

CHAPTER 2

GENERAL METHODOLOGY

STUDY OBJECTIVES

2.1 Main Objectives:

To establish, in children aged 6-71 months in South Africa, by socioeconomic status, by geographic and age distribution as well as by degree of urbanisation:

- 2.1.1 The vitamin A status;
- 2.1.2 The iron status;
- 2.1.3 The anthropometric status; and
- 2.1.4 The immunisation coverage status.

2.2 Subsidiary Objectives:

To establish in the same population:

- 2.2.1 The prevalence of visible goitre; and
- 2.2.2 The prevalence of breastfeeding.

The study was approved by the ethics committees of all institutions involved in the study, to which the protocol was submitted, in accordance with their existing policy for research. Written parental or guardian consent for drawing blood was obtained prior to the household/child being included in the study. Consent forms explaining the purpose of the study were drawn up and translated into the various languages, as appropriate for the different provinces (Appendix 2.1).

SAMPLING STRATEGY

Study Population

The survey population consisted of all children aged 6-71 months in South Africa. SAVACG decided to obtain a national probability sample of these children, which implies that each child in the population would have a chance of being selected and that the probability of being selected would be known. A multistage probability sample of households was drawn in such a way that reasonably precise estimates could be obtained for each province. Since the population size per province varies considerably, the sample was disproportionately stratified by province, namely a similar number of households was selected in each province; this implies that children had a greater or lesser probability of being selected depending on the population size of a given province. Provincial weighting factors were, therefore, used in the analysis of the data.

Sample Size

Separate calculations were made to obtain the required sample size of children and the required number of blood samples to be drawn.

Sample size of blood specimens

Assuming a prevalence of low (<20 µg/dL) serum vitamin A concentration of 25% and requiring the estimate to be within 8,5% of the true value with 95% confidence, the sample size per stratum had to be 100. Estimates were required for 48 strata, based on urban/rural, province, socioeconomic status and age, leading to 4 800 blood samples. Allowing for a 15% non-response rate, the required sample size amounted to 5 520 blood samples.

Sample size of children

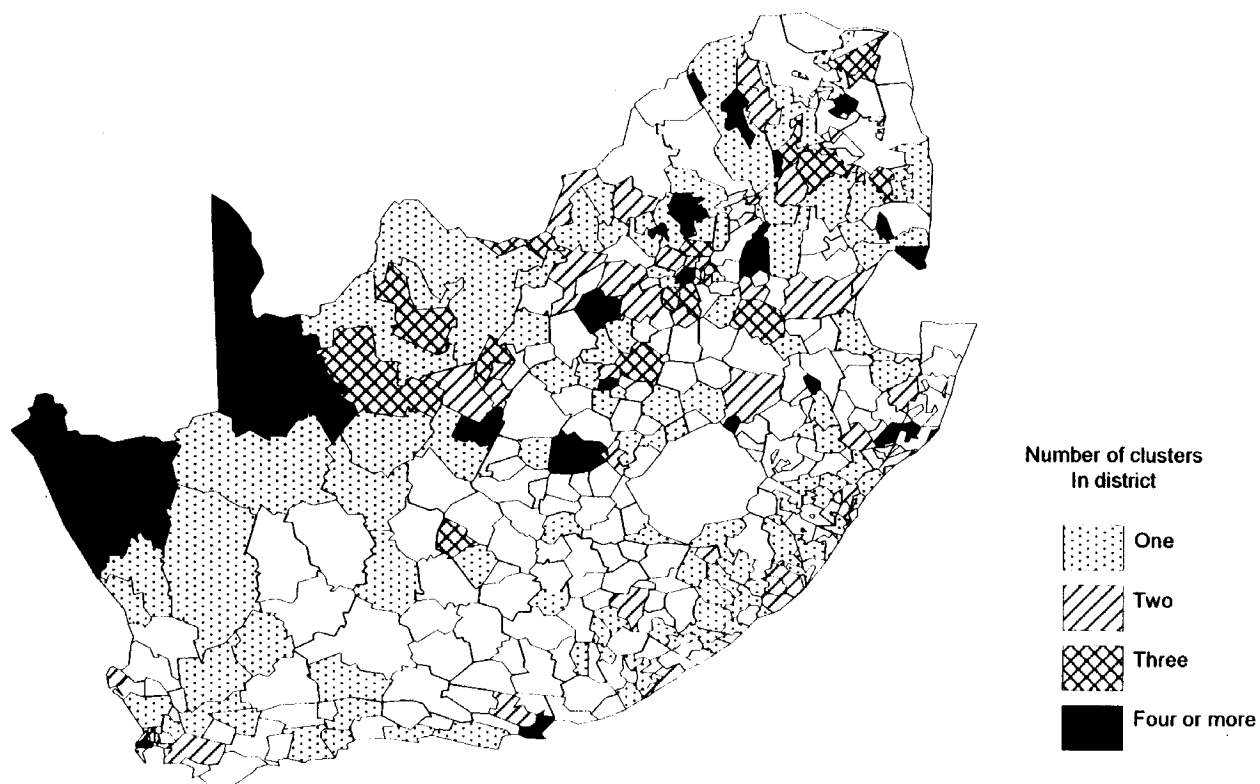
Requirements concerning estimates of immunisation coverage amongst one year old children were used to determine the total number of children in the sample. Assuming a coverage rate of 50 percent and requiring the estimate to be within 10 percent of the true value with 95% confidence, the sample size per province had to be $(1,962)(0,52)/(0,12) = 96$, if a simple random sample of children were to be drawn. Since it was obvious that some form of cluster sampling would be more appropriate, the required sample size was doubled to 200 one year old children per province (assuming a design effect of 2,0). It was decided that, within each of the nine provinces, forty clusters should be selected and that the sampling strategy should aim to include six one year old children per cluster. This led to 240 one year old children per province. Assuming that the number of children aged 6-71 months is 5,5 times the number of one year old children, a total of 1 320 children 6-71 months would be selected per province. For the whole country, the sample size would

