

Indicators for Assessing Iodine Deficiency Disorders and their Control Programmes

Report of a Joint WHO/UNICEF/ICCIDD
Consultation, 3-5 November 1992

Review version, September 1993

This report deals in a detailed way with the purposes, principles, procedures and indicators for surveillance in relation to both the epidemiology of iodine deficiency disorders, and programme implementation indicators as regards the main intervention (salt iodization) for preventing IDD. It demonstrates and describes how the indicators and procedures will need to vary according to the purposes of a survey or surveillance activity.

Precise guidelines are given on such important issues as the characteristics and criteria for choice of various indicators (clinical and biochemical), selection of age/physiological groups for surveys, and sample size in IDD surveys. It presents a new simplified (3-stage) version of the classical 5-stage goitre grading system, and an outline of the principles for use and interpretation of ultrasound measurements of thyroid size.

It summarizes recommended procedures as regards urinary iodine and thyroid-related hormone estimations. Criteria for determining the severity of an IDD endemic are defined in a simple table.

For salt iodization, a tabular summary is given of the recommended levels of salt fortification according to conditions of climate, salt packaging and daily consumption. Procedures for monitoring salt iodine content at the factory or importation level, and at the district and household levels, are outlined.

Finally a succinct table summarizes the main indicators recommended for use in tracking progress towards the World Health Assembly/World Summit for Children goal of eliminating IDD by the end of this decade.

This condensed report therefore contains much valuable information in the form of guidelines and didactic material for national programme managers, and for various organizations which are supporting national efforts for eliminating IDD.



World Health
Organization



United Nations
Children's Fund



International Council
for the Control of
Iodine Deficiency Disorders

Table of Contents

Preface	vi
1. Introduction	1
2. Purposes of Surveillance for IDD	2
3. Selecting Target Groups for Surveillance (Whom to Measure)	3
3.1 Criteria for Target Group selection	3
3.2 Target Groups for IDD Surveillance	4
4. Surveillance Methods for IDD (How to Measure)	6
4.1 Assessing the prevalence of IDD	6
4.2 Identifying high risk areas	7
4.3 Monitoring and evaluating IDD control programmes	8
4.4 Measuring progress towards long-term micronutrient goals	9
5. Interpretation and Presentation of Results	10
6. The Selection of Appropriate Indicators (What to Measure)	11
6.1 Criteria for Indicator Selection	11
6.2 Outcome Indicators of IDD	12
Clinical Indicators	12
Thyroid Size	12
Palpation	12
Ultrasonography	14
Cretinism	16
Biochemical Indicators	16
Urinary Iodine	16
Blood constituents	19
Thyroid Stimulating Hormone (TSH)	19
Thyroglobulin (Tg)	22
Summary of Outcome Indicators	23
6.3 Process Indicators of IDD Control Programmes	23
Salt Iodization Programmes	23

Techniques for measuring salt iodine levels	24
Monitoring salt iodine levels	25
Procedures for monitoring	26
Other IDD Control Programmes	27
7. Criteria for Tracking Progress Towards World Health Assembly/World Summit Goals	28
Annex 1 List of Participants	29
Annex 2 Sample Sizes for IDD Prevalence Surveys	30

Abbreviations

ELISA	Enzyme-linked Immunosorbant Assay
EPI	Expanded Programme of Immunization
ICCIDD	International Council for the Control of Iodine Deficiency Disorders
IDD	Iodine Deficiency Disorders
IQ	Intelligence Quotient
LQAS	Lot Quality Assurance Sampling
MCH	Maternal and Child Health
PAMM	Programme Against Micronutrient Malnutrition
ppm	parts per million
PPS	Probability Proportionate to Size
Tg	Thyroglobulin
TSH	Thyroid Stimulating Hormone
UNICEF	United Nations Children's Fund
WHO	World Health Organization

List of Tables

	page
1. Framework for considering target groups for IDD surveillance	5
2. Proposed classification of goitre	14
3. Proposed epidemiological criteria for assessing the severity of IDD, based on prevalence of goiter in school-aged children	14
4. Comparison between palpation and ultrasound in grading small thyroids	15
5. Normative thyroid volume size	15
6. Proposed epidemiological criteria for assessing severity of IDD based on median urinary iodine levels	19
7. Summary of IDD prevalence indicators and criteria for a public health problem	23
8. Recommended levels of iodine in salt	25
9. Proposed criteria for assessing adequacy of salt iodization programmes	26
10. Criteria for tracking progress towards the World Summit goal of virtual elimination of IDD	28

Preface

The consultation of which this is the Report, was convened by WHO in November 1992 on behalf of the three organizations (the International Council for the Control of Iodine Deficiency Disorders, ICCIDD; the United Nations Children's Fund, UNICEF; and the World Health Organization, WHO, all of which have been working most closely together over the last 8 years or so in an effort to combat one of the most ancient and insidious scourges of the human race – iodine deficiency disorders. Iodine deficiency has serious effects on the physical development of children, on young-child mortality and on the reproductive performance of women with increased rates of abortion, stillbirth and congenital abnormalities. But it is primarily the impairment of intellectual development and in severe cases mental defect ranging up to cretinism, which have mobilized the world – scientists, public health administrators and political leaders alike – to call and work vigorously for the elimination of IDD.

"Indicators" looks like a dry topic, but both the subject and participation in the consultation-meeting were characterized by intense enthusiasm. Considerable momentum has already been generated world-wide towards the elimination of IDD within this decade, and it is imperative to measure progress and assess whether and when the goal of elimination has really been reached. Only by clear definition of appropriate indicators and criteria, can the adequacy of progress towards this goal be assessed.

This consultation was marked by a refreshing new look at the whole subject of IDD indicators. Even the hallowed WHO classification of goitre came under critical review! Participants were in favour of facilitating an increasing role of anyone strategically situated in the field, to identify IDD problems, assess in a simple way their severity and the adequacy of measures to combat iodine deficiency through the main intervention – salt iodisation. Thus it was agreed that while the classical 5-stage grading of goitres (0 Ia Ib II and III) remains valid, a simpler grading into three grades (0, 1, and 2) would be much easier and more pragmatic for field use, for instance for screening purposes by school teachers trained in the method.

This text provides also the best outline internationally agreed, to date, of practical procedures for verifying the adequacy of salt iodisation on a national and local scale.

The vexing issue of numbers of subjects and biological samples required for IDD surveys is creatively tackled, taking into consideration the various purposes for which surveys are conducted – sampling methods and sizes vary according to the aims of the survey. Relatively clear and epidemiologically adequate guidance is given on this subject for the first time, and reference to appropriate sources for further details.

The assessment of IDD is constantly evolving and so also is the science of indicators of IDD. This report will certainly not be the last word on the subject. But it is a milestone along the way and which may well serve as an

excellent yardstick for measuring progress throughout this decade. Further experience with the use of the methods and procedures outlined here will enable the indicators and criteria to be further elaborated and refined over the next few years.

So, to the perceptive reader we offer in this little monograph a delicious guided tour through the fascinating microcosms of the epidemiology of IDD, and the march towards their elimination.

To the IDD-experts of the ICCIDD, UNICEF, WHO and associated institutions who contributed so richly to this subject, to the meeting and to the development of this report, we offer our heartfelt gratitude.

Special thanks are due to those who laboured longest over this Report, including Peter Greaves, Jonathan Gorstein, Suzine Pak, Kevin Sullivan, Rick Trowbridge and to each of the group leaders who worked on various sections in this Report.

