Relative validity of an interactive 24-hour recall to assess the intake of energy and nutrients among 15-month old Malawian children

Master's thesis

Taru Saloheimo
Division of Nutrition
Department of Food and Environmental Sciences
University of Helsinki
December 2011

UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of agriculture and forestry	Department of Food and Environmental Sciences, Division of nutrition

Author
Taru Johanna Saloheimo

Title
Relative validity of an interactive 24-hour recall to assess the intake of energy and nutrients among 15-month old Malawian children

Subject
Nutrition Science

Level	Month and year	Number of pages
Master's thesis	December 2011	90 pages + annexes

Abstract
The aim of this study was to assess the relative validity of an interactive 24-hour recall used in a large randomized controlled trial among 15-month-old Malawian children. Relative validity studies should always be done when a dietary method is applied to a new population. The interactive 24-hour recall is a modified version of the traditional 24-hour recall. It was developed in 1995 in order to improve the validity of 24-hour recall in measuring nutrient intake in poor rural areas in Africa. The modifications of the interactive 24-hour recall are designed to make remembering food items and estimating portion sizes easier for the respondent. In the interactive 24-hour recall, the day before the recall interview the respondents are asked to use standard sized bowls and cups when eating, and to mark off each eaten food item on a picture chart containing pictures of local foods. The picture chart is intended to reduce memory lapses, and the use of cups and bowls is intended to facilitate estimating quantities of consumed foods.

In this study, the relative validity of the interactive 24-hour recall was assessed among 44 children aged 15 months. The subjects were healthy children living in rural villages in Mangochi district, southern Malawi. To mimic the trial conditions, an intervention was included in this study. Half of the children were randomly assigned to receive a lipid-based nutrient supplement and the other half were assigned to follow their usual diet. The relative validity of the interactive 24-hour recall was assessed against a more accurate method of dietary assessment, the weighed food record (reference method). Intakes of energy, protein, fat, iron, zinc and vitamin A were measured by the two methods.

Statistical analyses were done to assess the relative validity of the interactive 24-hour recall at both the population and the individual level. Food level data were analyzed to describe memory lapses and inaccuracies in portion size estimates as sources of measurement error. Furthermore, a description of differential measurement error was calculated to see whether such a bias might exist between the two study groups. Correction values developed in an earlier study in Malawi were tested to see whether they can be successfully used for adjusting values obtained by the interactive 24-hour recall.

The relative validity of the interactive 24-hour recall was good at the group level, and moderate in estimating the intake of individuals. Memory lapses and portion size estimates introduced some errors to measurement. The interactive 24-hour recall performed well in assessing the intake of staple foods. Notable differential measurement error between the study groups did not exist. The correction values did not improve the results. Together, the results of the present investigation suggest that the interactive 24-hour recall is a relatively valid method for estimating the average energy and nutrient intakes of rural 15-month old Malawian children at the group level, and it is an acceptable method to be used to estimate selected individual nutrient intakes.

Keywords
Interactive 24-hour recall, weighed food record; relative validity, infants and young children

Where deposited
Division of nutrition

Further information
Supervisors: Elaine Ferguson, Maijaliisa Erkkola

HELSINGIN YLIOPISTO

Tiedekunta/Osasto	Laitos
Maatalous-metsätieteellinen tiedekunta	Elintarvike- ja ympäristötieteiden laitos, ravitsemustieteen osasto

Tekijä

Taru Johanna Saloheimo

Työn nimi

Vuorovaikutteisen 24-tunnin haastattelun luotettavuus energian ja ravintoaineiden saannin mittaamisessa 15 kuukauden ikäisillä malawilaislapsilla

Oppiaine

Ravitsemustiede

Työn laji	Aika	Sivumäärä
Pro gradu -tutkielma	Joulukuu 2011	90 sivua + liitteet

Tiivistelmä

Tutkimuksen tavoitteena oli arvioida suuressa interventiotutkimuksessa käytettävän vuorovaikutteisen 24-tunnin ruoankäyttöhaastattelun suhteellista luotettavuutta (engl. relative validity). Tutkimus tehtiin Malawissa 15-kuukauden ikäisillä lapsilla. Ruoankäytön tutkimusmenetelmien suhteellista luotettavuutta on tarpeen tutkia, kun menetelmiä käytetään uusissa väestöryhmissä. Vuorovaikutteinen 24-tunnin ruoankäyttöhaastattelu on perinteisestä 24-tunnin ruoankäyttöhaastattelusta muokattu versio. Menetelmä kehitettiin vuonna 1995, kun haluttiin parantaa 24-tunnin ruoankäyttöhaastattelun suhteellista luotettavuutta ravintoaineiden mittaamisessa köyhillä maalaisalueilla Afrikassa. Vuorovaikutteiseen 24-tunnin ruoankäyttöhaastatteluun sisällytettyjen muutosten tarkoitus on helpottaa vastaajia muistamaan kaikki syömänsä ruoat ja arvioimaan syötyjen ruokien määrät. Vuorovaikutteisessa 24-tunnin ruoankäyttöhaastattelussa tutkittavia pyydetään haastattelua edeltävänä päivänä täyttämään lomaketta, jossa on kuvia paikallisista ruoista. Tutkittavat merkitsevät syödyt ruoat lomakkeeseen ruksein ja käyttävät ruokailtaessa tutkimuksen antamia kulhoja ja mukeja. Kuvia sisältävän lomakkeen tarkoituksena on vähentää unohdusten määrää ja yhdenmukaisten kulhojen ja mukien tarkoitus on helpottaa annoskokojen arviointia.

Vuorovaikutteisen 24-tunnin ruoankäyttöhaastattelun suhteellista luotettavuutta arvioitiin 44:llä malawilaislapsella. Lapset olivat terveitä ja 15 kuukauden ikäisiä, maalaiskylistä Mangochin alueelta Etelä-Malawista. Interventiotutkimuksen asetelman jäljittelemiseksi osallistujat satunnaistettiin kahteen ryhmään. Toinen ryhmä sai maapähkinä-pohjaista ravintolisää ja toinen ryhmä noudatti tavanomaista ruokavaliota. Vuorovaikutteisen 24-tunnin ruoankäyttöhaastattelun luotettavuutta arvioitiin suhteessa punnittuun ruokakirjanpitoon. Molemmilla menetelmillä mitattiin lasten energian, proteiinin, rasvan, raudan, sinkin ja A-vitamiinin saannit.

Tilastolliset analyysit tehtiin sekä ryhmä- että yksilötasoilla. Ruokatason analyysien avulla selvitettiin ruoankäyttöhaastatteluissa tapahtuneet ruokien unohdukset ja lisäykset. Lisäksi arvioitiin annoskoko-arvioiden tarkkuutta ja mittausvirheen suuruutta tutkimusryhmittäin. Aiemmassa Malawilaistutkimuksessa kehitettyjen korjauskertoimen toimivuutta testattiin tutkimusaineistossa.

Vuorovaikutteisen 24-tunnin ruoankäyttöhaastattelun suhteellinen luotattavuus oli hyvä ryhmätasolla ja kohtalainen yksilötasolla. Virhelähteitä olivat muistivirheet ja epätarkat annoskoko-arviot. Vuorovaikutteisella 24-tunnin ruoankäyttöhaastattelulla saatiin hyvä arvio peruselintarvikkeiden käytöstä. Tutkimusryhmien välillä ei ollut eroa mittausvirheen suuruudessa. Korjauskertoimista ei ollut tutkimuksessa hyötyä. Johtopäätöksenä voidaan todeta, että vuorovaikutteinen 24-tunnin ruoankäytönhaastattelu on suhteellisen luotettava menetelmä mittaamaan 15-kuukauden ikäisten malawilaislasten energian ja ravintoaineiden saantia ryhmätasolla ja kohtuullisen luotettava yksilötasolla.

Avainsanat

Vuorovaikutteinen 24-tunnin ruoankäyttöhaastattelu, punnittu ruokakirjanpito; suhteellinen luotettavuus, alle 2-vuotiaat lapset

Säilytyspaikka

Ravitsemustieteen osasto

Muita tietoja

Ohjaajat: Elaine Ferguson, Maijaliisa Erkkola

Acknowledgements

This Master's thesis was done for the Department of International Medicine, University of Tampere, within an intervention trial carried out in Malawi. I would like to thank the families in Malawi who volunteered in this study.

I thank Per Ashorn for enabling this study. I have learnt a lot about science from him, and I admire his attitude towards doing research. I wish to express my gratitude to my supervisors Elaine Ferguson and Maijaliisa Erkkola. I feel grateful for having received guidance and constructive feedback from such experienced professionals. Jaimie Hemsworth deserves a big thank you for instructing me at the field, and for always helping me whatever the problem was during and after data collection.

I wish to thank John Phuka and the entire research team in Mangochi, Malawi. Working in Malawi was an amazing professional experience for me. Chiza Kumwenda's help with collecting all necessary details for this thesis is much appreciated. The data collection was a success, thanks to data monitor Patrick Ngosi and each and every one of the highly professional nutrition data collectors.

I thank Elina Kokkonen and Liisa Korkalo for their advice with the statistical analyses, and Jere Kasanen for his assistance with the graphics.