amount to 11 880 children. The sampling unit, however, would be households and not children. It was therefore necessary to estimate how many households had to be visited in order to find six one year old children. According to the 1991 census¹, an average of 5,83 persons reside in a household and 2,45 percent of the population is one year of age. These figures translate into finding six one year old children in 42 households. Assuming an arbitrary non-response rate of 20%, 52 households had to be selected in each cluster in order to obtain the required six one year old children.

Sampling Procedure

First stage: selection of clusters

Figure 2.1. Geographic distribution of clusters



A cluster was defined as an enumerator sub-district (ESD) as drawn up for the 1991 census. Forty clusters were selected per province, giving a total of 360 clusters (Fig. 2.1; Appendix 2.2). Within each of the 360 selected ESDs, fieldworkers had to list all households. From this list, the required number of households was selected. The listing of households was anticipated to be fairly laborious, and an alternative was considered, namely making use of the listing compiled by SALDRU², in 1993. The SALDRU sample also consisted of 360 ESDs, but these were distributed according to the population density rather than evenly across the nine provinces. SAVACG decided nevertheless to include as many SALDRU ESDs as possible, and to augment these with additional selections, where necessary.

The SALDRU sample consisted of more than forty ESDs for each of four provinces, namely Eastern Cape, KwaZulu/Natal, Northern Province and Gauteng. Per province, a systematic sample of forty SALDRU clusters was selected, with each ESD having the same probability of selection. Each cluster had to consist of at least 52 households (requirement of the present survey) over and above the 25 included in the SALDRU survey (in order to honour the agreement with SALDRU to avoid, as far as possible, restudying the same households). ESDs with an estimated population size of less than 450 inhabitants were combined with the next ESD on the list which was in the same category of urban or rural. This was done prior to the actual selection.

For the remaining five provinces, all available SALDRU ESDs were included except for two which comprised mining hostels and were, therefore, unsuitable for the present survey. After combining ESDs (as described above) it was found that an additional 92 clusters had to be selected, namely, Eastern Transvaal (17), Northern Cape (35), North West (12), Free State (17) and Western Cape (11).