1. Introduction

1. An indicator is something that draws attention to a condition. A discussion on indicators therefore must begin with an understanding of the conditions that are of concern, and whose attention is sought to be aroused, by whom, and for what purpose, before the indicators themselves can be discussed with much meaning. These issues are the subject of this report, with respect to IDD.
2. Three reasons have combined to make it timely to reconsider the matter of indicators. First, increasing scientific knowledge about IDD and accumulating experience of IDD control programmes make it appropriate to review and, if necessary, revise earlier judgments. Second, experience from other areas – control procedures in manufacturing industry, monitoring progress of the Expanded Programme on Immunization (EPI) – has provided new tools that can be applied with advantage to the area of micronutrients. And third, the adoption by governments of ambitious goals for the elimination of IDD by, if not before, the end of the century, has provided an impetus and an urgency for action.
3. This report recognises that IDD is not simply equivalent to goitre. IDD covers a broad range of effects, notably on the brain, and on intelligence. Iodine deficiency causes a marked reduction in IQ. Therefore indicators are recommended that are sensitive to these broader effects. Many countries will need training and technical support in their application, and this can be provided: information may be obtained from ICCIDD, WHO, UNICEF and PAMM.
4. The consultation of which this is the report was one of three, planned for each of the micronutrients for which WHO and UNICEF have specific goals. Much of the discussion reported in the first five sections is valid in general terms – also for iron and vitamin A. Most of the last two sections are specific for iodine. In view of the interest that attaches to micronutrients as such, and the potential gains to be had from a combined approach to assessment and monitoring, it may become feasible to develop integrated guidelines for country action that enable all three micronutrients – in countries where deficiency of each presents a public health problem – to be tackled together. For example, some lack of precision may be accepted if a common sampling frame reduces costs and improves the logistical capacity to measure progress towards micronutrient goals. There is no point in continuing to amass data when a carefully planned survey covering a minimum number of people produces information that is sufficiently reliable. As the chairman reminded the meeting, it is more important to be roughly right, than precisely wrong.

The list of participants is given in Annex 1.

2. Purposes of Surveillance for IDD

5. The surveillance of iodine deficiency disorders (IDD) may serve a variety of distinct purposes, including: assessment of the magnitude and distribution of IDD prevalence; the identification of high risk populations; evaluation of control programme activities; and the tracking of progress towards long range goals. The surveillance design employed, indicators used, and the approach to interpreting the data will vary according to the specific purpose at hand.
6. **Assess the prevalence of IDD:** One of the fundamental purposes of IDD surveillance is to determine the magnitude and distribution of IDD within a population. This assessment can provide a baseline for long term monitoring as well as a basis for advocacy to highlight the extent of IDD problems and to stimulate action, including the allocation of resources required for IDD elimination.
7. **Identify high risk areas for intervention:** IDD assessment is also critical to programme development by identifying high risk communities. This type of surveillance activity is primarily concerned with the identification of priority areas for intervention, thereby resulting in more efficient use of resources.
8. **Monitor and evaluate IDD control programmes:** Another fundamental purpose of surveillance is to evaluate programme implementation and impact. Indicators can be measured that assess the extent of programmatic activity as well as the impact on specific outcomes.
9. **Measure progress towards long term micronutrient goals:** Many countries have joined in working towards the goal of eliminating IDD as part of achieving child health and development goals for the year 2000. Surveillance activities can provide a quantitative basis for assessing progress towards those goals.

3. Selecting Target Groups for Surveillance (*Whom to Measure*)

10. A wide variety of target groups, including newborns, infants, preschool children, school-aged children, and certain groups of adults might serve as the focus for IDD surveillance. The selection of the optimal group or groups depends on a number of considerations, including vulnerability, representativeness, accessibility, and potential usefulness for surveillance of multiple health problems. These criteria are considered below.

3.1 Criteria for Target Group selection

11. **Vulnerability:** In order to serve as a sensitive indicator, the target population must be vulnerable to the deficiency. Three aspects of vulnerability are:

- Extent of exposure to deficiency
- Severity of health consequences due to deficiency
- Degree of clinical or biochemical responsiveness to deficiency and to interventions.

12. **Representativeness:** The issue of representativeness arises in two contexts:

- Is the target group used for surveillance representative of all persons in the same age/sex group in the community? For example, if children examined in school are the target group, are they representative of all children of school-age in the community? It may be that those who are in school come from more advantaged or better educated families, and their risk of IDD may be lower. This aspect of representativeness is generally referred to as internal validity.
- Is the IDD status of the target group representative of the status of the community as a whole? It may be that the apparent prevalence of IDD in the target group tends to overestimate or underestimate the prevalence in the community. This aspect of representativeness is generally referred to as external validity.

13. **Accessibility:** Another criterion for target population selection is the accessibility of the population for measurement. Easily accessible populations, such as children in school, women in MCH clinics, or newborns seen in hospital, may be useful for surveillance. Using these relatively accessible target groups will tend to facilitate surveillance and reduce logistic costs, but the most accessible groups may not always prove to be fully representative or the most vulnerable.

14. **Usefulness for surveillance of multiple micronutrient and other health problems:** It may prove advantageous if the target group selected for IDD surveillance can also serve for the assessment of other nutritional and health problems. For example, an IDD surveillance based in schools might also serve for the surveillance of iron status or helminth infections. However, school-aged children may be less useful for the assessment of vitamin A deficiency and anthropometry.

3.2 Target Groups for IDD Surveillance

15. Applying the above criteria to potential target groups provides a view of some of their advantages and disadvantages for IDD surveillance. A framework for considering these characteristics in relation to various potential target groups is presented in Table 1.

16. **Newborns:** The use of newborn screening for identification of congenital defects is well established in many developed countries and is being instituted in some less developed countries. Where instituted, the regular collection of blood spot samples can serve as a major source of information for IDD surveillance because these samples may be used to assess thyroid stimulating hormone (TSH) status.

17. Surveillance may also be done more practically by collection of blood spot samples from cord blood in a sample of newborns in iodine deficient areas. Blood spots could be collected by trained birth attendants in households, local health posts or hospitals as part of primary health care for newborns. The representativeness of the sample in this type of surveillance would depend on the coverage of newborns by the primary health care system. A surveillance system of this design could be highly effective and would make efficient use of the existing primary health care infrastructure.

18. **Infants and preschool children:** Infants and preschool children are also highly vulnerable to the effects of IDD and may be useful for surveillance of other health problems in addition to IDD. However, they may not be highly accessible except in MCH clinics, where issues of representativeness may arise. Some indicators of IDD, including goitre by palpation, may be relatively difficult to assess in this age group. There may be opportunities to coordinate IDD surveillance and supplementation with the Expanded Programme on Immunization (EPI).

19. **School-aged children:** School-aged children are a useful target group for IDD surveillance because of the combination of high vulnerability, easy access, and usefulness for a variety of surveillance activities. Adolescents develop an enlarged thyroid in response to iodine deficiency and can readily be examined in large numbers in school settings. Other health concerns in

this age group, including helminth infections, anaemia, and behavioural risks can be assessed at the same time as IDD and educational interventions may be implemented. The major concern which arises in the consideration of school-aged children for surveillance is that children who do not attend school are not represented, which may lead to biased prevalence estimates. It may however be possible to use the school premises for the assembly of such children, and also younger preschool children, on special occasions.

20. Pregnant women in MCH clinics: Pregnant women are a vulnerable group whose iodine status is particularly critical because of the susceptibility of the developing foetus to iodine deficiency. When seen in the context of a primary health care setting, pregnant women are quite accessible and a wide variety of other health measures can be assessed simultaneously. Representativeness may be a problem depending upon the level of access to and utilization of health care services by the women who are at highest risk.

21. Adults in households: Assessing adult women and men in the context of household surveys provides an opportunity to gain a sample of a population vulnerable to IDD. Above the age of 30 goitre and TSH levels are not reliable indicators of the current iodine intake. Depending upon the circumstances, accessibility may be limited because of the expense and logistical constraints associated with performing household surveys. Representativeness will depend on the extent to which men and women may work outside the home; in particular it may be difficult to find men at home during the day.

Table 1. Framework for considering target groups for IDD surveillance

	Vulnerability	Representativeness*	Accessibility	Usefulness for other Surveillance**
Newborns	High	Intermediate	Intermediate	Intermediate
Preschool children in MCH clinics	High	Intermediate	High	High
Preschool children in households	High	High	Intermediate	High
Children in schools	High	Intermediate	High	High
Pregnant women in MCH clinics	High	Intermediate	High	High
Adult women in households	Intermediate	Intermediate	Intermediate	Intermediate
Adult men in households	Intermediate	Low	Low	Low

*Level of representativeness depends on access or coverage (see text)

**Usefulness of group for surveillance of other nutrition and health problems

4. Surveillance Methods for IDD

(How to Measure)

22. As mentioned in section B, there are several purposes for the surveillance of IDD. Different data collection methods will be necessary depending upon the purpose of surveillance. Potential data collection schemes for each of the purposes of surveillance are described below. When designing a surveillance system, it is important to determine whether the primary purpose of the survey is to derive a prevalence estimate, or to identify high prevalence areas. Each of these two purposes usually requires a different type of survey method, and studies that attempt to fulfill both purposes tend to be inefficient and fail to answer adequately either question well. In general the first step is to perform a prevalence survey. If the results of the prevalence survey indicate that an IDD problem exists and it is decided that identifying high prevalence areas is important, then a lot quality assurance sampling (LQAS) survey would be useful.