I wish to thank my sisters and all my wonderful friends for their understanding and for helping me keep worry-free through the process of writing this thesis. Lastly, I would like thank my mother and father for always supporting me throughout my studies.

Taru Saloheimo
Helsinki, December 2011

Contents

1 Abbreviations and definitions....3
2 Introduction....4
3 Theoretical framework....7
3.1 Measuring food consumption and nutrient intakes in infants and young children....7
3.1.1 Diet during the first two years of life....7
3.1.2 Challenges in measurement....9
3.1.3 Surrogate reporters....10
3.2 Measurement errors specific to dietary recalls....11
3.2.1 Random error vs. systematic error and their sources....11
3.2.2 Differential measurement error....14
3.2.3 Selected issues in dietary assessment in developing countries....15
3.3 The interactive 24-hour recall method....17
3.3.1 Reasons for developing an interactive 24-hour recall....17
3.3.2 Description of the method....18
3.3.3 Strength and limitations of interactive 24-hour recall vs. other methods....20
3.3.4 Assessing the relative validity of dietary recalls....21
3.3.5 Statistical assessment of relative validity....24
3.3.6 Previous relative validity studies in young children in developing countries....29
4 Aim of the study....34
5 Subjects and methods....35
5.1 Study design....35
5.1.1 The main study....35
5.1.2 The present study....35
5.2 Dietary methods....43
5.2.1 Weighed food record....43
5.2.2 Interactive 24-hour recall....44
5.3 Data preparation and statistical analysis....46
5.3.1 Data preparation....46
5.3.2 Statistical analysis....48
6 Results....51
6.1 Characteristics....51
6.1.1 Subjects....51
6.1.2 Surrogate respondents....53
6.2 Comparison of two interactive 24-hour recalls....54
6.3 The relative validity of interactive 24-hour recall....55
6.3.1 Extent of agreement between methods on a group basis....55

6.3.1.1 Wilcoxon matched-pair signed-rank test 55
6.3.1.2 Application of correction values developed by Thakwalakwa et al. 56
6.3.2 Extent of agreement between methods on an individual basis 57
6.3.2.1 Spearman rank correlation coefficient and partial correlation 57
6.3.2.2 Cross-classification into tertiles 58
6.3.2.3 Bland-Altman analysis 59
6.3.3 Sources of measurement error 66
6.3.3.1 Intrusion and omission of food items, and portion size estimates 66
6.3.3.2 Description of differential measurement error 69
7 Discussion of results 71
7.1 Aims of the study and main findings 71
7.2 Sample vs. target population 71
7.2.1 Internal validity 71
7.2.2 External validity 72
7.3 Relative validity of interactive 24-hour recall in this study 73
7.3.1 Results of the present study 73
7.3.2 Comparison with previous studies and recommendations 74
7.4 Secondary objectives of the study 75
7.4.1 Memory lapses and inaccuracies in portion size estimates 75
7.4.2 Differential measurement error 76
7.4.3 Correction values developed by Thakwalakwa et al. 77
7.5 Strengths and limitations 78
8 Conclusions 80
9 References 82
10 Annexes 91

1 Abbreviations and definitions

HSA	Health surveillance assistant
Infant	Child less than 12 months of age
i-24HR	Interactive 24-hour recall
i-24HR-I	Baseline interactive 24-hour recall
i-24HR-II	An interactive 24-hour recall concerning the same day of food intake as WFR
LNS	Lipid-based nutrient supplement
RA	Research assistant
Young child	Child between 12 and 23 months of age
WFR	Weighed food record
24HR	24-hour recall
12-i-24HR	An interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours

2 Introduction

Recent estimates suggest that different forms of child undernutrition - stunting, wasting, intrauterine growth restriction and micronutrient deficiencies - are responsible for more than a third of deaths in children younger than five years (Black et al. 2008). In addition to survival in childhood, undernutrition has long-term consequences. They include shorter adult height, lower attained schooling, reduced adult income and decreased offspring birth weight (Victora et al. 2008). Promoting child health and investing in nutrition are considered as essential for achieving the key development goals of the international community (UN 2000).

The period from pregnancy to two years of age is a crucial window of opportunity for reducing undernutrition and related disease burden (Bryce et al. 2008). Action at the national level should focus on this age segment (Bryce et al. 2008). Promoting breastfeeding and improving the quality of additional foods given to children are listed among interventions known to be effective in reducing child undernutrition (Bhutta et al. 2008). However, to truly eliminate child undernutrition, long-term investments to improve education, economic status, and empowerment of women are required (Bhutta et al. 2008).

Children younger than two years should receive breast milk, and starting from the age of six months, a variety of additional foods along with continuing breastfeeding (WHO 2003). These complementary foods are often given too early (Lauer et al. 2004), and their quality is poor (Gibson et al. 2010). Compliance with breastfeeding recommendations is also low (Lauer et al. 2004). Food consumption surveys are indispensable sources of nutrient intake data. Although laboratory, anthropometric and clinical studies can also be used to assess some aspects of undernutrition, the first stage of any nutritional deficiency can only be identified by dietary assessment methods (Gibson 2005a). Information on food consumption is also needed for developing food-based dietary guidelines and studying the relationship between diet and chronic disease.

To measure food intake, qualitative, semi-quantitative or quantitative methods could be used (Thompson & Byers 1994). Information is needed on absolute food intake when assessing total energy intake and the intake of nutrients present in a variety of sources. For example, zinc and iron are widespread in so many foods that measuring their intake is not possible using simplified questionnaires. In many settings, the method of choice for collecting quantitative data is the 24-hour recall (24HR). In the 24HR, the respondent is asked to list and quantify all foods eaten during the previous day. The 24HR has many advantages: it is inexpensive, easy to implement and has low respondent burden. However, the 24HR is more feasible among some people than among others, the preceding often being educated, lean adults (Klesges, Eck & Ray 1995) and the latter being young children (Fisher et al. 2008, Thakwalakwa et al. 2011), the elderly (Madden, Goodman & Guthrie 1976), or people from poor rural areas (Alemayehu, Abebe & Gibson 2011), for example.

Memory faults and inaccurate estimates of portion sizes of foods are the main reasons why 24HR has not performed very well in assessing the food intake of rural African respondents (Ferguson et al. 1989; Dop et al. 1994; Ferguson, Gibson & Opare-Obisaw 1994; Ferguson et al. 1995; Alemayehu, Abebe & Gibson 2011). Perhaps the concept of the 24HR is just not evident for a rural African respondent in the way that it is for a researcher who thinks of food as a source of energy and nutrients. In different cultures people attach very different meanings to food (Cassidy 1994). To improve the feasibility of the 24HR in rural settings, a modification of the 24HR - termed an interactive 24-hour recall (i-24HR) - was developed in 1995 (Ferguson et al. 1995). The modifications are designed to make remembering food items and estimating portion sizes easier for the respondent. So far, there is only one study about the performance of the i-24HR among children aged less than two years (Thakwalakwa et al. 2011).

One way to determine whether a method like the i-24HR really works is to compare it to another source of food consumption information. In nutritional epidemiology, these method comparison studies are called relative validity studies (Nelson 1997). Assessing the absolute validity of a dietary method for all nutrients is not possible because a true external reference measure would be needed. At present, such a measure does not exist in nutrition (Nelson 1997). A relative validity study will tell whether the method being studied understates, exaggerates or provides similar estimates of food intake as the comparison method. For the i-24HR, the weighed food record (WFR) is the most appropriate comparison method. The relative validity of a method is very much affected by the characteristics of the population that it is applied to, and the nutrient of interest (Ferguson, Gibson & Opare-Obisaw 1994). The relative validity of a method can be different even in two areas very close to each other (Ferguson, Gibson & Opare-Obisaw 1994). Hence, it is very important to reassess the relative validity of any method before it is used at large scale, for example in a large epidemiologic study. Valid estimates of nutrient intakes are essential for carrying out good-quality research.

The aim of the present study was to assess the relative validity of the i-24HR among 15-month old rural Malawian children, a group in which chronic malnutrition is highly prevalent (NSO Malawi & ORC Macro 2005). Malawi is a landlocked country in sub-Saharan Africa with a population of 13.1 million people (NSO Malawi & ORC Macro 2005). Life expectancy at birth is 52 and 54 for males and females, respectively, and mortality of under-five children is 10.0 % (WHO 2010). Approximately a fourth (23.6 %) of under-five children are underweight and half (48.3 %) are stunted (NSO Malawi & ORC Macro 2005).

This work was done for a large clinical trial in southern Malawi that investigates lipid-based nutrient supplements (LNS) in the prevention of linear growth failure in infants and young children (ClinicalTrials.gov identifier: NCT00945698). It is known that LNSs are effective in treating severely undernourished children (Collins & Sadler 2002, Diop et al. 2003, Manary et al. 2004, Sandige et al. 2004) but less is known about their efficacy in preventing undernutrition. Earlier results from Malawi showed that a 50 g daily dose of LNS reduced the incidence of severe stunting among 6-18 month old infants (Phuka et al. 2008).

