Section two: Methodology

- Research design
- Instruments and scales
- Pilot study
- Ethical clearance
- Sample design
- Sample
- Data collection

2. Methodology

2.1 Research design

The study employed a triangulation of several research methods. Firstly, there was significant involvement of key stakeholders in conceptualisation and design of the study and its execution. A Technical Task Team (TTT) consisting of researchers from the HSRC-led consortium and representatives from all ELRC stakeholder groups, the DoE (national and provincial), Sadtu, Naptosa, SAOU, Natu and the SACE jointly developed the terms of reference for the study, the research proposal and an advocacy plan for both facilitating entry of fieldwork teams into educational institutions and for dissemination. This ensured a 'buy-in' into the research process by all ELRC stakeholders and the SACE, which was necessary to achieve a high response rate. The same approach was also used with FET campus/human resources managers and deans responsible for education at various universities in connection with accessing of FET-based educators and education students respectively. As much collaboration as possible took place throughout the study.

Secondly, formative research was undertaken using focus groups and key informant interview methods among educators throughout the country. These qualitative methods helped to determine the themes to include in the questionnaire.

Thirdly, there was also a once-off consultation with HIV/AIDS educators, researchers and experts from the HSRC-led consortium to acquaint the task team with research work in the area that was being undertaken in the country. This process also helped to inform the final protocol for the study.

Fourthly, and most importantly for the present report, a cross-sectional survey among educators and student educators employing the second-generation surveillance method that combines the measurement of behavioural and biological indicators within the same study was used. A behavioural risks questionnaire-based survey was conducted concurrently with HIV testing to determine the association between the two.

Fifthly, and finally, completion of the institutional questionnaire by the principal or rector involved them undertaking some archival research to look up school records to obtain the information required.

2.2 Instruments and scales

2.2.1 Individual questionnaires

The questionnaire was developed during a lengthy process: As a first step, indicators were identified based on epidemiological and prevention intervention models (FHI 2000; Mertens, Caraël, Sato, Cleland, Ward & Smith 1994; Rietmeijer, Lansky, Anderson & Fichtner 2001). Secondly, the design of the questionnaire was informed by information collected during a series of qualitative interviews (focus groups with educators and in-depth interviews with experts) that dealt with a range of education issues, including work-related aspects. This qualitative phase was conducted across the country with a cross section of educators and other experts. Thirdly, the questionnaire also reflected the needs of the DoE and the ELRC. The first draft of the questionnaire was tested during a pilot survey. Based on the feedback from the pilot survey, the questionnaire was adapted

by shortening it as well as changing some questions that proved to be problematic during the pilot survey. The actual questionnaire administered during the study consisted of the following components:

a) Biographical data of the respondent;
b) Teaching responsibilities and work load;
c) Impact of HIV/AIDS on educators and their work;
d) Absenteeism from work;
e) Morale and job satisfaction and how this was influenced by HIV/AIDS;
f) Training and support received by educators;
g) Substance use;
h) Violence within institutions;
i) Sexual behaviour of the respondent;
j) Male condom accessibility;
k) HIV/AIDS knowledge;
l) Communication about HIV/AIDS;
m) Risk perception on the part of the respondent;
n) Voluntary counselling and testing (VCT) services;
o) Stigma;
p) TB;
q) Health and medical service.

Using the educator questionnaire as a basis, a much shorter questionnaire was developed for use among students at tertiary training institutions. Obviously a number of sections dealing with work-related experiences were not applicable to students.

The final questionnaires were translated into Afrikaans, SeTswana, isiXitsonga, SePedi, isiZulu and IsiXhosa.

2.2.2 Scales used in the study

Scales were developed on the basis of a literature review, focus groups with educators, expert interviews, pilot testing, and statistical analysis (for example, item-total correlations and factor structure). Summated scales were developed by adding scores on all variables loading on a component. To verify that these items measured the same attribute, it was necessary to calculate Cronbach's alpha. If Cronbach's alpha was =>0.70, the items measured the same attribute. Since this study was exploring some of these relationships, a Cronbach's alpha of 0.60 or greater was considered sufficient to determine reliability.