4.1 Assessing the prevalence of IDD

23. There are two fundamental principles in assessing the prevalence of IDD (or of anything else): to collect the minimal amount of data that are representative of the target population, and to provide a stable prevalence estimate within a desired level of precision. The survey method to be used depends upon many factors, including the target group (e.g., newborns or school children), the survey site (e.g., households or schools), and the size of the geographic area of interest. In large countries, it may be desirable to perform prevalence surveys at a provincial level. However, the number of geographic units to be studied should be kept to a minimum.

24. A common method used for immunization and anthropometric household-based surveys is the "probability proportionate to size" (PPS) cluster method. This method is useful when census data are of low quality. In general, all villages and cities are listed and a systematic sampling scheme is used based on the cumulative population. This sampling scheme assures that larger villages and cities are more likely to be selected than smaller ones. Thirty clusters (of households, schools, etc.) are selected at random. The selection of thirty clusters is to assure a valid estimate of the prevalence; collecting substantially fewer can lead to estimates that differ dramatically from the true prevalence. The number of individuals to sample within each cluster depends upon the prevalence of the condition, the level of precision desired (based on the type of confidence interval and width of the confidence interval), the design effect (a measure of the variability of the prevalence between clusters), whether inference at the cluster level is desirable, and the precision of the method of surveillance.

25. The estimated prevalence would be the prevalence of an abnormal biochemical or clinical indicator of IDD. If there is no available information on the likely prevalence, 50% is used in calculating the sample size needed.

(For a given sample size, the maximum variance occurs when the prevalence is 50%.) The sample size can be reduced if the expected prevalence is substantially less than or greater than 50%. The type of confidence interval is traditionally 95% or 90%. The absolute precision on either side of the prevalence estimate is traditionally $\pm 5\%$ or $\pm 10\%$, for prevalence of 50% (please see however the note on this subject in Annex 1). There has been limited experience in estimates of the design effect of indicators for IDD, but values have been in the range of two to four. As an example, commonly used estimates and values would be:

Estimated prevalence	50%
Confidence interval	95%
Absolute precision	$\pm 5\%$
Relative precision	$\pm 10\%$
Design effect	3

The sample size based on this information would be 40 children per cluster for a total of 1200 children (30 clusters x 40 children per cluster). This is sufficient to establish whether IDD is present. But if the intention is to assess the degree of severity, the sample size has to be larger: in most cases 75 children per cluster would be sufficient. For more explanations, details and tables of sample sizes, please see Annex 2.

26. Surveys of schools or health clinics could use either a PPS or a two-stage cluster survey. In the latter, generally the schools or clinics would be selected at random (not by PPS) from a listing, and then individuals would be selected at random from within selected schools or clinics. Either a fixed number of individuals could be selected from each school or clinic or a proportion of the population could be selected. One important distinction between a PPS and two-stage cluster survey is that the latter requires an accurate count of the population under study in each cluster. The population counts are used to weight the results in proportion to cluster size. The sample size for a two-stage cluster survey would be the same as that described for a PPS survey.

27. The sampling of neonates may be more complex, but whatever method is devised a random selection of neonates should be assured.

4.2 Identifying high risk areas

28. In some situations a goal of surveillance may be to identify areas of high prevalence in order to focus intervention activities. A difficulty with identifying high prevalence areas is that IDD tends to occur in geographic foci, and a large number of sites may need to be surveyed in order to find the high prevalence areas. An efficient survey method for screening a large number of sites is "lot quality assurance sampling" (LQAS). For example, if

school children were the target group for surveillance for identifying areas with a high prevalence of goitre, LQAS would be appropriate. In order to find the high prevalence areas, every school within a geographic area would be surveyed, and within each school a sample of school children would be palpated. If a large number of children had goitre, then the area would be identified as "high prevalence". To continue the example, suppose that the estimated prevalence of goitre among school children in a region is estimated to be 10%, and the Ministry of Health is interested in identifying schools with a prevalence greater than 30% in order to focus intervention efforts. How many children would need to be palpated in each school, and at which point would a school be classified as having an IDD problem? Using standard WHO tables (assuming a significance level of 5% and power of 90%), in each school 33 children would need to be palpated, and if five or more children had goitre, the school would be classified as having an IDD problem (i.e., a "rejected" lot); if fewer than five had goitre, the school would be classified as not having a problem (i.e., an "accepted" lot). When schools are either "rejected" or "accepted", there are two possible errors: a school could be accepted as not having an IDD problem although it truly had an IDD problem, or rejected although it did not have an IDD problem. The use of LQAS in this situation is to minimize the first error because this error may lead to not intervening in areas that are in need of intervention. The second type of error is of relatively less importance. In the example above there is only a 5% chance that, if fewer than 5 students out of 33 had goitre, the true prevalence in the school would be greater than 30%.

29. In order to minimize the number of schools to be screened, the study area could be reduced to include only areas at higher risk of having IDD, such as rural mountainous areas. If all schools in an area are included in an LQAS survey, there are ways in which this information can be used to derive a prevalence estimate for the area.

4.3 Monitoring and evaluating IDD control programmes

30. The main control programmes for IDD are fortification of salt with potassium iodate (or iodide) and supplementation with iodized oil capsules. The procedures for monitoring salt iodization are outlined in paragraphs 82 to 86. For iodized oil programmes, in areas where iodized oil is recommended for specific target groups (such as young children and women of child-bearing age), LQAS surveys can be useful in determining the coverage of the population.

4.4 Measuring progress towards long-term micronutrient goals

31. Periodic prevalence surveys as described earlier in this section are necessary to measure the change in prevalence over time. To measure progress towards long-term micronutrient goals requires that the surveys are representative of the population. Because the surveys need to be repeated, they should be as simple to perform and analyze as possible and based on the minimum number of individuals required to provide stable prevalence estimates within the desired level of precision.

5. Interpretation and Presentation of Results

32. Many indicators of IDD measured on a continuous scale (e.g., urinary iodine levels and TSH) are not normally distributed and the use of means and standard deviations are likely to be inappropriate. With some non-normally distributed indicators it may be possible to transform the data (e.g., by logarithmic transformation) before calculating means and standard deviations. For other indicators transformations may not work and the presentation of a median or other percentiles and use of nonparametric statistics would be appropriate.

33. If possible the full distribution of results should be presented, in addition to a measure of the central tendency (mean or median) and the use of cut-off points to describe the upper or lower tail of the distribution. The distribution of individuals at the extremes of a distribution can be characterized by using standard cut-off points and tabulating the prevalence of values below or above cut-off values. Several cut-off points may be used to provide an impression of the magnitude of the problem occurring at different levels of the distribution. For example, lower cut-off points may be selected to highlight the most extreme cases, while higher cut-off points may be useful to capture the proportion of the population that may be at risk of an inadequate iodine status, but not necessarily severely deficient.

6. The Selection of Appropriate Indicators (What to Measure)

34. Indicators are basically of two types: outcome indicators, that provide a measure of IDD status, and process indicators, that measure the condition or progress of implementation of an IDD control programme. Outcome indicators may refer to exposure (as urinary iodine excretion, reflecting iodine intake), or to impact, whether on morphology (as estimated by thyroid size, or thyroglobulin - Tg) or function (estimated by thyroid hormones, or thyrotropin - TSH). Alternatively, outcome indicators can be categorized according to whether the assessments are clinical or biochemical. Once the target population(s) for assessment are defined, the selection of particular indicators should be based on consideration of the criteria discussed below, as well as on the specific objectives of the surveillance.