The trial uses the i-24HR to assess the dietary intake of respondents aged 15 months. The participants of the present study were not participating in the clinical trial, but were recruited from the same region. To simulate the clinical trial, the participants of this study were randomized into two groups. One group received LNS and the other group followed their habitual diet. The relative validity of the i-24HR was determined by comparing it to WFRs. It was then assessed whether the relative validity of the i-24HR differs between the two subgroups, i.e. whether the response to dietary assessment methods differ according to group assignment (Kohlmeier & Bellach 1995). The intervention might have introduced a differential measurement bias in the i-24HR related to errors in estimating the quantity of LNS consumed, or to an influence of an intervention on respondent responses. Furthermore, sources of measurement error at food level were explored to support the aggregate nutrient analysis using a recently introduced framework (Smith et al. 2007). This study provides an estimate about the relative validity of the i-24HR in assessing the intake of energy, protein, fat, iron, zinc and vitamin A, in one particular age group and in one geographic location in Africa.

3 Theoretical framework

3.1 *Measuring food consumption and nutrient intakes in infants and young children*

3.1.1 Diet during the first two years of life

The nutritional needs of children change remarkably during the first two years of life (Brown, Dewey & Allen 1998). The World Health Organization (2003) recommends practicing exclusive breastfeeding from birth to six months of age. The term exclusive breastfeeding is used when all energy and nutrients are provided by breast-milk, with the exception of small amounts of vitamins, minerals or medicines (WHO 2008). After six months of age, other foods and liquids are needed to meet the nutrient needs of infants (WHO 2003). The process of introducing new foods to infants is defined as complementary feeding (PAHO 2003). However, continuing breastfeeding alongside giving other foods is beneficial for young children until two years of age (WHO 2003).

The total energy requirements of children aged 6 to 8 months, 9 to 11 months and 12 to 23 months are approximately 2.57 MJ, 2.87 MJ and 3.74 MJ per day, respectively (Dewey & Brown 2003). The energy needs from complementary foods are 0.85 MJ per day at 6 to 8 months of age, 1.28 MJ per day at 9 to 11 months of age, and 2.29 MJ per day at 12 to 23 months of age (Dewey & Brown 2003). Figure 1 demonstrates how the proportion of energy intake from complementary foods increases gradually.

Given the relatively small amount of complementary foods consumed before the age of two years, complementary foods need to be rich in nutrients (PAHO 2003). Animal source foods should be eaten daily and fortified complementary foods and supplements, as needed, used to complement the diet (PAHO 2003). Yet, in many areas cereals and other plant-based foods prepared as thin gruels provide the basis for young children’s complementary foods (Gibson, Ferguson & Lehrfeld 1998). Previous research suggests that by 12 months, most children are physically able to consume the same foods as the rest of the family in substantial amounts (Kersting et al. 1998, USDA 2000). Hence, special transitional foods with liquid or semi-solid consistency may only be needed between 6 and 11 months of age.

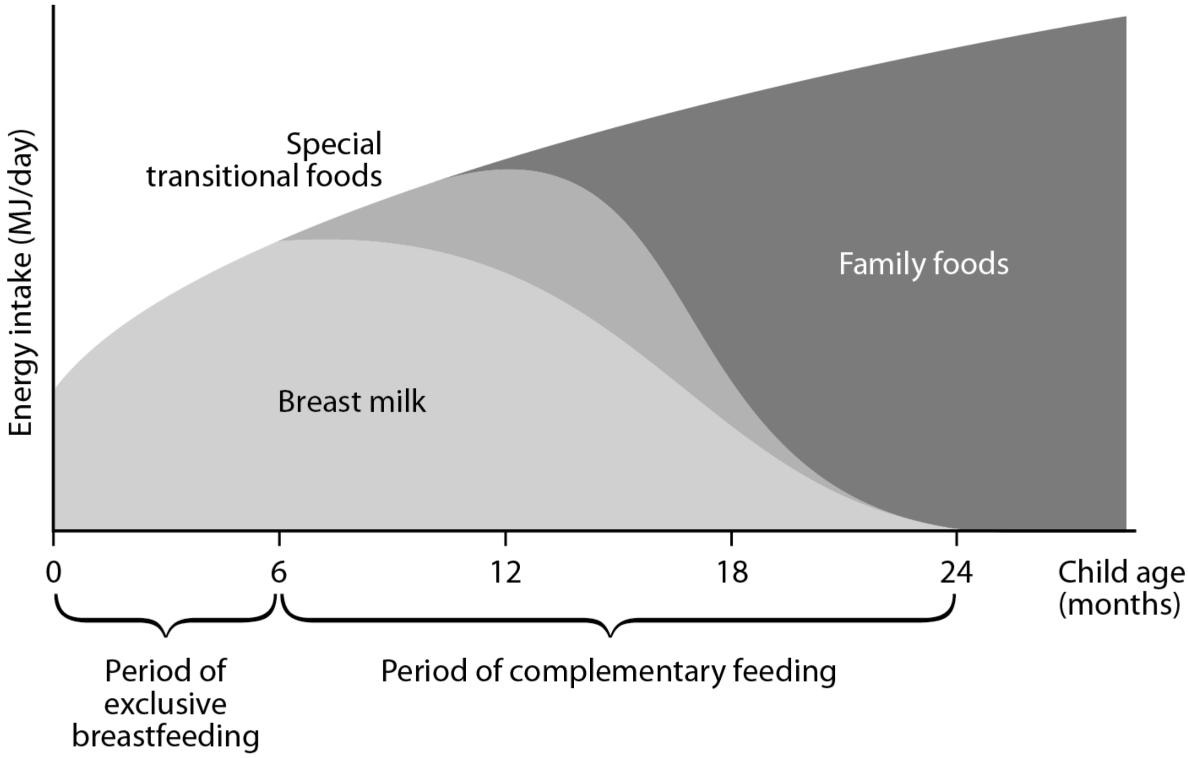

Figure 1. Possible contribution of different food sources to young children's energy intake in relation to age. Adapted from Brown, Dewey & Allen 1998.

Household food availability and access to food are important factors associated with infant feeding practices (Saha et al. 2008). In addition to what is available, the delivery of appropriate complementary foods is determined by components of care-giving; how food is fed, when food is fed, who is giving the food, and where food is fed (Pelto, Levitt & Thairu 2003). Time constraints, knowledge and beliefs, and social pressure may constrain caregivers in their ability to provide care (Pelto, Levitt & Thairu 2003). Maternal education is associated with better complementary feeding practices (Guldan et al. 1993). It should be noted that parents are not the only ones who decide how their child is fed (Aubel, Touré & Diagne 2004, Bezner Kerr et al. 2008). For instance, a study conducted in northern Malawi revealed that grandmothers commonly gave herbal tea - which is not a recommended complementary food - to their grandchildren aged less than one month, and it was done to protect the child from illness believed to be caused by the "promiscuity" of the mother or father (Bezner Kerr 2006).

Hence, although breastfeeding and complementary feeding recommendations are universal, local practices vary. Some regions and countries have higher breastfeeding rates than others. Lauer et al. (2004) reviewed infant and young child feeding data from cross-sectional national surveys performed between 1991 and 2001 in 135 developing countries. In most of the surveys breastfeeding estimates were based on the 24HR. Two fifths (39 %) of infants aged six months or younger were exclusively breastfed. The prevalence of continued breastfeeding was 86 % and 68 % for children aged 6 to 11 and 12 to 23 months, respectively. Feeding patterns of Malawian children appear to be slightly closer to the recommendations compared with the average estimates presented by Lauer and colleagues. According to a recent national survey in Malawi, more than half (53 %) of infants younger than six months were exclusively breastfed, almost all (99 %) infants were breastfed for at least a year, and 80 % were breastfeeding toward their second birthday (NSO Malawi and ORC Macro 2005). The Malawian survey is also based on data obtained by the 24HR. The quality of complementary foods in many developing countries is poor because adequate intakes of several nutrients are difficult to achieve from traditional, plant-based complementary foods (Gibson et al. 2010). In southern rural Malawi, complementary diet of young children has been described as inadequate and especially lacking in iron, zinc and calcium (Hotz & Gibson 2001).

3.1.2 Challenges in measurement

Population-based indicators of child feeding practices are useful tools for nutrition programs and interventions. While the first indicators to measure breastfeeding practices were launched in 1991 (WHO 1991), progress in developing indicators to evaluate complementary feeding practices has taken place more recently. In 2003, Ruel et al. urged developing simple, valid and reliable tools to assess the adequacy of complementary feeding practices. Thereafter, the World Health Organization has published a set of eight core indicators and seven optional indicators that cover the entire range of child feeding practices, including appropriate complementary feeding (WHO 2008). The core indicators include early initiation of breastfeeding, exclusive breastfeeding under six months of age, continued breastfeeding at one year, introduction of solid, semi-solid or soft foods, minimum dietary diversity, minimum meal frequency, minimum acceptable diet, and consumption of iron-rich or iron-fortified foods.