The various scales and their reliability using Cronbach's alpha (α) are listed below:

- Religious involvement index (α=0.74)
- Job satisfaction scale (α=0.71)
- Job stress (α=0.52)
- Educator support index (α=0.82)
- Violence at school index (α=0.68)
- Alcohol use scale (AUDIT) (α=0.78)
- Self-efficacy scale for HIV risk behaviour (0.78)
- HIV/AIDS knowledge index (α=0.52)
- HIV and sexuality communication comfort index (α=0.61)
- HIV risk perception scale (α=0.58)

- TB social distance scale (α=0.69)
- HIV/AIDS stigma (α=0.68).

A detailed analysis is provided in Appendix 1.

2.2.3 Learning areas

During the interview educators were asked an open-ended question about what subject they are currently teaching (see right column in Table 2.1) and this was grouped into the current/future learning area (see left column in Table 2.1), except for foundation phase, according to the DoE. Many educators did not respond with particular subject but with 'foundation phase' or all subjects, and thus, 'foundation phase' was added in the subject groups.

Table 2.1: Learning areas and subject groups

Learning area	Subjects
Foundation phase	Numeracy Literacy, Communication, Linguistics, Creative writing Life skills
Foundation languages	Afrikaans, English, IsiXhosa, IsiZulu, SePedi, SeSotho, SeTswana, SiSwati, TshiVenda, XiTshonga, Ndebele
Additional languages	Arabic, Classical Greek, French, German, Gujarati, Hebrew, Hindi, Italian, Latin, Portuguese, Tamil, Telegu, Urdu
Arts and culture	Dance Studies (Dance, Dance Performance, Practical Ballet) Design (Design, Graphic Art) Dramatic Arts (History of Theatre, Costume & Literature, Speech & Drama) Music (Music, Music Performance, Anatomy & Music, Music Composition) Class Music Visual Art (Art, History of Art, Painting & Decorating, Painting, Sculpture)
Economic and management science	Accounting Business Studies (Economics, Business Economics, Entrepreneurship, E-commerce) Hospitality Studies (Hotel-keeping & Catering, Restaurant Studies, Reception Studies) Travel and Tourism Hairdressing, Beauty Care/cosmetic, Administration
Social sciences	Geography History Human and Social Science

Life orientation	Health Education, Sex Education, Religious Education, Scripture, Bible Education Civic Responsibility, Family Guidance, Life Skills, Guidance, Physical Education, Movement Education, Sports Activities Information Skills, Library/librarianship, Library Science, Media Guidance, Media-user Guidance Jewish Studies, African Studies, Biblical Studies
Mathematics	Mathematics (Functional Mathematics, Mathematics, Additional Mathematics, Commercial Mathematics) Mathematical Literacy
Natural sciences	Physical Sciences (Functional Physical Science, Physical Science) General Science Environmental Studies Maritime Studies Life Sciences (Biology, Physiology, Physiology and Hygiene) Agricultural Sciences (Gardening, Field husbandry, Equine studies, Applied Agricultural Science, Agricultural Economics, Farm Mechanics, Animal Husbandry)
Technology	Computer Studies (Computer Application Technology, Information Technology) Civil Technology (Bricklaying & Plastering, Building Construction, Plumbing & Sheetmetalwork, Woodwork, Woodworking, Technika Civil Consumer studies (Cookery & Nutrition, Home Economics, Housecraft, Needle & Clothing) Compu-typing, Typing Electrical Technology (Electrician work, Electronics, Technika Electrical, Technika Electronics) Technical Drawing, Technology, Design, Mechanical Technology (Fitting and Turning, Motor Body Repairing, Metalwork, Metalworking, Technika Mechanical, Motor Mechanics, Motor Vehicle Construction, Welding & Metalworking)
Special	Remedial Adult Basic Education and Training (ABET)
Other	Methods of Teaching, Education Didactics & Pedagogic, General Teaching Method, Outcomes-based Education, Management Skills/School Organisation Commercial Company Law, Mercantile Law, Law of Criminal Procedure and Evidence, Introduction to Criminology, SA Criminal Law Psychology, Sociology, Anthropology, Introduction to Ethnology, Research EBW Zoology, Botany Child study/Paeds/Kindergarten

2.2.4 HIV testing

Participants in this study had the option of either providing a blood specimen or a specimen of oral fluid for HIV testing. This approach was used in order to maximise participation. All specimens were linked to the questionnaire by means of a bar code, which enabled an HIV result to be linked to data from the questionnaire, but no HIV result could be linked back to any individual, thus ensuring anonymity and confidentiality.