6.1 Criteria for Indicator Selection

35. **Acceptability:** The acceptability of an indicator for a particular target population is a critical factor to be considered in indicator selection. Some measures, such as assessment of thyroid size by palpation, may be widely acceptable. Other measurements, such as drawing venous blood for biochemical determinations, may be quite unacceptable, especially in certain target groups such as infants and children. Acceptability is also an issue for field staff performing the tests. Again, drawing blood samples in populations with a high prevalence of HIV infection involves some level of risk, or perceived risk, that should be considered in indicator selection.

36. **Technical Feasibility:** Technical feasibility involves a number of factors including:

- ease of data or sample collection
- sample storage and transport requirements
- transportability and ruggedness of field equipment
- availability of personnel to obtain specimens, e.g. phlebotomist.

37. **Cost:** Costs associated with the use of certain indicators include:

- capital costs for facilities and equipment
- recurring costs for supplies and reagents (cost per test)
- maintenance costs
- training costs
- personnel, administrative and related expenses.

38. **Performance:** Another criterion for indicator selection is the performance in identifying IDD status. Useful measures of indicator performance include sensitivity, specificity, and reliability.

39. **Interpretation and Availability of Reference Data:** Interpretation of IDD status depends on the availability of reference data. Reference data assist in establishing cut-off values and prevalence levels for use in identifying public health problems. The presence of reference data is useful in selecting indicators and target groups for surveillance, as these will enhance interpretation across different studies.

6.2 Outcome Indicators of IDD

Clinical Indicators

Thyroid Size

40. The size of the thyroid gland changes inversely in response to alterations in iodine intake, with a lag of 6 to 12 months. The traditional method for determining thyroid size has been palpation. Ultrasonography provides a more precise and objective method. Both methods are described below. Issues common to palpation and ultrasound are not repeated in the section on ultrasound.

Palpation

Feasibility

41. Palpation of the thyroid is important in assessing the prevalence of goitre. The costs are associated with mounting a survey and the training of individuals, which is relatively easy to conduct. Costs will vary depending upon the availability of health care personnel. Feasibility and performance vary with the different target groups.

42. *Neonates:* It is neither feasible nor practical to assess goitre by either palpation or ultrasound. Performance is poor.

43. *School-aged children:* Preferably children 8 to 10 years of age should be studied, but if an adequate number to ensure statistical precision cannot be obtained the range may be extended from 6 to 12 years of age. The smaller the child, the smaller its thyroid and the more difficult it is to perform palpation: thus there is a practical reason for not measuring very young age groups. It is important to provide guidelines for field managers on the percentage of attendance in schools required in order to gain a representative assessment of IDD prevalence. In communities where low school attendance rates pose a problem, household surveys may be useful for identifying school-aged children.

44. *Adults:* Pregnant and lactating women are of most concern. Pregnant women are especially sensitive to marginal iodine deficiency. Pregnant women are relatively accessible, as they often come to maternal clinics. They are also the prime target group for IDD control activities.

Interpretation

45. A modification to the previous classification system for goitre, in which five grades were defined, is recommended. The old grades 1A and 1B are combined into one group, and grades 2 and 3 are combined into another (Table 2). Table 3 gives the epidemiologic criteria for establishing the severity of IDD based on the prevalence of goitre in school-aged children. It should be understood that "mild" is a relative term, and does *not* imply that this category of IDD is of little importance.

46. It is recommended that a total goitre rate (TGR, goitre grades 1 and 2) of 5% or more in primary school children (age range approximately 8 to 10 years) be used to signal the presence of a public health problem. This recommendation is based on the observation that in a normal, iodine replete population the prevalence of goitre should be quite low, generally less than 1%. The cut-off of 5% allows some margin for the inaccuracy of goitre assessment and for goitre that may occur in iodine-replete populations due to other causes such as goitrogens and autoimmune thyroid diseases. The cut-off level of 10% that was previously recommended has been revised downward because it has been found that goitre prevalence rates between 5% and 10% may be associated with a range of abnormalities, including subnormal circulating levels of thyroid hormones and elevation of TSH in the population, clearly representing a public health concern.

47. The specificity and sensitivity of palpation are low in grades 0 and 1 due to a high inter-observer variation resulting in misclassification of up to 40%, as was demonstrated in many studies with experienced examiners. This diminishes the usefulness of palpation in populations with few visible goitres (grade 2). Ultrasonography will provide a more precise measurement of thyroid size.

Table 2. Proposed Classification of Goitre

Grade 0:	No palpable or visible goitre.
Grade 1:	A mass in the neck that is consistent with an enlarged thyroid that is <i>palpable but not visible</i> when the neck is in the normal position. It moves upward in the neck as the subject swallows. Nodular alteration(s) can occur even when the thyroid is not enlarged.
Grade 2:	A swelling in the neck that is <i>visible when the neck is in a normal position</i> and is consistent with an enlarged thyroid when the neck is palpated.

Table 3. Proposed Epidemiological Criteria for Assessing the Severity of IDD based on the prevalence of goitre in school-aged children.

	Mild IDD	Moderate IDD	Severe IDD
Prevalence of goitre (TGR)	5-19.9%	20-29.9%	≥30%

Ultrasonography

Feasibility

48. Ultrasonography is non-invasive and safe. It is a specialized technique in which individuals have to be trained. A trained worker can perform up to 200 examinations per day. The thyroid volume can be easily calculated using a calculator or a microcomputer during data entry. The portable ultrasound equipment is relatively rugged but requires electricity; it can be operated from a car battery using a transformer.

Cost

49. Portable ultrasound equipment with a 5 MHz transducer currently costs US \$12,000. This price is expected to decline with the availability of smaller machines. A one-week training programme is available.

Performance

50. Compared to palpation, ultrasonography provides a more precise measurement of thyroid volume (Table 4). This becomes especially significant when the prevalence of visible goitres is small, and for monitoring iodine control programmes where it would be expected that thyroid volumes will decrease over time. For practical reasons, school-aged children between the ages of 8 to 10 years should be examined, although this range can be extended to 6 to 12 years if necessary. In younger children the thyroid is more difficult to examine, especially in children less than six years of age, when a 7.5 MHz transducer is required to get an adequate resolution.

Table 4. Comparison between palpation and ultrasound in grading small thyroids.

	Goitre grade			
	0	1a	1b	2
Number of subjects graded by palpation	105	88	101	10
Subjects with grade confirmed by ultrasound				
Numbers	63	72	89	10
Percentage	60	82	88	100

Interpretation

51. Results of ultrasonography from a population under study should be compared to normative data (collected in Africa, Europe, Latin America and South-East Asia) from populations with sufficient iodine intake (above 150 µg iodine per day). Normative data are presented in Table 5. In an iodine-replete population, the expected prevalence of thyroid sizes greater than the 97th centile would be 3%, and this figure can be compared with the observed prevalence. In addition, the median (50th centile) thyroid volume may be useful.

Table 5. Normative Thyroid Volume Size (ml)

	Centile	
	50th	97th
Women	8.0	18.0
Men	10.0	25.0
Children (ages in years)		
17	7.9	16.0
16	6.8	14.0
15	5.8	12.0
14	5.0	10.5
13	4.3	9.0
12	3.8	8.0
11	3.2	7.0
10	2.8	6.0
9	2.3	5.0
8	2.0	4.5
7	1.6	4.0
6	1.4	3.5

Cretinism

Biological features

52. Endemic cretinism, the most dramatic consequence of iodine deficiency, is caused by thyroid hormone deficiency during development of the central nervous system. The manifestations of endemic cretinism include mental retardation, a pathognomonic gait, squinting of the eyes and deafness. In some populations, continued postnatal thyroid hormone deficiency results in the classic clinical signs of hypothyroidism including severe growth retardation, skeletal retardation, and sexual immaturity.

Feasibility

53. Because cretinism is a clinical diagnosis of a disorder with a spectrum of presentation from mild to devastatingly severe, it is difficult to identify all of the affected individuals in a population. In fact, the more mildly affected cretins (sometimes referred to as subclinical cretins) may not be diagnosed except by clinical experts. A significant amount of time may be required to perform the necessary physical examination.

Performance

54. Cretinism prevalence is not a sensitive indicator of a population's iodine status. As noted above, identification of cases is difficult and may require expert clinical skills. The more severely affected cretins, though easily identifiable, probably only represent the 'tip of the iceberg' in terms of IDD case finding.