Unfortunately, such indicators are not sufficient for research studies aiming to measure the actual food consumption of children. Accurate assessment of the food intake of children is challenging, especially when children are receiving both breast-milk and complementary foods (Piwoz et al. 1995). Moreover, most of the methodological data on children’s food consumption deals with children older than two years of age (Livingstone, Robson & Wallace 2004, Forrestal 2011). The conventional method for assessing breast-milk intake entails weighing the child before and after each feeding (Arthur, Hartmann & Smith 1987). However, the method is inaccurate and difficult to apply in field conditions (Savenije & Brand 2006). The introduction of the stable isotope technique was an important advance in the measurement of breast milk intake. The technique was pioneered by Coward and co-workers in 1982. Briefly, the amount of breast-milk consumed by the child is assessed by giving the mother a drink of deuterium labelled water, and following the disappearance of the deuterium from the mother and its appearance in the child (IAEA 2010). The breast-milk intake is proportional to the rate of deuterium appearance in the child, measured by urine or saliva collections. Although the deuterium oxide technique represents a substantial improvement over test-weighing, it is not yet widely used. By the end of 2007, approximately 20 studies had conducted deuterium oxide measurements of breastfeeding (da Costa et al. 2010).

When it comes to quantitative assessment of complementary food consumption of infants and young children, several issues may complicate obtaining accurate data. Spilling food, which may occur when learning to self-feed, can impede the accuracy of portion size estimation regardless of the dietary method used.The transition to self-feeding takes place before the age of 2 years, but there is variability in ages at which individual children show selected self-feeding skills (Carruth et al. 2002; 2004). Furthermore, the pace of overall motor development may be culture-specific (Papalia, Olds & Feldman 2007). In Uganda, for example, children commonly walk at 10 months, as compared with 12 months in the United States (Gardiner & Kosmitzki 2005). Regarding dietary methods that rely on the respondent’s memory, remembering quantities eaten by the child accurately may be difficult because the amount of food consumed by young children is not large. Children less than two years may also be frequently fed compared with older children making it possibly challenging to remember each separate feeding session.

3.1.3 Surrogate reporters

Surrogate reporters are needed when complementary food consumption of young children is measured via the 24HR. Mothers are usually the primary respondents in studies done in Africa, although sometimes additional information is probed from other persons to get all necessary details about the child’s diet (e.g. Ferguson et al. 1989, Dop et al. 1994, Gibson &

Opare-Obisaw 1994, Olinto et al. 1995, Gewa, Murphy & Neumann 2009, Thakwalakwa et al. 2011). The time mothers spend away from their children is a major factor that may reduce the accuracy of the 24HR. Gewa and collaborators (2007) found that Kenyan mothers were not aware of foods that their school-aged children ate outside the home. A question arises whether young African children are habitually fed by others than the primary caregiver. It is known that a large kin group is important for a traditional African family, and that the extended family is dedicated to raising the children besides the nuclear family (Kayongo-Male & Onyango 1984).

The importance of multiple caregivers in child feeding has been documented in a group of foragers in Congo (Fouts & Brookshire 2009). Twenty-two children between two and four years of age were observed to determine who feeds them. Mothers were the single highest contributor to child feeding, but combined contributions from other individuals - such as fathers, grandmothers, aunts and siblings - was higher than that of mothers. Having younger siblings, especially having a new infant predicted the feeding of two to four year old children by others than the mother. Although a group of foragers is not totally comparable with rural communities nor are children aged two to four years fully comparable to younger children, findings from Congo are valuable information when better data are lacking. A study conducted in a high-income country showed that a 24HR of a child's food intake is likely to be more complete if information is gathered from more than one respondent (Eck, Klesges & Hanson 1987). Together these data suggest that surrogate reports of a child's diet may lack accuracy simply because surrogate respondents may not observe all occasions when the child is eating. Additionally, surrogate respondents may be susceptible to over- or underreporting the dietary intake of the index subject just as when recalling of their own food intake.

3.2 *Measurement errors specific to dietary recalls*

3.2.1 Random error vs. systematic error and their sources

Measurement errors can be introduced at any point in a dietary study. Broadly speaking, there are two types of measurement errors - random and systematic. When assessing the dietary intake of a group of individuals with a single measurement, the intake of some individuals may be overestimated, while the intake of others may be underestimated. If the mean for the group is nonetheless correct, the error in measurement is random by nature. This type of error is termed random between-person error; individuals within a group are affected by errors, but randomly. If the individuals within the group are affected by a systematic error that biases the measurement error in the same direction, the average would be significantly different from the true mean; the mean would be either lower or higher than the true mean. A systematic between-person error would be said to exist.

The best way to control for both random and systematic errors is to include quality-control procedures into the study assuring that all procedures are done in a standardized way. Increasing the number of observations in a survey can compensate for the effect of random errors. The spread in the measurements, as shown by the standard deviation, estimates the likely size of random error in a single measurement. Also, expressions of statistical significance are sensitive to the degree of random errors. It is harder to show statistical significance if the measurement is strongly affected by random errors. Quantifying systematic error is done using an external reference method. There is literature on how to correct for the effect of systematic errors after a study has been completed, but for some reason, these strategies are rarely carried out (Kohlmeier & Bellach 1995, Kipnis et al. 2002).

Interviewer biases

When performing dietary recalls, i.e. the 24HR or the i-24HR, interviewers can introduce errors to measurement if they use probing questions incorrectly, record responses erroneously, omit food items intentionally, or there is no rapport between the interviewer and the respondent, among other things (Fowler & Mangione 1990). Recruiting and training local people to serve as interviewers is preferable so that communication will be more effective. Using interpreters is another possible source of error in measurement (Kigutha 1997). The design of a dietary study should allow assessing any potential interviewer bias so that statistical methods can be applied to correct for this type of measurement error (Slimani et al. 2000).

Respondent memory lapses and incorrect estimation of portion size

Individuals participating in the study may consciously or unconsciously deny or exaggerate their food intake (Willett 1998). Respondents may have memory lapses resulting in omission or addition of food items. Memory lapses relating to consumption of beverages (Alemayehu, Abebe & Gibson 2011), snack foods (Ferguson, Gibson & Opare-Obisaw 1994, Ferguson et al. 1995), as well as fish and millet-sorghun (Dop et al. 1994) have been documented in developing country settings. In most rural communities, meal frequency and the time of eating are not rigidly fixed (Kigutha 1997). Furthermore, meal patterns vary according to the area and season.

Portion sizes may be estimated incorrectly when respondents fail to quantify accurately the amount of food consumed (Alemayehu, Abebe & Gibson 2011). Most foods are prepared at home in rural households and industrial, packaged products are not consumed often. Mistakes in estimating amounts of rice (Dop et al. 1994), porridge (Ferguson et al. 1989), nsima (a maize-based staple food) and legume relish (Ferguson et al. 1995), and banku (staple prepared from corn and cassava) (Ferguson, Gibson & Opare-Obisaw 1994) have been problematic in some studies.

Respondent biases may also arise if the respondent misunderstands the purpose of the study. In resource-poor settings, it is very important to clearly state the objective of the study. Based on experiences in Kenya, Kigutha (1997) believes that biases may be introduced if participants think that they will get food or financial aid if they report low food intakes, and that overreporting of food intake may occur if subjects want to impress their interviewers.

Underreporting or overreporting food intake is common in dietary surveys. Underreporting is the most studied form of reporting bias and has been observed in many nutrition surveys in North America and Europe (Beaton, Burema & Ritenbaugh 1997). Interestingly, some data implies that underreporting of food intake by the 24HR is not universal, at least among adult Egyptian and Indonesian women (Harrison et al. 2000, Winkvist, Persson & Hartini 2002). However, few studies concerning underreporting have been done in developing countries (Scagliusi, Ferriolli & Lancha 2006), contrary to the large number of studies done in affluent ones.

Coding errors and errors in handling of mixed dishes

Recording data from dietary recalls is subject to coding errors. Coding errors can occur when amounts of foods are converted to grams and when assigning codes for food items (Gibson 2005d). Handling of mixed dishes is a further source of error. Errors may appear when the mixed dish is broken down into raw ingredients and when taking into consideration changes in weight and nutrient retention due to cooking (Gibson 2005d). Availability of accurate recipe data for calculating the quantities of raw ingredients consumed from mixed dishes is important. However, there is very little literature on nutritional composition of mixed dishes used in African countries (Sharma et al. 2007). Average recipes of mixed dishes are often calculated for small geographical units such as provinces because areas within one country can be different from each other in terms of food habits (Sharma et al. 2007). On the other hand, rural respondents often have first-hand knowledge of exact ingredients and recipes (Solomons and Valdés-Ramos 2002).

3.2.2 Differential measurement error

Systematic errors in measurement that are related to the outcome of the study are the most serious problem in nutritional epidemiology (Willet 1998). As an example, Bellach and Kohlmeier (1998) found that cases underreported fat and energy intake but controls did not. This type of error is called differential measurement error. It occurs when respondents react differently to a dietary assessment method used within a study (Kohlmeier & Bellach 1995). Literature on differential measurement error usually deals with case-control studies (Kohlmeier & Bellach 1995, Thürigen et al. 2000, White 2003) but could also be relevant for cohort studies, for example, if those with symptoms of disease change their habitual diet before future diagnosis (White 2003). The concept would seem applicable to randomized clinical trials, too, if the group allocation is thought of as the outcome (Natarajan et al. 2010). Some studies suggest that knowledge about an intervention may boost social desirability responses in an intervention group more than in a control group (Miller et al. 2008, Natarajan et al. 2010), but some studies have found no such effect (Harrington et al. 2009).