SALDRU kindly made available a list of all ESDs in these five provinces, which they had entered on computer from paper versions provided by Central Statistical Services and the Department of Economic Affairs in ex-Bophuthatswana. The selection of additional clusters from these lists was done using a similar methodology as that employed by SALDRU³. Within each province, the ESDs were sorted by district and by the proportion of the population that was Black. A systematic sample of the required number of ESDs was selected with probability proportional to the estimated population size of each ESD. If the population of a selected ESD was estimated to be less than 300, then it was combined with the next ESD on the list which was in the same urban/rural category. To balance the probabilities, the reverse was also done, namely, if the ESD on the list immediately before a selected ESD had a population of less than 300, then it was combined with the selected cluster.

Eleven selected ESDs had to be replaced; of these, four could not be entered because of violence and seven consisted of mining hostels. For those provinces in which only SALDRU clusters were included, a random selection of the remaining SALDRU clusters determined the replacement ESD. In the other provinces, the next ESD on the list (in the same district and the same urban/rural category) was chosen.

Approximately fifty clusters had large populations. These were areas of which aerial photographs had been taken for the 1991 census. For each such cluster, one photograph was randomly selected from all those representing the area. If this was found to be unsuitable, for example, containing many trees and bushes but no structures indicating the presence of habitation, then another photograph of the same ESD was randomly selected. On the photograph, an area comprising at least 52 structures was selected.

Second stage: selection of households

Fieldworkers were given descriptions of selected ESDs, in many instances merely the boundaries, as they existed during the 1991 census. A list of all households or stands in each ESD was compiled. For selected SALDRU clusters, fieldworkers had to ensure that the household listing was still valid and amend it, if necessary.

The sampling strategy had originally been designed in such a way that all households (and therefore all children) within the province would have the same probability of being selected. A communication problem at central level led to a deviation from the original design in the field. Fieldworkers were instructed to calculate a selection interval by dividing the number of households on the list by 52. Using a random start, households were systematically selected using this interval. Only households indicated in the listing were visited and no substitutions were made. If the selected household had been included in the SALDRU survey, the next household was visited; this was done because of the agreement made with SALDRU that we would avoid, as far as possible, including the households SALDRU used, in case SALDRU had to restudy any of their chosen households. If more than one family lived in the household or on the property, they were all included in the study. If there was no response at a selected household, fieldworkers were requested to make one further attempt, after which the household would be considered as a non-response.

Third stage: selection of children

All children aged 6 to 71 months in selected households were included in the sample.

Fourth stage: selection of children from whom blood samples were drawn.

Children were eligible to have their blood drawn if the following conditions were met: the parent or legal guardian gave consent, the child did not have a high temperature ($>38^{\circ}\text{C}$), and the child had not received high dose vitamin A supplementation in the preceding 6 months. Within each household, all children within the age group 6-71 months were included in the study. From the first 16 children for whom permission was obtained, blood was drawn for biochemical analysis. If the sixteenth child had siblings in the target age group, then blood was drawn from them as well. The vitamin A₂ analogue was administered to the first child in each cluster for whom permission to draw blood was given.

METHODOLOGY

An instruction manual was compiled and used by all fieldworkers throughout the study (Appendix 2.3). The purpose of the manual was to ensure standardisation and uniform procedures.

Questionnaires

A household and individual child questionnaire was completed for each household and each child within the household (Appendix 2.4).

The household questionnaire was completed for each family per household visited and contained mostly socioeconomic information i.e living conditions, water and electricity supply, highest level of education and present occupation of the mother. The number of children between the age of 6-71 months was also indicated and numbered (where the youngest was allocated number 1 and the oldest child the highest number).

The individual child questionnaire was completed for each child aged 6-71 months. Information obtained on this questionnaire included the date of birth and sex of the child, immunisation status, duration of breastfeeding, clinical eye and goitre examination results, anthropometric data and, if blood was taken from the child, the section on the medical and hospitalisation history was also completed.

After completion of the questionnaires by the fieldworkers, the data were checked by the team/provincial coordinator before the questionnaires were sent for data entry to the Directorate of Epidemiology, Department of Health.