Interpretation

55. Cretinism prevalence may be most interesting as a historical record of a community's exposure to severe iodine deficiency. While cretinism results from deficiencies during intrauterine life and early childhood, it is most easily diagnosed in later childhood and adulthood. Therefore, in a qualitative sense, the presence of cretins in a community, even if the prevalence is very low, is significant because this documents that individuals were exposed to a marked environmental iodine deficiency sometime in the past. However, this does not necessarily reflect the current iodine status of the population, although it may have considerable advocacy value.

Biochemical Indicators

Urinary Iodine

Biological features

56. As most iodine is excreted in the urine, the urinary iodine level is a good marker of dietary iodine. But since an individual's urinary iodine varies day by day the data can only be used for a population estimate.

Feasibility

57. Acceptability is very high and spot urine samples are easy to obtain. They give as accurate an estimate for populations as do 24 hour collections.

58. Collection and transport: small amounts (0.5 – 1.0 ml) are required and may be collected in tubes with screw tops, and do not need refrigeration. The iodine content remains stable throughout transport to the laboratory. Samples can be stored in the laboratory for several months before the actual determination is made. The specimens can be refrigerated upon arrival in the laboratory to avoid nuisance.

59. Determination: many techniques have been used for the analysis (a manual which describes these in detail has been prepared by the ICCIDD). The simplest methods, quite adequate for epidemiological surveys, require less than one milliliter of urine. The sample is digested in chloric acid and its iodine content is measured by its catalytic action in the reduction of ceric ammonium sulfate (yellow) to the cerous (colorless) form. The result is expressed as a concentration ($\mu\text{g I/dl}$ urine). Relating it to urine creatinine is cumbersome, expensive, unreliable and unnecessary. A single technician can do at least 150 samples per day. The total instrument cost is about US\$ 3,000, and the total cost per sample has been estimated to be \$ 0.50-1.00 including labour costs. Since casual samples are used, it is important to measure at least 75 samples from a group to allow for different degrees of subject hydration and other biological variations between individuals. Other methods digest the urine more completely, but are complicated and take more time and cost to perform.

Performance

60. The recommended methods are able to detect urinary iodine levels as low as 0.5-2 $\mu\text{g/dl}$ with a coefficient of variation under 10%. Laboratory techniques require training, but are not difficult. ICCIDD provides a network of reference laboratories. Meticulous attention to avoid iodine contamination at all stages is essential. As in all surveys for estimating prevalence, population samples must be representative.

Interpretation

61. The simple methods make it feasible to do large numbers of samples and to characterize the distribution according to a number of cut-off points and intervals. The cut-off points proposed for the classification of iodine deficiency into different degrees of public health significance are shown in Table 6. For full interpretation frequency distribution curves are necessary.

62. As to sample size, it has been advised elsewhere (ICCIDD Manual no. 3) that a minimum of 40 is required. This is enough to give an indication that an endemic is indeed associated with low urinary iodine intakes; if the values were not low, one could suspect for instance a role of

goitrogens. But for purposes of monitoring, this sample size would not be enough, as explained in the next paragraphs.

63. If the distribution of urinary iodine values were normal and the variance $10 \mu\text{g/dl}$ (squared), a simple random sample size of 40 would yield a 95% confidence limit of $\pm 1.0 \text{ m}\mu/\text{l}$ of the mean. If a cluster sampling method is used, to achieve similar confidence limits, this would require a sample size of 80 or 120 (i.e. a design factor of 2 or 3).

64. However in practice the distribution of urinary iodine values is often far from "normal" with appreciable numbers of very high and/or very low values. So in principle, for urinary iodine values, one quotes median values rather than a mean and standard error. The indicator of "elimination" is a median value for iodine concentration of $10 \mu\text{g/dl}$, i.e. 50% of the samples should be above $10 \mu\text{g/dl}$. The same Tables as given for prevalence of goitre rates apply, in principle (see Annex 1). The "design effect" applicable for goitre prevalence surveys (see Annex 1) is probably also applicable for urinary iodine values, at least in an initial survey, but may decrease over time, with implementation of a salt iodisation programme.

65. As an IDD prevention programme progresses, goitre rates become progressively less useful as criteria of elimination, and urinary iodine levels relatively more important. But the dispersion of urinary iodine values should become less, so a design effect of 2 may be acceptable under these conditions. If a relative precision of 20% above or below 50% is required (i.e. the required interval is 40-60%), the number of subjects required is $96 \times \text{Design effect}$, i.e. provisionally 192, say, 200. Future experience should provide a more adequate basis for predicting required sample sizes.

66. At a relatively advanced stage in an IDD prevention programme there would probably not be need for district-by-district survey data and a province may be the unit of coverage. Hence the overall number of samples required in a given country will not be excessive.

Table 6. Proposed Epidemiological Criteria for Assessing the Severity of IDD based on median urinary iodine levels.

Median value ($\mu\text{g/dl}$)	Stage of IDD
<2.0	Severe IDD
2.0-4.9	Moderate IDD
5.0-9.9	Mild IDD
≥ 10.0	No deficiency

Blood constituents

67. The measurement of two blood constituents are discussed in this document, thyroid stimulating hormone (TSH) or thyrotropin, and thyroglobulin (Tg). In a population survey, the collection of a blood spot on filter paper could be used to measure TSH and Tg. Issues common to TSH and Tg will not be repeated in the section on Tg. The thyroid hormones thyroxine (T4) and triiodothyronine (T3) are not discussed, since their measurement is not recommended for surveillance purposes.

Thyroid Stimulating Hormone (TSH)

Biological features

68. Iodine is essential for the synthesis of thyroid hormones which are necessary for normal brain and neurological development. The kinetics of the thyroid hormone receptor in the pituitary gland mimic the kinetics of thyroid hormone receptors in the brain. At low enough levels, the concentration of thyroid hormones in the pituitary gland stimulates the release of TSH which is then detectable in blood. Therefore, serum or whole blood TSH levels directly reflect the availability and adequacy of thyroid hormone. A TSH level is the best diagnostic test for the determination of hypothyroidism. An elevated TSH level in newborn or infant blood is of fundamental concern because this indicates that the level of thyroid hormone is inadequate during this critical time of brain development.

Feasibility

69. The methodology for the determination of TSH levels, from both dried whole blood spots on filter paper and serum, is well established. Whole blood from any source is acceptable for spotting onto filter paper of a certified grade. Because only a few drops of whole blood are required, a finger, heel, or earlobe prick are the most common sources. For subjects, use of sterile equipment (either lancets for blood spot collection or needles

and syringes for serum collection) is important. For the persons collecting and processing specimens, standard procedures for handling blood products or objects contaminated with blood should be followed. The risk of HIV or hepatitis infection from dried blood spots is extremely low.

70. **Timing:** Among newborns, blood samples can be collected from cord blood at the time of birth or heel prick of infants after 3 days (72 hours). Collection during the first 3 days (<72 hours) of life from the heel is not recommended because levels could be elevated as a result of the birth process itself. After the first three days of life, the timing of the specimen is not critical. Blood samples may be obtained from pregnant women during prenatal care visits, or from school-aged children during school-based surveys. However, further study of the distributions of TSH among these older populations is needed to better understand their inter-relationship.

71. **Transport:** The blood spots are easy to transport. It is important that the spot be dry before storage or shipment. Filter papers, usually stored in a plastic bag, can be transported in the normal postage system and are stable for periods of up to 6 weeks even in a hostile environment of high temperature and humidity. Customs clearance may be required for the international transport of dried blood spots.

72. **TSH Assay:** TSH in the blood spot can be measured by commercially available assay kits. The enzyme-linked immunosorbant assay (ELISA) methodology is recommended because of lower equipment cost, longer shelf life of reagents (6 months) and high sensitivity (<2 mU/l). Direct linkage of the ELISA system to a microcomputer is recommended because this allows a high throughput and facilitates data management for quality assurance and public health decision making. Laboratory staff need training in laboratory management and quality assurance. Laboratories should participate in an external quality control programme. (An international programme is available through PAMM, based at the Centers for Disease Control, Atlanta, Georgia, USA.)

73. **Cost:** Current estimates for the costs of laboratory equipment and reagents are given below. Labor costs of collection and processing will vary with local economies.