Differential measurement error is a major source of bias in epidemiological studies and results in an over- or underestimation of risk (White 2003). One measure of differential measurement error - termed factor A - is the ratio of the observed mean difference in dietary intake between cases and controls to the true mean difference in dietary intake, or:

$$Factor\ A = 1 + \frac{a - b}{c - d}$$

where

a = observed mean intake minus true intake for cases

b = observed mean intake minus true intake for controls

c = true mean intake for cases

d = true mean intake for controls

Observed intake denotes intake measured by the test method and true intake denotes intake measured by the reference method (White 2003). If factor A is positive, it gives the proportion of over- or underestimation. A value of 1.6 would mean that the dietary intake measured with error overestimates the true case-control mean difference by 60 percent. If A is 0.2, then only 20 percent of the estimated true difference between cases and controls in intake is observed by the exposure measured with error. A negative value means that the mean difference in intake between cases and controls has changed signs; that is, association between diet and disease is in the opposite direction to the true association. Differential measurement error should always be interpreted in terms of the magnitude of its effect on the measure of association in question. Assessing the effect of differential error on the risk of disease can be easily quantified when certain simplifying assumptions are made (White 2003).

3.2.3 Selected issues in dietary assessment in developing countries

Although there are practically always some errors in dietary data, measurement errors are not an inherent property of any method. The error can vary for a single method when applied to different population groups which vary by, for example, level of education. Even within one's own society, sex, social class, language, ethnicity, age, income, education and personal history affect how each individual perceives reality, what meanings are attached to words and acts, and how concepts such as good diet are understood (Cassidy 1994). In cross-cultural research, the world view of the western researcher may not match that of the people that the researcher wants to understand. To do good science, researchers need to be aware of their own assumptions and value systems.

Traditional cultural practices reflect values and beliefs held by members of a community for periods often over many generations. All cultures have rules and beliefs about foods (Simoons 1994) and they seem to be plenty in Africa (Simoons 1994, Kruger & Gericke 2003, Onuorah & Ayo 2003). A qualitative exploration in rural South Africa showed that mothers did not consider food intake behavior as being important in assessing the health status of their children (Kruger & Gericke 2002). The type of foods, variety, nutritional value or quantity of foods given to the children was not given any consideration. Cultural factors had a powerful influence on feeding practices.

The meanings of words are another culture-specific issue. The term food is a very general category which does not have an equivalent in all languages. In some settings people distinguish rice from “food” (Wolff 1965). When asked about yesterday’s food intake people will report “food” intake, that is, everything else but rice. Rice is valued so much that it is almost considered more than food. In these cases asking about rice consumption in addition to food consumption is crucially important in order to get a complete picture of the diet. In addition, some completely different foods may be called the same name or similar foods may be called by different names (Loria et al. 1991).

Seasonality is a major factor determining food availability in tropical areas where agriculture is highly dependent on temperature and rainfall variation. Scientists must be familiar with this issue and understand how seasonal variations affect food intake (Ferguson et al. 1993, Mitchikpe et al. 2009). Seasonal variability in nutrient intake may be greater in subsistence communities relative to western populations and to more urbanized developing country populations, which underscores the importance of considering seasonal variations during study design (Nyambose, Koski & Tucker 2002). Researchers should also be aware of peoples’ attitude towards questioning. In some societies it is frightening to have a person probing for information because only government authorities question people (Cassidy 1994). In some cultures mixed-ethnicity social situations provoke shyness (Paulhus, Duncan & Yik 2002).

Research has shown that misreporting reflect users' participation in social patterns that award differential status to certain foods. In a Norwegian study done among children aged 12 and 24 months, parents overreported the consumption of foods considered healthy like bread, fruit and potatoes, and underreported the consumption of foods considered less healthy such as cakes and sweets (Andersen et al. 2003; 2004). Underreporting of foods rich in fat and carbohydrates is common in western contexts (Heitmann & Lissner 1995, Lafay et al. 2000). On the basis of expectations of appropriate eating behavior, admitting to eating certain foods might be difficult for many people holding different kinds of beliefs or convictions.

Culturally sensitive research recognizes differing values and works to establish good communication between the researcher and the respondent (Cassidy 1994). A culturally sensitive method has some specific characteristics. It has few words on it, the respondent is asked for information that is comprehensible and logical to him or her, and the method is flexible and open allowing for making modifications (Cassidy 1994). The 24HR is an example of a culturally sensitive dietary method. Since the 24HR requires recall of everything eaten, there is no need to categorize foods into groups, or to guess beforehand which foods respondents are likely to eat. Explaining the basic idea of the 24HR is easy, and quantities of foods can be asked when the respondents understand what is wanted of them. Even with the 24HR some challenges remain: how to estimate individual intake from shared serving dishes, and how to estimate quantities from nonstandard eating and serving tools like hands (Harrison 2004).

3.3 *The interactive 24-hour recall method*

3.3.1 Reasons for developing an interactive 24-hour recall

The 24HR is used in epidemiologic investigations when estimates of absolute energy and macronutrient intakes are required (Buzzard 1998). The 24HR is more accurate in some populations than in others (Gibson 2005b). In Africa, the WFR has been the method of choice for collecting quantitative dietary data because of difficulties intrinsic in collecting 24HR data. The accuracy of 24HRs is affected by the fact that making quantitative estimates of foods has proven difficult to African respondents who eat their food from one plate of food shared with other family members (Rutishauer 1973). Additionally, snack food consumption by children has been poorly captured when using the 24HR (Ferguson et al. 1989).

Some studies have suggested, however, that the 24HR can be used for estimating dietary intakes on a group basis (Rutishauer 1973, Ferguson et al. 1989, Ferguson, Gibson & Opare-Obisaw 1994). Rutishauer (1973) employed a picture chart of local foods and samples of foods actually consumed to reduce memory lapses and improve estimates of portion sizes consumed. Against this backdrop, Ferguson and co-workers (1995) developed a new, modified version of the 24HR – termed an i-24HR - and pilot-tested it among 60 pregnant rural Malawian women. A detailed description of the i-24HR follows below.

3.3.2 Description of the method

The i-24HR is a modification of the 24HR developed for assessing dietary intakes of rural populations in developing countries (Ferguson et al. 1995). The modifications reduce the number of memory lapses, facilitate estimating quantities of consumed foods, and improve the visual impression of foods eaten. The main difference between the i-24HR and the traditional 24HR is that in the i-24HR the respondents prospectively record the food items that they or their children consume during the day of interest. Food items are recorded by marking off a chart that contains pictures of local foods. Thus, the respondent does not have to be literate.

The use of the i-24HR is discussed in detail in a manual prepared by Gibson and Ferguson (2008). Essentially, the i-24HR is comprised of three distinct parts done on three consecutive days. First, respondents are prepared for the i-24HR (day 1). Preparing respondents should take place two days before the recall interview in a group or in an individual training session. During the session, the purpose of the i-24HR should be explained, estimating portion sizes rehearsed and instructions on how to complete picture charts should be given. Cups, bowls and picture charts should be provided for use on the following day.

Second (day 2), respondents take part in collecting the dietary data the day before the recall interview - this can be considered as the interactive part. The respondents (or their child) are asked to eat from plates and bowls given to them. Eating from a separate plate is thought to help the respondent visualize foods eaten because eating foods from shared plates is a common practice in many African areas. Respondents are also asked to mark off all eaten food items on a picture of local foods. The picture chart should depict commonly consumed local foods of the season. It can be prepared from drawings or photographs. The picture chart will be used for comparison with the recall interview to reduce memory lapses.

In the third part of the i-24HR (day 3), a recall interview is made about the foods and drinks consumed on the previous day. The recall interview is carried out at the home of the respondent and performed by a trained research assistant (RA) who helps the respondent recall foods and drinks consumed by asking neutral questions. The recall interview of the i-24HR is a 4-pass approach. In the first pass, respondents are asked to list all foods they (or their child) consumed on the previous day. In the second pass, information is probed for the food preparation methods and other details about the food. In the third pass, the respondents are asked to estimate the amount of each food consumed. In the fourth pass, after the RA has written down all foods and the amounts consumed listed by the respondent on a form, the RA asks the respondent to pass the picture chart he or she had filled on the previous day. The RA compares the information on the form to the information on the picture chart and discusses possible discrepancies. If a food item appears on the picture chart but is not among the foods just listed, the RA and the respondent will discuss whether this food item was actually consumed or not. Corrections are made to the form accordingly. If a food item on the form has not been marked off on the picture chart, the same procedure will follow. The picture chart is intended to reduce the number of additions and omissions of foods.

Gibson & Ferguson (2008) also discuss alternatives for estimating portion sizes of foods consumed. Salted food models, actual foods, water, household utensils, play dough and tape measures can be used, among other things. However, using actual foods or salted food models are the preferred option because the amount eaten is easier for the respondent to visualize. Estimated portion sizes are converted into weight equivalents depending on how the quantities were estimated. For instance, actual foods and salted food models are weighed directly while water equivalents are multiplied by the density for the food or beverage consumed.