HIV testing was conducted by Contract Laboratory Services (CLS) of the University of the Witwatersrand Medical School and CD4 testing was undertaken at the same medical school by the Department of Molecular Medicine and Haematology. Both laboratories are registered with the South African National Accreditation System (SANAS) and both participate in national and international quality control programmes. See Appendix 2 for more details.

Blood specimens were tested for HIV on the Abbott AXSYM third generation HIV 1 / 2 g0 testing system. Oral fluid specimens were obtained by using the 'Orasure' oral fluid collection device and these specimens were tested using the Vironostika HIV Uni-Form II Oral Fluid testing system.

Only a single test was conducted per specimen. In the case of blood specimens that gave borderline readings, the result was confirmed by using the Biorad HIV test. Specimens that remained borderline on the Biorad system were reported as 'indeterminate'. Oral fluid specimens that were borderline were not repeat tested and were reported as indeterminate.

2.3 Pilot study

A pilot study was conducted among 438 educators, 393 of whom also provided biological specimens. Three modes of questionnaire administration and three methods of collecting biological specimens were tested in 33 schools located in North West and Western Cape provinces. The three modes of questionnaire administration tested were self-administered, group-administered with the assistance of the interviewer, and interviewer-administered. The three methods for biological specimen collection were giving venous blood serum specimen, oral mucosa transudate (oral fluid or saliva specimen) or a choice between the two. On the basis of the findings the decision was taken to: (1) use the interviewer-administered approach as a means of obtaining better quality of data using a questionnaire; and (2) give respondents a choice of either providing a blood specimen or an oral fluid specimen, with preference for the blood. This increased the response rate for the former over the latter.

2.4 Ethical clearance

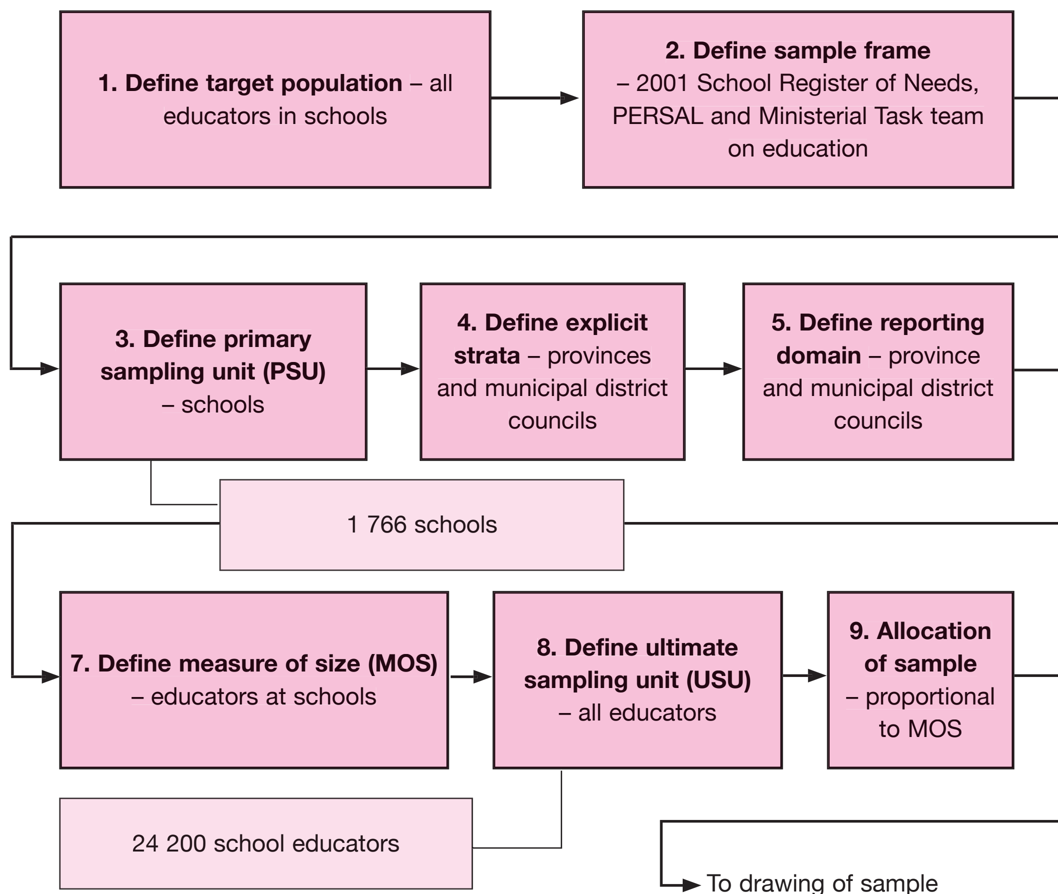
Ethical approval for conducting the study was obtained from the HSRC's Ethics Committee (Application Number REC2/20/08/030). Informed consent was obtained separately for agreeing to participate in the interview and for providing a specimen for HIV testing. In addition, the result of the HIV test for each participant was linked anonymously to