- TSH ELISA laboratory with state-of-the-art computer based system and software capable of processing 90,000 tests / year / technologist
US\$ 15,000
- TSH ELISA laboratory hardware & software capable of processing up to 5,000 tests / year / technologist
US\$ 5,000

- TSH assay kits including titer plate, collection materials (paper & lancets), and reagents procured in the United States

US\$ 0.50-1.00 / test.

The same equipment can be used to perform ELISA based assays for surveillance of other diseases including various micronutrients and infectious diseases.

Performance

74. **Sensitivity:** The term 'sensitivity' has two distinct meanings depending on whether it is used in a laboratory or epidemiologic context. In a laboratory context, 'sensitivity' refers to the improved ability of the TSH whole blood spot assay kits to detect levels of TSH into the full physiologic range including values as low as 1-2 mU/l. Earlier TSH assay systems were sensitive and reliable only for values over approximately 20 mU/l. The relatively recent commercial availability of the 'sensitive' TSH blood spot assay kits now permits the determination of mild to moderate IDD which may present with TSH levels less than 20 mU/l.

75. In an epidemiologic context, 'sensitivity' refers to the ability of the TSH assay to identify cases of IDD among populations. 'Specificity' refers to a screening test's ability to correctly identify those people without the disease. The specificity of TSH for the screening of IDD has not been clearly quantified. However, the number of false negative tests (individuals with IDD who test negative) is probably very low. Using the newer sensitive whole blood TSH assays, individuals with mild IDD and mild elevations of TSH can now be detected. Elevation of TSH values in individuals is associated with all causes of primary hypothyroidism. Causes of TSH elevation other than iodine deficiency include goitrogen ingestion, congenital hypothyroidism (CH), and autoimmune thyroiditis. CH is relatively uncommon - approximately 1 in 4,000 newborns around the world. Autoimmune thyroid disease is less uncommon, particularly in western countries. Goitrogen ingestion is usually regional and easily identified in the environment, and in any case does result in a relative deficiency of iodine because requirements are greatly increased. IDD screening programmes are not designed to follow-up individuals but to direct population-based interventions. TSH is an excellent indicator for neonatal hypothyroidism, but its efficacy in older groups is less certain.

Interpretation

76. Reference data for TSH are available among newborns because this data is routinely collected as part of newborn congenital hypothyroid screening programmes. TSH values are currently reported in whole blood units or serum units. It is critical that all reports and discussion of TSH distributions specify the units employed. Congenital hypothyroid screening and IDD surveillance require different TSH cut-offs. A TSH cut-off of

20-25 mU/l whole blood (approximately 40-50 mU/l serum) is commonly used to screen for congenital hypothyroidism. IDD may be present with TSH levels which are only mildly elevated. While further study of iodine-replete populations is needed, a cut-off of 5 mU/l whole blood may be appropriate for epidemiologic studies of IDD. Populations with a substantial proportion of newborns with TSH levels greater than the cut-off could indicate a significant IDD problem (see Table 7).

Thyroglobulin (Tg)

Biological Features

77. Insufficient iodine intake induces the proliferation of thyroid cells, which results in hyperplasia and hypertrophy. This leads to an enhanced turn-over of thyroid cells, which release Tg into the serum. Tg in serum changes inversely with iodine intake in all age groups.

Feasibility

78. The acceptability is high, and the collection and transport of samples are simple, identical with those for TSH (see previous section). The technique for determination is similar to that for TSH using a Tg antibody instead of one for TSH, although methods are not yet commercially available. The costs are comparable to those for TSH. Training will be required in laboratory techniques.

Performance

79. The available methods can detect levels as low as 2 ng/ml serum with a coefficient of variation close to 5%. Tg changes rapidly after an alteration of iodine intake in *all* age groups, and may be a more sensitive indicator than TSH, as the following observations suggest. Tg rises in individuals with an insufficient iodine intake, even under conditions where TSH falls or is suppressed due to functional autonomy as frequently happens with longstanding iodine deficiency. After iodine depletion, Tg will rise before TSH shifts to higher values and long before goitre develops. Following iodine supplementation Tg normalizes before the thyroid volume has decreased.

Interpretation

80. Individuals (children and adults) with sufficient iodine intake show a median Tg serum level of 10 ng/ml and an upper limit of normal of 20 ng/ml in most assay techniques. The results obtained from a survey should be expressed as a median and as the percentage of Tg levels above 20 ng/ml.

Summary of Outcome Indicators

81. Table 7 provides a summary of cut-off points and prevalences that are considered indicative of potential public health problems in regard to IDD.

Table 7. Summary of IDD Prevalence Indicators and Criteria for a Public Health Problem

Indicator	Target Population	Severity of Public Health Problem (prevalence)		
		Mild	Moderate	Severe
Goitre Grade > 0	SAC*	5.0-19.9%	20.0-29.9%	≥30.0%
Thyroid volume >97th centile by ultrasound	SAC	5.0-19.9%	20.0-29.9%	≥30.0%
Median urinary iodine level (µg/dl)	SAC	5.0-9.9	2.0-4.9	<2.0
TSH >5 mU/l whole blood	newborns	3.0-19.9%	20.0-39.9%	≥40.0%
Median Tg (ng/ml serum)	C/A**	10.0-19.9	20.0-39.9	≥40.0

*SAC = school-aged children **C/A = children and adults

6.3 Process Indicators of IDD Control Programmes

82. As programmes are implemented for the control of iodine deficiency disorders in a country, it is important to establish mechanisms for monitoring and evaluating these activities. Included in such monitoring protocols will be both indicators associated with the process of the programmes, as well as indicators of the impact realized through the implementation of the control activities. Depending upon the specific characteristics of the IDD control programmes, distinct indicators will need to be considered and different techniques employed in their monitoring. Although most countries will have iodized salt as the primary control activity, other programmes, where implemented, will need to be monitored, even if they are used as short-term measures while salt iodization is being established.

Salt Iodization programmes

83. All countries with a significant public health IDD problem should undertake a situation analysis of salt available for human and animal consumption, from points of production (or importation) through distribution channels to the consumption level. Such salt is referred to as *food grade salt*, which includes crude salt for direct edible use by people, and for

livestock, and refined salt for edible use and for industrial use in most processed foods (such as biscuits, instant foods, etc., but not in some sauces).

84. The situation analysis should include a list of the major salt producers or importers, production/import/export statistics, and information on salt quality, packaging, transport and storage, retail marketing, prices and household consumption. These data need to be updated periodically, e.g., every two years. In addition, periodic monitoring will have to be undertaken at different points along the distribution chain of iodized salt, to check that the iodine concentration levels are adequate, and if necessary to prompt corrective actions. Such actions may include the improvement of packaging, transport or storage of iodized salt at different points before it reaches the household.

85. The iodization of salt involves the addition of a small quantity of iodine (30 to 100 mg of iodine per kg of salt, or parts per million - ppm), usually in the form of potassium iodide or potassium iodate. The Joint FAO/WHO Expert Committee on Food Additives noted in 1990 that "potassium iodate has been shown to be a more suitable substance for fortifying salt than potassium iodide, because of its greater stability, particularly in warm, damp or tropical climates" (WHO Technical Report Series No. 806, Annex 5).

Techniques for measuring salt iodine levels

86. There are essentially two techniques for measuring iodine levels in salt:

- i) *Standard titration method* - This method is conducted in laboratories. Iodine is liberated using sulphuric acid. The free iodine is titrated with sodium thiosulphate, using starch as an indicator. Slightly different techniques are employed, depending on whether the iodine is in the form of iodate or iodide. Facilities for this method are normally available in most countries in a public health or standards laboratory, but some other laboratory may need to be equipped and enabled to develop competence in its operation. The titration method is preferred for checking of salt batches produced in factories, or on arrival in the country, and in general where accurate testing is required. But it is too time-consuming and expensive for purposes of routine monitoring throughout the country, for which the second method may be more suitable.
- ii) *Rapid-test kits* - These comprise bottles of starch solution (stabilized), of which one drop is placed on the salt. If the salt is alkaline a neutralizing solution is first applied. The intensity of the blue colour which develops indicates the salt iodine level, up to 50 or 100 ppm, depending on the kit used, with an accuracy of ± 10 ppm. Most of the rapid-test kits which are presently available can detect the presence of

iodate only. Details of available kits may be obtained from WHO, UNICEF or ICCIDD.

