Wight D**, Plummer ML, Mshana G, *et al*. Divergent sexual norms for young people in rural northern Tanzania. 12th International Conference on AIDS/STD in Africa, Ouagadougou, Burkina Faso, 9–13 December 2001. Abstract Book p301, no 12PT6–526.
- 18 **Mshana G**, Wamoyi J, Plummer ML, *et al*. Sexual negotiation and decision-making among young people in rural northern Tanzania. 14th International AIDS Conference, Barcelona, Spain, 7–12 July 2002. Abstract no E11595.
- 19 **Plummer ML**, Wight D, Wamoyi J, *et al*. 2002. Sexual violence, pressure and HIV in rural Mwanza, Tanzania. 14th International AIDS Conference, Barcelona, Spain. 7–12 July 2002. Abstract no WeOrE, 1282.
- 20 **Peeling RW**, Toye B, Jessamine P, *et al*. Pooling of urine specimens for PCR testing: a cost saving strategy for *Chlamydia trachomatis* control programmes. *Sex Transm Infect* 1998;**74**:66–70.
- 21 **Siegel S**, Castellan NJ. *Nonparametric statistics for the behavioral sciences* (2nd edition). New York: McGraw-Hill, 1998.
- 22 **Lambert H**. Methods and meanings in anthropological, epidemiological and clinical encounters: the case of sexually transmitted disease and human immunodeficiency virus control and prevention in India. *Trop Med Int Health* 1998;**3**:1002–10.
- 23 **Wood K**. Questioning sex in a South African township. Paper presented at the Association of Social Anthropologists Decennial Conference on Anthropology and Science, Manchester, UK, July 2003.