questionnaire data using bar codes. Finally, voluntary counselling and testing (VCT) for HIV testing was not provided as part of the study. Instead, those interested in finding out about their HIV status were given a referral card to go to the nearest primary healthcare centre that provided VCT services free of charge.

2.5 Sample design

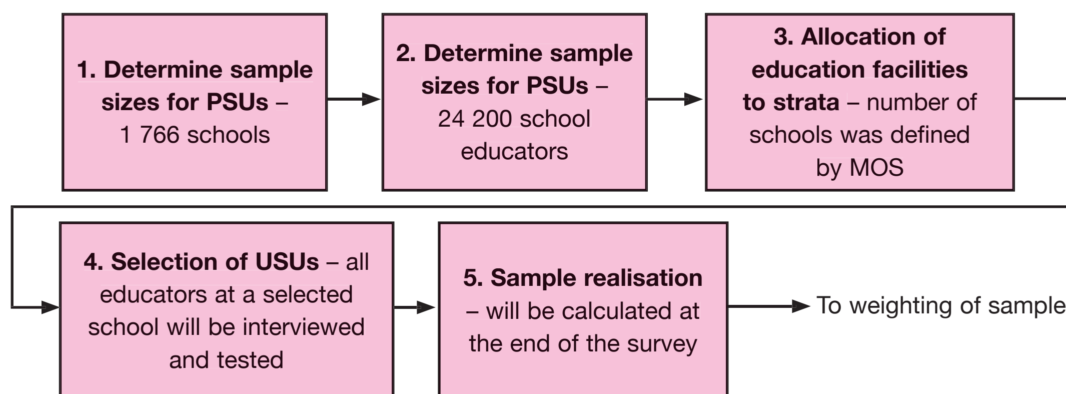
The steps used in designing and drawing the sample for the epidemiological survey are shown in Figure 2.1 and 2.2 below.

Figure 2.1: Steps in the sample design



The intention of the study was to report the results of the national survey at a provincial level and for the school sample also at the district/metropolitan council level. District councils were used because the boundaries of education regions have not been finalised for the country, and district/metropolitan councils have an important local government responsibility in terms of provision of education services.

Figure 2.2: Steps in the drawing of the sample



2.6 Sample

The target population for the study were identified as teachers at public schools, further educators at the FET colleges and students studying for an educational qualification at universities and technikons. In this report only the schools sample is presented.