Clinical Examination

Eye examination

A history of night blindness was taken by asking the parent/guardian whether the child has any difficulties in seeing or finding things in the dark. Each child's eyes were checked for signs of vitamin A deficiency. Any abnormalities found were compared with the WHO's vitamin A deficiency eye charts provided to the fieldworkers and were recorded on the questionnaires. The various categories of vitamin A deficiency eye signs available were: normal, corneal xerosis, Bitot's spots, corneal scar, keratomalacia and blindness (Appendix 2.5).

Goitre examination

Each child's neck was checked for visible goitre with the neck in the neutral and extended position and the results indicated on the questionnaire. Any abnormality found was brought to the attention of the group/provincial coordinator for a second opinion.

Interventions

If abnormalities were found on examination or on inspection of the Road to Health card, the child was referred to the nearest health facility. Children found to have a serum vitamin A concentration $<10 \text{ } \mu\text{g/dl}$, were followed up through the local clinic/hospital and were given vitamin A supplements as follows: first dose (200 000 IU) on day 1, a second dose (200 000 IU) 24 hours later, and a third dose (200 000 IU) 4-6 weeks later. Children younger than 12 months of age received vitamin A supplements using the same administration schedule but with half the dose (100 000 IU). Children with a haemoglobin concentration $<8 \text{ mg/dl}$ were referred to their local doctor for evaluation and appropriate treatment.

Human Resources

A national and nine provincial coordinators were enlisted to organise and manage the fieldwork. Their main task was to provide on-going support for the survey in their province and to be available to solve problems as they occurred. The coordinators recruited and trained fieldworkers. The coordinators and fieldworkers were recruited from the various Departments of Health, universities, local authorities and communities. The fieldteams were mostly multi-disciplinary incorporating nurses, dietitians, optometrists, doctors, medical specialists, health inspectors, health educators, community health workers and members of the public. They undertook this survey alongside their normal duties and did an excellent job under sometimes difficult conditions.

Pilot Study

Pilot studies were done in each province during May and June 1994. The purpose of these studies was to familiarise the fieldworkers with the methodology and to solve any problems that might occur. Aspects that were specifically addressed included the suitability of the questionnaires, the use and standardisation of all instruments, the willingness of parents to allow blood to be drawn from their children, the logistics of blood transportation to the analytical centre in Tygerberg Hospital, University of Stellenbosch, and the time needed to complete the survey per cluster. Discussion and solution of the problems encountered during these studies equipped coordinators to anticipate and deal with problems which they might experience during the survey.

Transport

Vehicles were obtained from the Department of Health or local authorities. In most provinces the number of cars available was insufficient and had to be supplemented with hired cars. On some occasions, private cars were used.

Fieldwork

Fieldwork was done from June 1994 to January 1995, with the major part of the work completed during July-October 1994. Some areas had to be substituted according to a set protocol if the area was a single sex hostel, hospital or retirement home or if the level of violence caused concern for the safety of the fieldworkers. Go-slow industrial action delayed the completion of the fieldwork in certain provinces. Some obstacles, mainly transport and the identification of the exact household from the supplied maps, were encountered in varying degrees in all the provinces, whereas other obstacles were unique to specific areas; in the provinces with more than two former health administrations entering the process of amalgamation, the survey was complicated by first having to establish contact between these departments.

Standardisation

Fieldworkers

Various measures were taken to ensure that the information obtained by all the fieldworkers was reliable:

- The 9 provincial coordinators were uniformly trained with reference to all the measurements that had to be done and procedures to be followed.
- All the fieldworkers were uniformly trained by their respective provincial coordinators.
- The training sessions were followed by pilot studies in all the areas, where problems that were encountered could be addressed and solutions found.
- During the execution of the study, all the fieldworkers were working under the supervision of a group/team coordinator. The function of this person was not only to prevent protocol violations, but also to check all the questionnaires for completeness and correctness as well as to repeat all measurements in five randomly selected households per cluster in order to ensure their correctness. This information served as a double-check system to ensure accuracy.

Laboratories

All the laboratories involved in the analysis of the full blood count used standard techniques and were standardised according to existing routine procedures. The one laboratory involved in the analysis of serum vitamin A and 3,4-didehydroretinol (vitamin A₂) was standardised respectively against the National Institute of Standards and Technology (Standard Reference Materials Program, Gaithersburg, USA) and the Department of Biochemistry, Iowa State University, Iowa, USA. The importation of vitamin A₂ and its administration to children was approved by the Medicines Control Council (reference 16/7/3/2A and 26/8/1/2/1, respectively).