Monitoring salt iodine levels

87. Iodization may take place inside the country, at the main production sites, or outside, with already-iodized salt being imported. There is inevitably some loss of iodine from salt between the points of production and consumption, and these losses are often much higher in non-industrialized countries, where the quality of salt packaging, storage and transportation conditions are suboptimal. Countries should establish the expected iodine content of salt at different levels of the distribution system, taking account of climatic conditions, types of packaging and habitual daily salt consumption. Guidelines are given in Table 8, calculated in each case to provide individuals with 150 µg iodine per day. The iodine concentration in salt needs to be monitored at two or three levels on a regular basis. The overall responsibility for quality control within a country should be clearly identified: often it will be vested in the Ministry of Health through its Primary Health Care Department and regional or provincial and district health departments. Criteria for assuring the adequacy of programmes are proposed in Table 9.

Table 8. WHO-UNICEF-ICCIDD-Recommended Level of Iodine in Salt

Parts of iodine per million of salt, i.e. microgrammes per gramme,
milligrammes per Kg
grammes per tonne

Climate and Daily salt consumption (g/person)	Packaging			
	Required at factory (external)		Required at retail sale	
	Bulk (Sacks)	Retail (Plastic bags)	Bulk (Sacks)	Retail (Plastic bags)
Warm moist 5 g	100	80	80	60
10 g	50	40	40	30
Cool dry: 5 g	80	60	60	40
10 g	40	30	30	20

N.B. 168.6 mg of KIO_3 contains 100 mg of iodine.

Table 9. Criteria for Assessing Adequacy of Salt Iodization programmes

Process Indicator	Criterion of Adequacy
A. Factory or importer level	
1. Percent of food grade salt claimed to be iodized	100%
2. Percent of food grade salt effectively iodized	≥ 90%
3. Adequacy of internal monitoring process	≥ 90%
4. Adequacy of external monitoring process*	10-12 monthly checks per producer/importer, per year
B. Consumer and district level	
1. Percent of monitoring sites with adequately iodized salt	Adequate in 90% of samples
i) households (or schools)	
ii) district headquarters (including major markets)	
2. Adequacy of monitoring process**	90% or more

*Corrective action systematically taken within 3 hours in 90% of cases, following the lot quality assurance methodology

**Monitoring undertaken in 90% of districts in each province, at both household and district level

Procedures for monitoring

88. Factory level: The responsibility for routine monitoring rests with the factory itself (internal salt monitoring). The recommended procedure is to carry out hourly monitoring with the rapid-test kit, and at least once daily by titration. The observations should be recorded systematically in a register indicating the date, time, batch number and iodine content of the salt. However, the government (Ministry of Health, Industry or Commerce, or the Standards Organization or other designated body) must maintain periodic checks (external salt monitoring) as well: at least once monthly, by titration, comparing the result with the factory's own test. During inspection, the manufacturer's records should be verified for adequacy of internal monitoring and variations in iodine levels.

89. Distributor and wholesaler level: The major distributors should be sensitized on the subject and provided with rapid-test kits to check the iodine levels in the salt before it is released for retail sale. This is especially important in larger countries or in situations where transportation results in a long time lag between production and consumption. Regular monitor-

ing at three monthly intervals may be advisable: if there are deficiencies, they should be notified to the provincial or district health department.

90. **Community and district level:** Public health inspectors or nurses at district level will often be responsible for this monitoring activity, the objective of which is to verify by the use of rapid-test kits that adequate concentrations of iodine are attained in salt, especially at the consumer level. Salt monitoring at the district and community level should be used as an opportunity for education and the promotion of information regarding the importance of consuming iodized salt. When checks at those levels show inadequate salt iodine concentrations, further spot checks should be made at successively higher levels to identify at what point losses are occurring.

91. Three approaches are recommended for the monitoring of salt iodization quality at the district level, based on markets, on schools and on household surveys. In the two former, more specific information may be obtained about the distribution of inadequate salt samples within the district, which may be used to highlight 'hot spots' in which problems are likely to be occurring. For schools with an enrollment of 100 to 1000 pupils at least 35 salt samples will have to be collected and tested to detect if more than 20% of the population they serve have access to inadequately iodized salt (indicated if 4 or more samples are inadequate). In monitoring based on household sampling, it will be possible to ascertain whether there is a problem for the district as a whole, but it would not be possible to gain any information about the patterns within the district. For each district at least 10 houses in each of 10 remote villages should be randomly selected for spot testing, with a new selection if possible every 4 months.

Other IDD Control programmes

92. Other potential IDD control programmes include iodization of water, fortification of foods other than salt, and supplementation with iodized oil. If other foods are fortified, then an evaluation of that food, similar to the one described for iodized salt, should be devised. In supplementation programmes, it is important to ensure that the high-risk population has an adequate level of coverage. LQAS surveys can be useful (see paragraph 28).

7. Criteria for Tracking Progress Towards World Health Assembly/ World Summit Goals

93. Table 10 presents criteria that are recommended as core indicators to be used for purposes of monitoring national progress towards the goal of virtual elimination of IDD, endorsed by the World Summit for Children. These criteria include both IDD status indicators and an IDD control programme process indicator, because it is important to ensure the sustainability of the control of iodine deficiency in the whole population rather than focusing entirely on reaching goals on the basis of measuring IDD status of a particular target group. Furthermore, monitoring the status of salt iodine is a useful first step in tracking progress towards the goal of elimination. It is recommended that countries use this process indicator, as well as one indicator of thyroid size and one biochemical indicator.

Table 10. Criteria for Tracking Progress Towards the World Health Assembly/ World Summit Goals: Elimination of IDD as a Public Health Problem

Indicator	Goal
1. Salt iodine Proportion of food grade salt effectively iodized	> 90%
2. Thyroid size In school children 8-10 years of age: Thyroid volume by ultrasound, Proportion > 97th centile OR Total goitre rate by palpation	< 5% < 5%
3. Biochemical TSH: Proportion of newborns with levels > 5mU/l whole blood OR Median urinary iodine (mcg/dl)	< 3% > 10.0

Annex 1 List of Participants

Dr D. Alnwick
Senior Nutrition Advisor
UNICEF
New York, New York 10017
USA

Dr J. Dunn
Box 511, University of Virginia
Charlottesville, VA 22908
USA

Dr J. Gorstein
University of Michigan
Ann Arbor, MI
USA

Dr P. Greaves
2 The Plantation
London SE3 0AB
UK

Dr R. Gutekunst
Im Felde 10
W-2430 Neustadt
Germany

Dr B.S. Hetzel
c/o Health Development Foundation
8th Floor, Samuel Way Building
Adelaide Medical Centre for Women
and Children
72 King William Road
North Adelaide
5006 Australia

Dr G. Maberly
Emory University School
of Public Health
1462 Clifton Road, NE
Atlanta, GA 30322
USA

Dr V. Mannar
87 Hillcroft Drive
Etobicoke
Ontario M9B 4X8
Canada

Dr G. Ndossi
Tanzania Food and Nutrition Centre
P.O. Box 977
Dar es Salaam
United Republic of Tanzania

Ms S. Pak
University of Michigan
Ann Arbor, MI
USA

Dr C.S. Pandav
Department of Medicine &
Endocrinology
Centre for Community Medicine
All India Institute of Medical Sciences
Ansari Nagar
New Delhi 110029
India

Dr C. Thilly
Ecole de Santé Publique
Case postale 590
Route de Lennik 808
1070 Bruxelles
Belgium

Dr F. Trowbridge
Director, Division of Nutrition
National Centers for Chronic Disease
Center for Disease Control
Atlanta, GA 30333
USA

WHO Secretariat

Dr. K.V. Bailey, Nutrition, Geneva
Dr G.A. Clugston, Nutrition, Geneva
Dr N. Cohen, Expanded Programme on
Immunization, Geneva
Dr A. Verster, Regional Advisor,
Nutrition, Eastern Mediterranean
Regional Office

Annex 2 Sample Sizes for IDD Prevalence Surveys

Confidence interval. For epidemiological surveys, normally one works with a confidence interval of $\pm 95\%$. This means that one accepts the possibility of an error, outside the range of precision decided upon (see below), five times out of a hundred, i.e. once in twenty surveys.