Two data sets were available as potential sample frames from which a sample of educators (teachers) could have been drawn. The first was the DoE's School Register of Needs (SRN) which contained data from surveys in 1996 and 2000 and the second, a database extracted from the government's PERSAL system and made available by the firm, Price Waterhouse Coopers. The SRN data set contained records for 26 713 schools with a total of 356 749 educators. The data set obtained from Price Waterhouse Coopers contained 363 650 records pertaining to state-employed school educators and administrative staff in education offices. Due to the size of the file, the exact number of 'educators' could not be determined.

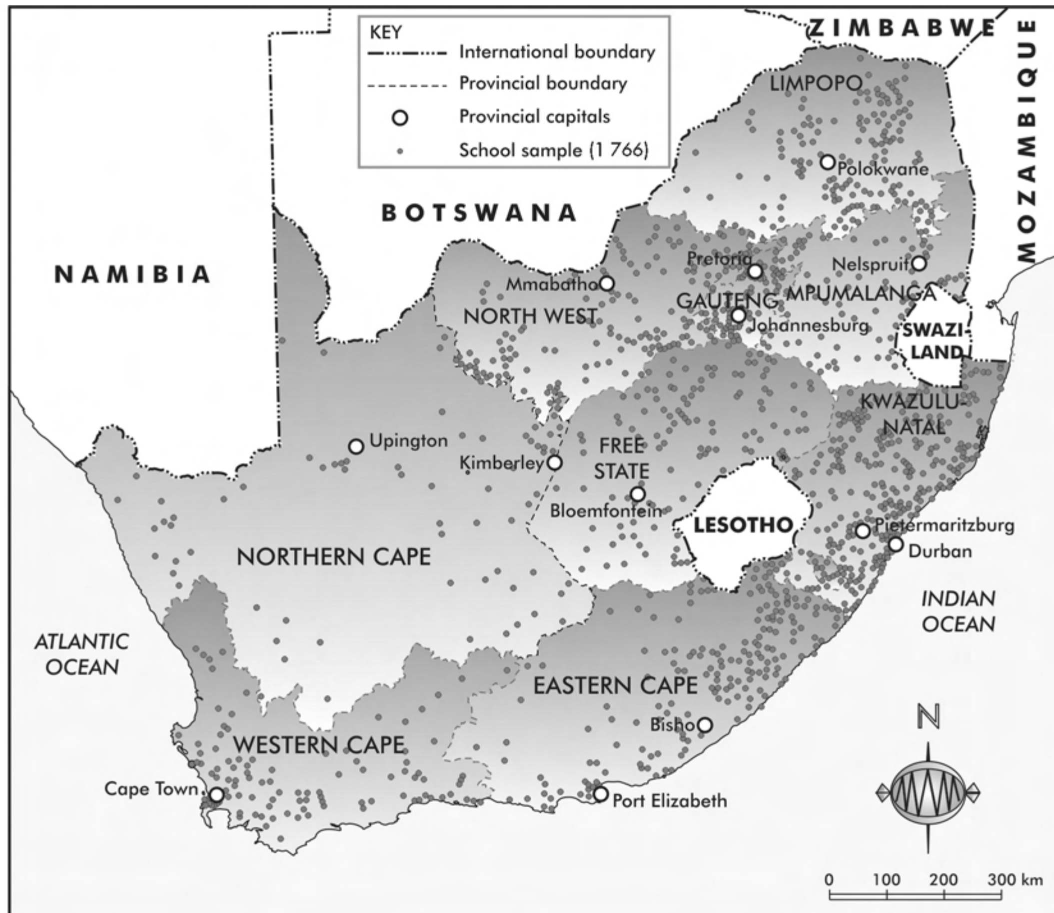
An initial sample of 1 750 schools was drawn from the SRN data set. When these were individually compared to the educator information captured in the PERSAL system, the number of state-paid educators at 248 schools could not be identified. The geographical distribution of these 'missing' schools was such that their exclusion would have resulted in compromising the representivity of the sample.