In the pilot study done among Malawian women, the i-24HR was repeated twice and one of the two recalls was compared with WFRs done on the same day of food intake and the two recalls were compared (Ferguson et al. 1995). Picture charts were used and portion sizes estimated by using salted replicas for main meal food items, and tape-measures and monetary equivalents for other foods. Data analysis showed that the median intakes of calcium, iron, zinc and manganese obtained by the i-24HR and the WFR were similar (Wilcoxon matched-pair signed-rank test, $p > 0.05$). The i-24HR slightly underestimated median intakes of energy, protein, fat and copper compared to the WFR (Wilcoxon matched-pair signed-rank test, $p \leq 0.05$). The authors concluded that underestimation of one portion of cereal staple was the main source of error. Due to this finding, the authors suggested that respondents should be trained to estimate portion sizes before conducting the recall interviews. This recommendation is included in the manual on i-24HR written by Gibson & Ferguson (2008). After the pilot study, the i-24HR has been used in several settings (e.g. Gibson & Huddle 1998, Hotz & Gibson 2001, Aziz & Hussein 2005, Alemaheyehu, Abebe & Gibson 2011, Thakwalakwa et al. 2011) and is being used by a clinical trial in Malawi (ClinicalTrials.gov identifier: NCT00945698).

3.3.3 Strength and limitations of interactive 24-hour recall vs. other methods

Two groups of dietary methods are available for assessing the food intake of individuals (Gibson 2005c). Recalls and records are used for measuring the quantity of foods consumed over a one-day period. Repeat recalls or records can be done to account for day-to-day variability in food intakes. Food frequency questionnaires and diet histories are used for describing food consumption patterns over longer, less precisely defined periods of time. The choice of method depends on the objective of the study. The benefits of the i-24HR compared with other quantitative dietary methods are discussed below.

The i-24HR is a modified version of the traditional 24HR. In the 24HR, respondents are asked to recall the exact food intake of the previous day by a trained interviewer. The accuracy of the 24HR depends on the respondents' ability to remember and quantify consumed foods. The major benefit of the i-24HR compared with the 24HR is that it is not fully dependent on memory. In the i-24HR, the respondents are asked to participate in collecting the data by recording their food intake prospectively. Participants mark off food items on a picture chart when the foods are actually consumed. The picture chart is likely to reduce memory lapses, especially when eating is frequent and consuming snacks is common.

Another difference between the i-24HR and the 24HR is the use of standard sized plates and bowls. Respondents are asked to eat from plates and bowls provided to them so that they would have a better visual impression of foods eaten. If foods are eaten from a separate plate instead of a shared plate, it may be easier to estimate the amount consumed. On the other hand, there is no surprise effect in the i-24HR unlike in the 24HR, if the respondent has not been notified about the timing of the upcoming interview. Hence, it cannot be ruled out that

that the respondents of the i-24HR alter their true eating patterns. Conducting the i-24HR also requires more resources compared with the 24HR as there are preparations before the recall interview. The validity of either 24HR or i-24HR has not been extensively tested among rural African populations. If the 24HR would yield as accurate results as the i-24HR, the former should be used in order to save resources.

Dietary records consist of estimated records and WFRs (Gibson 2005c). Respondent's ability to write is a perquisite for the estimated or weighed food record when respondents record the data themselves. Thus, they are not feasible among illiterate populations. For the WFR, the respondent does not have to have any special skills if foods are weighed by a RA. The i-24HR does not require special abilities either. Both the WFR and the i-24HR are completely open-ended allowing for prospectively recording all consumed foods in detail. However, the WFR is more accurate because portion sizes are measured directly while the i-24HR relies on the visual impression that the respondent has about each consumed food item.

On the other hand, the WFR is more likely to cause respondents to alter eating behavior than the i-24HR. The i-24HR may also cause respondents to alter eating behavior, but in the WFR the presence of a RA at the home for the entire day is likely to be more disruptive. Respondents may simplify their true food intake in order to facilitate weighing of foods. Yet, even if WFR would alter true eating behavior, the foods actually consumed are measured accurately. The i-24HR is subject to reporting bias because amounts of food consumed are not measured at the time of consumption. The possibility of over- or underreport food intake remains. However, the i-24HR is less burdensome for respondents compared with the WFR. Conducting a WFR requires more resources than doing an i-24HR. Although the WFR is the most precise and accurate method for estimating food and nutrient intakes of individuals (Gibson 2005c), it is resource-intense, when RAs measure food intakes, and has high respondent burden. Using the WFR is difficult to implement on a large scale.

3.3.4 Assessing the relative validity of dietary recalls

Nutritional epidemiology investigates the relationships between diet and disease. Accurate measures of diet are essential for observing true diet-disease associations. Ideally, a dietary method is able to measure food consumption of individuals accurately. Validity in dietary assessment refers to the degree to which a measurement is an accurate measure of what it claims to measure (Nelson 1997). Unfortunately, performing dietary assessment always intervenes with the life of the subject in some way; whether it is knowledge of the upcoming interview, sensations of unease, or beliefs about how one should be eating. Thus, even if the measurement itself is accurate there is no guarantee that the measured food intake represents the true usual food intake of the subject (Block 1982). Due to the nature of dietary assessment, one can only assess the so called relative validity of measurement as opposed to absolute validity. Relative validity studies usually precede large epidemiologic studies. A relative validity study can be used to decide whether a method planned to be used in a large epidemiologic study gives estimates that are sufficiently accurate.

The relative validity of a dietary measurement can be assessed by comparing the method of interest (test method) to another dietary method (reference method). The choice of the reference method varies according to the test method. First of all, the reference method should be more accurate than the test method (Bailey 1978). Second of all, the sources of error of the test-reference method pair should be independent of each other (Nelson 1997). Also, the reference method must measure the food intake over the same time frame as the test method. The 24HR is a retrospective method that relies on the respondent’s memory and ability to quantify portion sizes accurately. The WFR is a prospective method in which all foods eaten by the subject are weighed and recorded. Thus, the WFR is considered as the most appropriate reference method for estimating the relative validity of dietary recalls (Gibson 2005b). With respect to sources of bias in measurement, the i-24HR and the 24HR are similar methods. Consequently, the WFR is the most suitable reference method also for the i-24HR (Gibson & Ferguson 2008).

Good agreement between the dietary intake results of the test and reference methods may not indicate validity but merely similar errors in both methods (Gibson 2005b). An alternative approach for assessing relative validity of dietary methods involves biochemical markers of nutrient intake (Gibson 2005b). Biomarkers are usually components of body fluids or tissues that have a direct relationship with dietary intakes of one or more dietary components. Biomarkers are used increasingly because they are independent of the measurement of food intake. The gold standard method of validating a dietary method is via triangulation – comparing the test method to a dietary reference method and biomarkers (Kaaks 1997). The use of biomarkers has helped in understanding the structure of measurement error in dietary assessment methods (Kipnis et al. 2002). Unfortunately, biomarkers exist only for a limited number of nutrients (Bates et al. 1997). Altogether, use of biomarkers is recommendable although several criteria need to be considered before a biomarker is adopted for use in dietary validation because the relationships between dietary nutrient intakes and status indices are complicated by interaction between nutrients and use of drugs, among other things (Bates et al. 1997). They also required considerable resources, which is often not feasible in research studies.

When assessing the relative validity of dietary methods, one should also collect information on factors that are known to be associated with food intake. Age, gender, social class, ethnicity, and marital status influence social behaviours and health, and findings relevant to one group cannot necessarily be extrapolated to other populations or subgroups (Macintyre & Anderson 1997). The frame of reference of the study should be used for identifying important variables associated with food intake (Nelson 1997).

Systematic underestimation of energy intake has been documented in high-income countries (Briefel et al. 1995, Klesges, Eck & Ray 1995). Psychosocial and behavioural characteristics may be the underlying factors that explain reporting bias (Maurer et al. 2006). Past research suggests that higher social desirability is a major factor that explains misreporting of energy intake (Maurer et al. 2006). Hebert et al. (1995) have defined social desirability as the tendency of an individual to answer to questions in keeping with social norms to avoid criticism (Hebert et al. 1995). Social norms vary according to the local culture and population under investigation (Alasuutari 2004). Therefore, knowledge on local cultural norms and habits is important when relative validity studies are planned.

Finally, carrying out a relative validity study should not be the end of the process. Since dietary data is guaranteed to contain measurement errors, the effect of measurement errors on the estimated relationship between diet and the outcome of interest should be taken into account (Kohlmeier & Bellach 1995, Margetts & Thompson 1995). The parameters that quantify the measurement error can also be used to estimate the effects of the degree of error on the power and sample size (White, Armstrong & Saracci 2008).

3.3.5 Statistical assessment of relative validity

Conventional approaches

Statistical assessment of relative validity is usually based on analysis at the nutrient level. Once the dietary intake values have been collected with the test and reference methods, the results obtained by the two methods are converted to energy and nutrients, and compared with each other. Several approaches are available for determining whether the differences between energy and nutrient intake measured by the test and reference methods are statistically significant. Available statistical methods are discussed in Nelson (1997). Nonetheless, the objectives of the study determine the broad types of statistical analysis required.

For studies using the i-24HR, two categories of objectives can be defined (Gibson & Ferguson 2008). First, if a study aims to characterize mean intakes of a group of individuals, the test method should yield similar mean intakes of nutrients as the reference method. In practice, agreement at the group level is assessed by comparing means (paired t-test) or medians (Wilcoxon matched-pair signed-rank test). If there is no difference between the means or medians at the chosen level of significance, the test method is valid at the group level. If the test method yields smaller or larger estimates than the reference method, the test method is said to under- or over-estimate the intake of the chosen nutrients.