Precision is a measure of how close an estimate is, or is required to be, in relation to the true mean value of a population parameter.

- i) Table 1 gives sample sizes for a given precision in terms of *absolute precision* of the survey result – mean or median, expressed as a proportion (or percentage). The vertical columns labelled “p” give the anticipated (or actual) prevalence in terms of proportion, i.e. 0.05 means 5%, 0.95 = 95%. The horizontal rows – “d” – give the desired (or actual) precision in absolute terms; i.e. expressed as absolute proportion (or percentage points) above and below the mean – for instance for row d = 0.10 this means ± 0.10 (i.e. 10 percentage points); if the anticipated value of “p” is 50% ($p=0.50$), the number of subjects to be chosen for the resulting estimate of “p” to lie between 50% \pm 10% (40% and 60%) is 96.
- ii) Table 2 gives the sample size in terms of *relative precision*, i.e. expressed as a proportion of the mean value (p) expected or obtained in the survey. The vertical columns are the same as in Table 1. The horizontal rows “e” give the desired (or actual) precision, in terms relative to (i.e. as a proportion of) the observed (p) value; for instance, for row e = 0.10, if the observed mean is 50% ($p=0.50$), the relative precision is $0.10 \times 0.50 = 0.05$ (or 5 percentage points); the number of subjects to be chosen for the resulting estimate of “p” to lie between 50% \pm 5% (45% and 55%) is 384.

Thus for instance an *absolute precision* of 10% means ten percentage points above or below the observed mean value, e.g. 50% \pm 10%.

A *relative precision* of 10% means a width of 10% of the observed mean value, above or below it; if the mean was 50%, the width for 10% relative precision is $\pm 5\%$.

It is usually better to use the *relative precision*. For instance while an absolute precision of 5 percentage points above or below 50% may be meaningful and appropriate, 5 percentage points (in absolute terms) above or below a level of 5% is not very meaningful or appropriate.

Depending on the type and circumstances of a survey, the degree of precision required may vary, e.g. in an initial goitre survey, with expected result of about 50%, a relative precision of 10% may be appropriate, i.e. 50% \pm 5%. If salt iodization has been implemented for some years and goitre rates are likely to be mostly down around 5%, a relative precision of 30% (which would mean an eventual “width” of $\pm 1.67\%$) may be adequate.

If the anticipated prevalence is entirely unknown, sample size should be estimated assuming that the result will be 50%, since this gives the largest sample size, compared with other outcomes.

Design effect. The figures in Tables 1 and 2 refer to a survey with strict random sampling of a given population. Very often, cluster sampling procedures are used, and are more practical – this is the procedure followed for example in many EPI and CDD surveys; and also in anthropometric and goitre surveys. For phenomena such as goitre whose distribution is patchy, cluster sampling may produce misleading results – if one happens to fall on clusters of high prevalence, for example. To avoid errors of this sort a larger number of subjects must be examined, in total. The numbers in the basic tables (1 and 2) should be multiplied by a factor called the “*design effect*” to allow for this possible lack of homogeneity in the population studied. It is an indication of the variation due to clustering. It is estimated by the ratio of the variance when *cluster sampling* is used, to the variance when *simple random sampling* is used. The design effect has been established at 3 for goitre surveys.

In all cases where the parameter is not normally distributed (i.e. Gaussian distribution), e.g. for urinary iodine excretion values, it is wiser to give median rather than mean values, or another form of presentation of the distribution of values, rather than the mean.

For other confidence intervals and more details, including guidance on other types of studies than simple prevalence surveys, including lot quality assurance sampling, please see:

S.K. Lwanga & S. Lemeshow (1991) *Sample size determination in health studies, a practical manual*, WHO Geneva.

(This book is a practical guide to the subject, with a minimum of background theory.)

S. Lemeshow et al. (1990) *Adequacy of sample size in health studies*, Chichester, John Wiley

(This book includes the statistical methodology of sample size determination.)

Table 1

Table 1. Estimating a population proportion with specified absolute precision

$$n = z_{1-\alpha/2}^2 P(1-P)/d^2$$

P	(a) Confidence level 95%																			
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	
0.01	1825	3457	4898	6147	7203	8067	8740	9220	9508	9604	9508	9220	8740	8067	7203	6147	4898	3457	1825	
0.02	456	864	1225	1537	1801	2017	2185	2305	2377	2401	2377	2305	2185	2017	1801	1537	1225	864	456	
0.03	203	384	544	683	800	896	971	1024	1056	1067	1056	1024	971	896	800	683	544	384	203	
0.04	114	216	306	384	450	504	546	576	594	600	594	576	546	504	450	384	306	216	114	
0.05	73	138	196	246	288	323	350	369	380	384	380	369	350	323	288	246	196	138	73	
0.06	51	96	136	171	200	224	243	256	264	267	264	256	243	224	200	171	136	96	51	
0.07	37	71	100	125	147	165	178	188	194	196	194	188	178	165	147	125	100	71	37	
0.08	29	54	77	96	113	126	137	144	149	150	149	144	137	126	113	96	77	54	29	
0.09	23	43	60	76	89	100	108	114	117	119	117	114	108	100	89	76	60	43	23	
0.10	18	35	49	61	72	81	87	92	95	96	95	92	87	81	72	61	49	35	18	
0.11	15	29	40	51	60	67	72	76	79	79	79	76	72	67	60	51	40	29	15	
0.12	13	24	34	43	50	56	61	64	66	67	66	64	61	56	50	43	34	24	13	
0.13	11	20	29	36	43	48	52	55	56	57	56	55	52	48	43	36	29	20	11	
0.14	9	18	25	31	37	41	45	47	49	49	49	47	45	41	37	31	25	18	9	
0.15	8	15	22	27	32	36	39	41	42	43	42	41	39	36	32	27	22	15	8	
0.20	5	9	12	15	18	20	22	23	24	24	24	23	22	20	18	15	12	9	5	
0.25	•	6	8	10	12	13	14	15	15	15	15	15	14	13	12	10	8	6	•	

*Sample size less than 5.

Table 2

Table 2. Estimating a population proportion with specified relative precision

$$n = z_1^2 \frac{p(1-p)}{e^2 P}$$

(a) Confidence level 95%

P	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	729904	345744	217691	153664	115248	89837	71344	57624	48953	38416	31431	25611	20688	16464	12805	9604	6779	4268	2022
0.02	182476	86438	54423	38416	28812	22409	17836	14408	11738	9604	7858	6403	5171	4116	3201	2401	1895	1067	505
0.03	81100	38416	24188	17074	12805	9860	7927	6403	5217	4268	3492	2846	2298	1829	1423	1087	763	474	225
0.04	45619	21609	13606	9604	7203	5602	4459	3602	2935	2401	1964	1601	1293	1029	800	600	424	267	126
0.05	29196	13830	8708	6147	4610	3585	2854	2305	1878	1537	1257	1024	827	659	512	384	271	171	81
0.06	20275	9804	6047	4268	3201	2490	1982	1601	1304	1067	873	711	576	457	356	267	188	119	56
0.07	14886	7056	4443	3136	2362	1829	1456	1176	958	784	641	523	422	336	261	198	138	87	41
0.08	11405	5402	3401	2401	1801	1401	1115	900	734	600	491	400	323	257	200	150	108	67	32
0.09	9011	4268	2688	1897	1423	1107	881	711	580	474	388	316	255	203	158	119	84	53	25
0.10	7299	3467	2177	1537	1162	898	713	576	470	384	314	256	207	165	128	96	68	43	20
0.15	3244	1537	968	683	512	398	317	256	209	171	140	114	92	73	57	43	30	19	9
0.20	1825	864	544	384	288	224	178	144	117	96	78	64	52	41	32	24	17	11	5
0.25	1168	553	348	246	184	143	114	92	75	61	50	41	33	26	20	15	11	7	•
0.30	811	384	242	171	128	100	79	64	52	43	35	28	23	18	14	11	8	5	•
0.35	596	282	178	125	94	73	58	47	38	31	26	21	17	13	10	8	6	•	•
0.40	456	216	136	96	72	56	45	36	29	24	20	16	13	10	8	6	•	•	•
0.50	292	138	87	61	46	36	29	23	19	15	13	10	8	7	5	•	•	•	•

* Sample size less than 5.