DATA ENTRY AND EDITING

Household and Individual Child Questionnaires

The provincial coordinators checked all completed questionnaires before sending them to the Department of Health, Pretoria, for data entry. Because the correctness of the study numbers was vital for linking the information of a child with the corresponding household data, these were all checked again at central level.

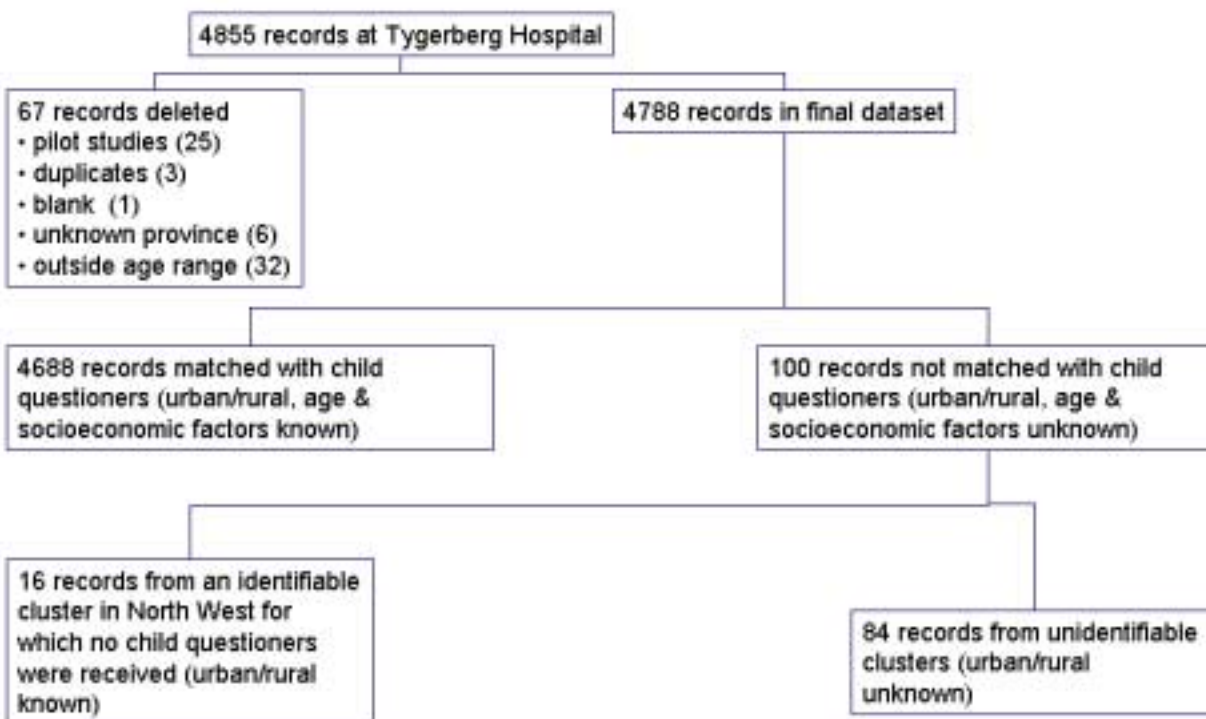
The data were entered on computer using Epi Info version 5.01b⁴. All data were keyed in twice, usually by different people, and then validated by comparing the two versions. Any differences were then checked against the original questionnaires and corrected.

In total, 11 579 child questionnaires were entered on computer. Forty-five of these indicated that the child's age was older than 71 months and 104 children were less than six months of age. These were excluded from further analysis leaving 11 430 questionnaires in the study sample.

A number of inconsistencies were found in the immunisation data. Approximately 850 questionnaires had at least one immunisation date in the wrong order. Where the error was fairly obvious, this was corrected. In those instances where no amount of logic could reconstruct the correct sequence of events, the offending date was replaced by an indication that the immunisation had been given.