The sample was thus redrawn and a total of 2 015 schools were identified. When compared to the PERSAL data set, the most recent number of state-paid educators could not be determined for 249 schools. These schools were excluded as it did not compromise the representivity of the sample as was the case with the original sample of 1 750. The final sample, consisting of 1 766 schools, had a total of 24 200 state-paid educators as potential respondents. The provincial distributions of the schools sample are shown in Table 2.2 and Figure 2.3.

Table 2.2: Breakdown of the sample of schools by province

Province	Number of SRN educators	Number of schools
Eastern Cape (EC)	3 193	263
Free State (FS)	1 700	165
Gauteng (GT)	3 998	171
KwaZulu-Natal (KZN)	4 893	383
Limpopo (LP)	2 570	212
Mpumalanga (MP)	1 940	118
North West (NW)	1 596	179
Northern Cape (NC)	1 603	99
Western Cape (WC)	2 707	176
Total	**24 200**	**1 766**

Figure 2.3: School sample



Source: HSRC GIS Centre

2.7 Data collection

2.7.1 Recruitment and training of field workers

In total, 436 field workers, comprising trained nurses and MA and PhD research interns, were appointed to conduct the fieldwork. To conduct the interviews and specimen collections, nurses registered with the SANC were employed. Interviewers were selected on the basis of their professional status (being registered with the SANC), their availability, representivity of the various racial and language groups of the educators, representivity of the various provinces, and successful participation in the training.

Five training sessions were conducted: in Cape Town from 21–24 February 2004, in Pretoria and Durban from 23–26 February 2004, in Bloemfontein from 1–3 April 2004 and Limpopo from 5–7 April 2004.

Nurses were trained to understand the purpose of the project and the project itself, the ethics of sensitivities to asking questions on human sexuality as well as the threats to the validity of the study and how to manage them, including the need for a high response rate. In addition, they were trained on questionnaire administration, HIV testing, disposal of medical waste and standard operating procedures for needle stick injuries.

2.7.2 Access into schools

Masters and doctoral research interns served as co-ordinators for the study. They were supervised by a national field manager and a project manager, all reporting to the principal investigator who was responsible for the conduct of the study. The scientific project director oversaw scientific issues. Several co-investigators participated in various aspects of the study. In all, a team of 21 scientists worked on the project.

To facilitate access to schools, the HSRC co-ordinators were supported by district officers of the DoE and labour unions. The field teams were supported by a national field manager, a project manager who tracked progress in the field, technical staff working on data management, as well as the principal investigators and the project director who oversaw the research process.

The DoE in the various provinces assigned co-ordinators who ensured that schools were informed about the study by means of faxes, and in some cases, during meetings with school principals in their respective provinces. Some difficulties were experienced in this process because of outdated school contact details, changes in staff in the provinces, and other circumstances. However, the majority of schools knew about the study before HSRC co-ordinators or supervisors contacted them.

District officials and circuit managers of the DoE played a facilitating role in verifying the contact details and addresses of schools and also assisted co-ordinators and supervisors in setting up appointments with the schools.

In some cases appointments were made over the telephone and in others pre-visits were made to set up appointments. During pre-visits or on arrival at the school for the actual interviews, the supervisors delivered the Head of Educational Institution questionnaires

and the information materials (comic book informing educators about the study, poster and flyer) on the purpose of the study.

Schools were visited between 9 am and 12.30 midday. However, on request of the principal, visiting times were adapted to suit the activities of the educators.

The school principals were given a Head of Educational Institution questionnaire on which they recorded the number of educators and learners, and absenteeism on the day of the visit by the fieldworkers. Follow-up was done to increase completion rates.

On the day of the visit, nurses wore epaulettes and nametags. Supervisors and the school principals identified appropriate areas to conduct the interview and draw biological samples. The supervisor introduced the study to all the educators before the interviews began.

Only those educators who consented to participate in the study were interviewed. VCT clinic cards were handed out to interviewees who wanted to know their HIV status; the cards referred them to a clinic to undergo a full VCT procedure.