Second, if a study wishes to determine the intakes of individuals, a more detailed analysis is needed. Correlation analysis, regression analysis, cross-tabulations and graphical investigations such as the Bland-Altman method (Bland & Altman 1986), among others, can be used for assessing relative validity at the individual level. There is no consensus on the most appropriate statistical methods for use in analysis at the individual level (Burema, Van Staveren & Feunekes 1995, Margetts & Thompson 1995). In addition, statistical measures are not unambiguous. Correlation coefficients can be high when agreement at the group level is poor, if a systematic bias exists. Some attempts have been made to establish guidelines about statistical measures as indicators of relative validity (Masson et al. 2002). Masson et al. (2002) have suggested that, at least when assessing the relative validity of food-frequency questionnaires, having more than 50 % of respondents correctly classified and less than 10 % of respondents grossly misclassified into thirds are desirable for nutrients of interest in epidemiological research. They also recommend Spearman rank correlation coefficients above 0.5. Most investigators advise exploring the data using several statistical methods.

Reporting-error sensitive analysis

Smith et al. (2007) criticize the conventional statistical methods that focus on comparing aggregate measures of energy and nutrient intake. Smith et al. (2007) point out that conventional approaches are indifferent to whether food items and amounts are reported correctly: reported information is converted to energy and nutrients regardless of whether food items were actually consumed. Thus, analysis at the nutrient level may misestimate reporting accuracy. Although some relative validity studies base their analyses solely on nutrients and do not report information on food consumption (Thakwalakwa et al. 2011), others have attempted to describe the magnitude of reporting errors using simple techniques. Reporting the frequency of memory lapses (Ferguson et al. 1989, Dop et al. 1994, Ferguson, Gibson & Opare-Obisaw 1994, Ferguson et al. 1995, Gewa, Murphy & Neumann 2009) and describing the accuracy of portion size estimates as a comparison of mean daily energy or gram intake from major food groups (Dop et al. 1994, Fisher et al. 2008, Alemayehu, Abebe & Gibson 2011) have been employed by some researchers.

Smith et al. (2007) advocate a new, more sophisticated approach for analyzing data from relative validity studies. They recommend evaluating the congruence between reported and reference information. Reported information is the set of item-amount pairs that a participant has reported eaten (e.g. cake, ½ serving), and reference information is the set of item-amount pairs that were actually eaten by the subject. Reported food intake should be broken down into intrusions (item reported eaten that is not in the reference information), matches (item reported that is in the reference information) and omissions (item in the reference information that is not reported eaten) (Figure 2, Table 1). Then, taking into account the amount of each food item, analysis should be based on five different categories of amounts: overreported amount from intrusions or matches, corresponding amount from matches, and unreported amount from matches or omissions. The framework is illustrated in Figure 2.

Suppose a subject has been observed to eat one serving of cake and ¼ banana, and reports having eaten ½ serving of cake and one apple. Converted to the five categories of amounts, we have ½ serving of cake in the category “corresponding amount from matches” since the cake was both eaten and reported, but only half of the eaten amount was reported. The unreported amount of cake from the match (½ serving) is treated as "unreported amount from matches". The ¼ banana is in the category "unreported amount from omissions" since the subject did not report eating any banana although a portion of banana was consumed. The apple is in the category "overreported amount from intrusions" because the subject did not eat any apple although reported doing so. The five categories of amounts can be based on servings, grams, or grams converted to energy or nutrients.

Grouping all food items into the five categories of amounts requires some efforts, but interpreting the data is much easier once the amounts are converted to three ratios (Table 1). First, the report rate is the ratio of the reported amount to the reference amount. Reported amount does not differentiate whether the food has been actually eaten or not. In fact, this ratio can be calculated for any relative validation study that reports estimates based on the test and reference methods separately. Second, the correspondence rate is the ratio of the corresponding amount to reference amount. This ratio shows how much of the reported intake was actually also eaten by the subject. If there are any reporting errors related to underestimating portion sizes or omitting some food items, the correspondence rate will be smaller than the report rate. The correspondence rate can never exceed 100 %. Third, the inflation ratio is the ratio of overreported amount to reference amount. The inflation ratio shows how much overreporting - whether it is reporting intrusions or exaggerating the amount from matches - inflates the report rate.

Although data on energy and nutrients are in most cases the epidemiologic interest, reporting-error sensitive analysis is relevant when studying dietary patterns and their relationship to health outcomes (Kant 2004), sources of nutrients and nutrient interactions (e.g. Shin et al. 2002, Baech et al. 2003, Choi et al. 2005), and differential reporting accuracy for different foods (Rosner & Gore 2001). Furthermore, analyzing variables that are sensitive to reporting errors gives valuable information about the method itself and provides insights into ways to reduce measurement error. Using the approach advocated by Smith et al. (2007) facilitates making comparisons of studies that resemble each other. Reporting-error sensitive analyses are used increasingly in relative validation studies (e.g. Paxton et al. 2011).

Table 1. Terminology on reporting-error sensitive analysis. Adapted from Smith et al. 2007.

Correspondence rate	The percentage of the reference amount to which the reported amount corresponds. It is the ratio of the corresponding amount to the reference amount, multiplied by 100. The correspondence rate is a genuine measure of accuracy with values between 0 and 100 %, with higher values reflecting higher accuracy.
Corresponding amount	The amount of a match that overlaps between the reported amount and the reference amount.
Inflation ratio	The ratio of the overreported amount to the reference amount, multiplied by 100. The inflation ratio quantifies inaccurate reporting. It has a lower bound of 0 %, but no upper bound since there is no limit on what a subject reports. Lower ratios indicate better reporting accuracy.
Intrusion	A food item reported eaten by the subject that is not in the reference set.
Match	A food item in the reference set that is reported eaten by the subject.
Omission	A food item in the reference set not reported eaten by the subject.
Overreported amount	The amount by which the reported amount of a match exceeds the reference amount, or the amount of an intrusion.
Reference amount	The amount of a food item in the reference set.
Reference information	The set of food items (reference set) and their amounts that were actually eaten by a subject.
Report rate	The ratio of the reported to reference amounts, multiplied by 100. Conventionally, values close to 100 % are considered as indicating high accuracy. The report rate of a subject is the sum of the correspondence rate and the inflation ratio.
Reported amount	The amount of a reported food item.
Reported information	The set of food items and their amounts reported by a subject.
Unreported amount	The amount by which the reported amount of a match falls short of the reference amount, or the amount of and omission.