Blood Results

Figure 2.2. Details of bloods drawn



The results of blood investigations were entered on computer at Tygerberg Hospital, University of Stellenbosch, also using Epi Info version 5.01b⁴. The file with 4 855 records was sent to the Department of Health, Pretoria, for analysis. A check was first done to see whether all these records matched up with the child questionnaire, using the study number as a link. Where this link could not be established, surnames and initials that had been keyed in on the laboratory dataset were used for matching with the original questionnaires in Pretoria. Sixty-seven records in the laboratory dataset had to be deleted. A further 100 records could not be matched with child questionnaires, but were left in the dataset since the province of origin was identifiable (Fig. 2.2).

DATA ANALYSIS

Weighting Factors

Provincial weighting factor

Table 2.1 Provincial Weighting Factors

	Northen Cape	Western Cape	Eastern Cape	KwaZulu Natal	Eastern Transvaal	Northern Province	Gauteng	North West	Free State	South Africa
No. of Children aged 1-4 years ¹	69 207	305 636	686 597	884 634	303829	554 170	629 506	354 583	287 148	4 075 265
No. of Children aged 6-71 months ²	95 160	420 250	944 071	1 216 372	417 765	761 984	865 571	487 490	394 829	5 603 492
Weighting factor	0,0170	0,0750	0,1684	0,2170	0,0746	0,1360	0,1545	0,0870	0,0705	1,0000

¹Source of mid-1994 population estimates: Directorate Epidemiology, Dep of Health (based on 1991 census)

²Proportional calculation based on 1-4 year old population

The sample was disproportionately stratified by province. This implies that children had a greater or lesser probability of being selected depending on the population size of a given province. Therefore, in order to obtain national estimates of the various parameters being investigated, the province-specific results were weighted by the corresponding population aged 6 to 71 months (Table 2.1).

Cluster weighting factor

The sample was designed to be self-weighting within provinces and the appropriate selection intervals to achieve this were calculated. Fieldworkers, however, calculated their own intervals in order to select exactly 52 households (this was in accordance with the instructions they received). Comparison of the two sets of selection intervals for those areas where they were available does not indicate a great discrepancy, and thus the sample could still be regarded as self-weighting. There was, however, one significant deviation from this design. Several clusters consisted of two ESDs. Instead of selecting a total of 52 households from the two ESDs, a number of fieldworkers selected up to 104 households. This would obviously have affected the probabilities of selection, and in order to accommodate it in the analysis, a cluster-specific weight was utilised. If the number of households in the cluster was 52 or fewer, the weight was 1. If the number of households was more than 52, the weight was 52 divided by the number of households.

Weighting factor for blood samples

The sampling strategy indicated that 16 blood samples should be obtained per cluster, irrespective of the number of households visited or the number of children in that cluster.

It follows that in the case of a smaller cluster the probability of a child being selected to have a blood sample drawn was greater than in the case of a larger cluster. To correct for this in the analysis, a weight equal to the number of children in the cluster divided by the number of blood samples drawn in the same cluster was introduced. For the 100 blood samples (Fig. 2.2) which could not be matched with child questionnaires, the average cluster-specific weight for the corresponding province was used.

Statistical Programs

All the analyses were done using Epi Info version 6.02; more specifically, the CSAMPLE module of Epi Info was used to obtain weighted estimates and confidence intervals.

REFERENCES

1. Population Census 1991. Households. Central Statistical Services. Report No: 03/01/25. Pretoria. 1992.
2. South African Labour and Development Research Unit (SALDRU). National Living Standards Survey. University of Cape Town. 1993.
3. South African Labour and Development Research Unit (SALDRU). National Living Standards Survey. Revised sample design: Technical description. University of Cape Town. 1993.
4. Dean AG, Dean JA, Burton AH, Dicker RC. Epi Info, Version 5: a word processing database and statistics programme for epidemiology on microcomputers. Centre for Disease Control. Atlanta, Georgia. USA. 1990.
5. Dean AG, Dean JA, Coulombier D, et al. Epi Info, Version 6: a word processing, database, and statistics program for epidemiology on microcomputers. Centre for Disease Control and Prevention. Atlanta, Georgia. USA. 1994.