After the interview, the interviewees were asked to give a blood specimen. If the respondents declined, they were requested to provide an oral fluid specimen instead. Barcodes on the questionnaire, specimen and tracking forms linked the questionnaire and specimen. The supervisor was responsible for ensuring that the appropriate medical waste disposal procedures were followed. In addition, the supervisors kept a fieldwork record sheet to record the number of interviews and samples as well as absenteeism. After the school visit, the specimen was immediately couriered to the laboratory for testing. A waste management company collected biological waste from the supervisors.

2.7.3 Quality control

Quality control in a study of this nature cannot be reduced to a single action. In effect, the quest to reduce errors in the study started with the meticulous process of questionnaire design, in order to ensure that the instrument assisted in the collection of quality data. The same applies to the extensive training provided to the fieldworkers.

Direct quality control was the responsibility of the provincial co-ordinators and field team supervisors. The supervisors ensured that the interviewer team visited the correct school, assisted in setting up the interviewing process, for example, by negotiating a time and place for the interviews and checking the completed questionnaires for obvious errors.

A next line of quality control occurred at the office where a team of editors went through the returned questionnaires to code open-ended questions and to ensure that the geographic and other details were correctly entered on the questionnaires. Quality control was also done during the testing of the blood or oral specimens (see 2.2.4 and Appendix 2), while double data entry procedures were used during data capturing to reduce errors to a minimum.

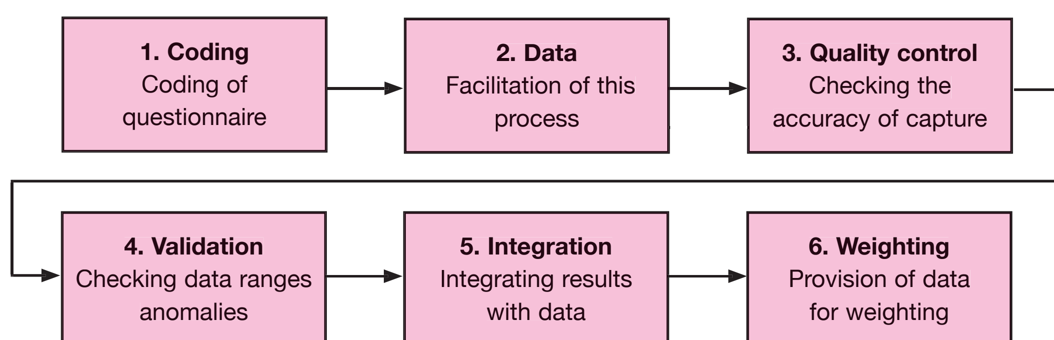
2.7.4 Needle stick injuries to nurses during fieldwork

A few (n=7) needle stick injuries occurred during fieldwork. Fortunately, none of the nurses became infected with HIV because the Standard Operating Procedures For Occupational Exposure were followed.

2.7.5 Data management

Data processing was managed for two questionnaires, namely, the educator and educational institutional sectors (see Figure 2.4).

Figure 2.4: Steps in data processing



The process involved coding the questionnaires and the electronic capture of the data. Biographical information on the questionnaires was also checked during the coding process. Coding involved the allocation of numeric values to text information on subjects trained for and being taught at different education levels (for example, primary, secondary and tertiary). Data was captured electronically and verified through the use of double capturing.

Quality control was done on data captured from the questionnaires. The validation process involved the writing of programmes to detect problems in the captured data and identifying potential outliers. This included checking whether the data was within the ranges stipulated by the project team. Duplicate records were checked as well as wrong codes captured for schools, districts, and provinces before correcting these errors. Checks were made for skip patterns and any errors detected were corrected.

Data management included liaison on a regular basis with the laboratory with regard to specimen results received and any problems that were encountered. As the specimen results were received they were integrated into the existing database of questionnaires. Further quality checks were done to ensure that the type of specimen indicated on the questionnaire was the same as that received from the laboratory. Two problems that had to be addressed were duplicated barcodes for specimens received from the laboratory and mismatches between the specimen's barcodes and data records in the database. Throughout this process new batches of questionnaires and specimen results were received as revisits to the different education institutions were completed, which necessitated processing their data, as has been described above.

To enable the weighting of the data, further reconciliation between data captured from the questionnaires and the actual situation at schools was performed. In other words, the number of educators at each of the schools had to be checked and, if necessary, updated to allow the weighting of the survey results. Tables were also generated to enable the realisation, non-contact/absenteeism and refusal rates to be calculated. The data was then transferred to the statistician for final weighting. Finally, the educational institution and educator data had to be merged.

2.7.6 Weighting of samples

Prior to analysis of data, the following weighting procedures were done to take into account the realised samples and non-responses. The steps required to weight the data are presented below in Figure 2.5.

Step 1 – Calculating the sampling weights: The Statistical Analysis System (SAS) procedure Surveyselect was used to draw the sample of schools in each district. The schools were drawn using probability proportional to size (PPS) sampling and the estimated number of teachers used as the measure of size (MOS). Therefore, the data file of drawn schools contained the selection probabilities as well as the sampling weights of these schools. The first step was to calculate the sampling weight of the schools in each district.

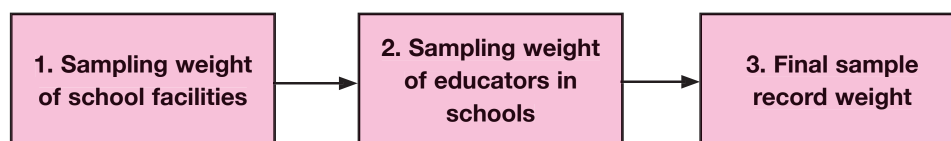
Step 2 – Provision of information on realised sample: With respect of all realised schools or the final samples used in this study, the following information was used:

- The total number of educators in the school;
- The actual number of educators in the school that participated in the study.

Step 3 – Calculating the educator sampling weight: The educator sampling weight was based on the counted number of educators in each school divided by the number of educators participating in the survey.

Step 4 – Calculation of final record weight: In this step, the above information was integrated and the final sampling weight for each data record was calculated. This weight was equal to the final school sampling weights (as given in Step 2) multiplied by the educators' sampling weight.

Figure 2.5: Steps used for weighting of the sample



2.7.7 Data analysis

The collected data in the form of questionnaires was captured on a database by a professional data-capturing company. The database was designed with validation and consistency checks.

After the data was received from the capturing company, programmes were run to validate the reliability of the data. The validation included consistency checks on the extent to which the skip patterns were followed in the data collection. The validation ensured that, for example, males did not answer questions designed for females and vice versa. The quality control of the data was carried out to ensure that there were no data-capturing mistakes. Suspect values were confirmed by drawing the archived physical questionnaire.

The data were converted to statistical packages such as Statistical Package for Social Scientists (SPSS), SAS and STATA. The exploratory analysis of the data was carried out in SPSS and SAS. The exercise included exploring the frequency distributions of all the variables to check the responses to each variable and check that the variables contained values in the accepted range of values. After the data had been edited, weights were calculated and attached to the complete dataset.

The weighted data were analysed using both SAS and STATA. However, an extensive analysis was carried out mainly in STATA, whilst most data manipulation was carried out in SAS. The analysis in STATA took into account the multilevel stratified cluster sample design of the study.

In the analysis, weighted percentages are reported. The reported sample size refers to the sample that was asked the target question. The two-sided 95% confidence intervals are reported. The p-value less or equal to 5% is used to indicate statistical significance. Both the reported 95% confidence intervals and the p-value are adjusted for the multi-stage stratified cluster sample design of the study. Multivariate logistic regression analysis was carried out on the selected determinants of HIV.

In the analysis annual income is reported from low, medium to high; low stands for from less than R40 000 to R60 000, medium = R60 000 to R132 000, and high = more than R132 000. Position in school is sometimes reported as junior or senior: junior = educator and senior = senior educator/education specialist/principal. Educational qualification is sometimes reported from low, medium to high: low stands for Grade 12 and under, medium = diplomas, high = first degree and above. Finally, if marital status is reported as single versus married, single stands for single/separated/divorced/widowed and married = married/co-habitating.