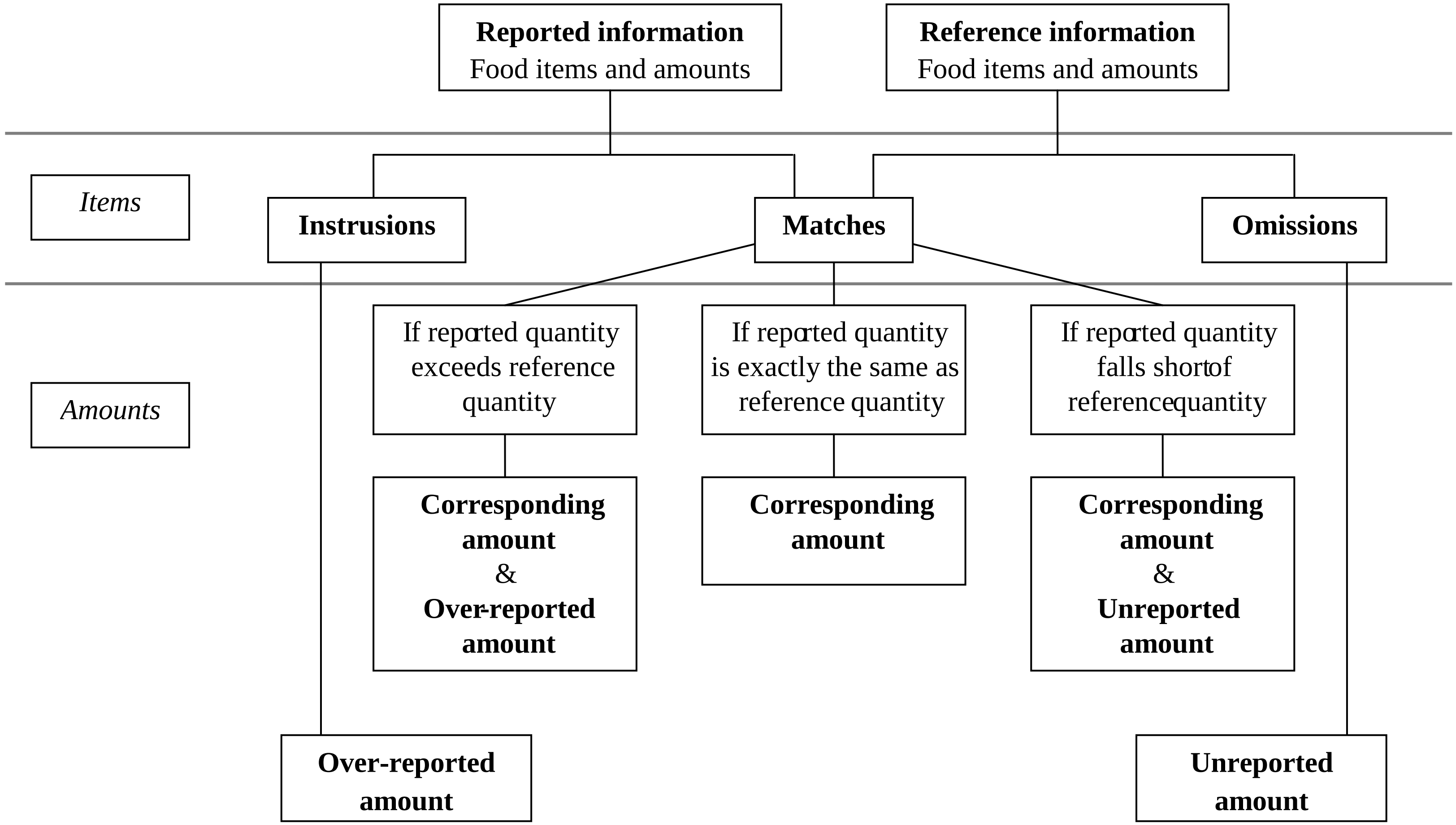

Figure 2. A framework for assessing portion size estimates and memory lapses as sources of measurement error. Congruence between reported and reference information is evaluated by classifying reported and reference items intrusions, matches and omissions. Amounts are classified as overreported, corresponding and unreported. Adapted from Smith et al. 2007.

3.3.6 Previous relative validity studies in young children in developing countries

The relative validity of the 24HR in young children has been studied by Dop et al. (1994) and Olinto et al. (1995), and the relative validity of the i-24HR by Thakwalakwa et al. (2011). Each of the studies compared the 24HR or the i-24HR with the WFR for one or two days of food intake. Surprisingly, there are not much comparable data available from high-income countries (Fisher et al. 2008). These studies are reviewed below and conclusions will be made about the evidence they give about the relative validity of dietary recalls among young children in resource-poor settings.

Dop and collaborators (1994) assessed the relative validity of the 24HR among 45 Senegalese children aged 11 to 18 months. A WFR and a qualitative observation were used as reference techniques against which the 24HR was judged. The qualitative observation was done on day 1. On day 2, the WFR was performed, as well as a 24HR concerning day 1. On day 3, the WFR was performed again, and a 24HR concerning day 2. On day 4, a 24HR concerning day 3 was done. Both the WFR and the qualitative observation were used for assessing the qualitative differences between reference techniques and the 24HR. Analysis by food group showed that most errors in the 24HR were omissions; foods that were eaten that were not reported. Omissions were presented by food groups as a dichotomous measure. Thus, if a food item was eaten many times and reported as eaten only once, no omissions were reported. As a proportion of consumers, omissions exceed 10 % for fish and millet-sorghum out of a total of seven food groups. Comparing portion size estimates, wheat products were the only food group whose portion sizes were significantly underestimated by 24HR (WFR mean 22.5 g per day vs. 24HR mean 18.3 g per day, paired t-test, $p=0.01$). This finding dealt with only one of the two days for which WFR and 24HR data were available.

Mean intakes of energy, protein, fat and carbohydrate were calculated for the two days. Mean differences between methods were <11 % of mean intakes. The significance of differences between mean intakes obtained by the 24HR and the WFR was examined with paired t-tests. Differences across methods in energy, protein, fat and carbohydrate intake were not statistically significant. Spearman rank correlation coefficients were calculated for the means of the two days. Correlation coefficients for fat, energy, protein and carbohydrate ranged from 0.70, 0.75, 0.75 to 0.80, respectively, indicating that there was strong association between the methods. Quantifying foods eaten from the household common pot as “handfuls” introduced a large random error to the nutrient intake estimates. The authors called for exploring the use of food models or actual foods as methods for measuring intake of rice. They concluded that the 24HR used in Senegal was accurate although it lacked precision.

Brazilian children living in slum areas were studied by Olinto et al. (1995). Food consumption of 50 children aged 6 to 19 months was measured with a single 24HR and compared with the results of a single WFR. The two methods measured the food intake on different days. Thus, inter-method differences at the individual level may reflect day-to-day variability in intakes instead of measurement error. The authors justified the choice of study design with the fact that the diets of children of low socio-economic status are very repetitive. The 24HR was performed at the beginning of the day and the same field worker stayed at the household for the rest of the day, weighing all food that the child ate.

Mean intakes were compared and paired t-tests used for analysing the significance of the differences between the 24HR and the WFR. Mean differences between the 24HR and the WFR ranged from 13 % (fat) to 24 % (protein). Recalled energy and protein ($p<0.001$) as well as fat ($p<0.005$) intakes were higher than weighed intakes. Overestimation was greater for underweight children than for normal weight children. Also, overestimation tended to be less marked for children aged 18 months and above, for girls and for educated mothers. Other sources of measurement error were not described. The authors recommended that future studies should take into consideration the possibility of overreporting among underweight children.

Thakwalakwa et al. (2011) conducted a study using the i-24HR in Malawi. A single i-24HR was compared with a single WFR for the same day of food intake to assess energy, protein, fat, iron, zinc and vitamin A intakes of 169 rural Malawian children aged 6 to 18 months. Guardians of the subjects were provided with picture charts before the recall interview, and some salted food models were used for estimating the amounts of foods eaten. The results from this study showed that recalled intake estimates were substantially higher than weighed intake estimates for energy and for each examined nutrient. Mean differences between i-24HR and WFR ranged from 13 % (energy) to 39 % (iron). Differences were statistically significant (paired t-test, each $p < 0.01$). The intraclass correlation coefficients varied from 0.42 (zinc and iron) to 0.83 (vitamin A).

Thakwalakwa and colleagues wanted to correct for the effect of systematic bias in measurement. This is one way of using the results of a validity study that precedes a larger epidemiologic study; the degree of measurement error is assessed and ways on how to correct for it are explored. Thakwalakwa et al. decided to develop an adjustment model that would yield correction coefficients for energy and each of the chosen nutrients. Then, multiplying the i-24HR values by the coefficients would provide an approximation of the mean values based on the reference method (WFR). The so called regression-through-the-origin model was used for developing these coefficients. The adjustment model showed that multiplying the mean energy, protein, fat, iron, zinc and vitamin A intake estimates based on the i-24HR by 0.86, 0.80, 0.68, 0.69, 0.72 and 0.76, respectively, predicted the mean values based on the WFR for rural Malawian children living close to Mangochi. The adjustment algorithms were calculated because they may be useful if their validity is ascertained in further studies in the same population.

A comparison of the previous relative validity studies of the 24HR and the i-24HR in children is provided in Table 2. Although each of the studies compared the 24HR or the i-24HR to the WFR, they differed in terms of what was measured, how the measurements were done exactly, as well as in terms of the population to which the method was administered to. In addition to differences in the characteristics of the subjects and measurement of breast milk intake (Table 2), the period of observation during which foods were weighed differed a little between the studies. RAs stayed at the household from 6 hours until 18 hours in the evening (Thakwalakwa et al. 2011), from 7 hours until bedtime (Dop et al. 1994), or for the entire day without specification about exact clock readings (Olinto et al. 1995). Also, means for quantifying portion sizes consumed varied; water and respondents' "handfuls" (Dop et al. 1994), foods available at the home (Olinto et al. 1995), as well as salted food models were used (Thakwalakwa et al. 2011). These differences illustrate the possible sources of variation in results.

Table 2. Comparison of studies that have assessed the relative validity of a dietary recall method against weighed food record among young children in developing country settings.

Author, year, setting	Method	n	Age range (months)	Breast milk intake	Dietary recall vs. weighed food record % difference[a] (p-value)[b] Energy	Protein	Fat	Iron	Zinc	Vitamin A	Carbohydrate
Thakwalakwa et al., 2011, rural Malawi	i-24HR	169	6-18	Not measured	13 (<0.01)	18 (<0.01)	34 (<0.01)	39 (<0.01)	28 (<0.01)	34 (<0.01)	-
Dop et al., 1994, suburban Senegal	24HR	45	11-18	Not measured							
Day 1					-0.2 (0.97)	-11 (0.13)	-3 (0.73)	-	-	-	2 (0.68)
Day 2					3 (0.53)	-4 (0.45)	6 (0.41)	-	-	-	4 (0.51)
Olinto et al., 1995, poor urban Brazil	24HR	50	6-19	Measured	17 (<0.001)	24 (<0.001)	13 (<0.05)	-	-	-	-

[a] (dietary recall-WFR)/WFR×100 for group mean intakes
[b] P-value of paired t-test between dietary recall and WFR
WFR, weighed food record; i-24HR, interactive 24-hour recall; 24HR, 24-hour recall; -, not assessed

Data from older children cannot be directly compared to studies done among young children because older children can participate in providing dietary data (Livingstone, Robson & Wallace 2004), and the eating habits of older children are different relative to young children. Nevertheless, the 24HR has given a valid indication of mean intakes of energy and several nutrients for older children in several developing country settings (Ferguson et al. 1989, Ferguson, Gibson & Opare-Obisaw 1994, Gewa, Murphy & Neumann 2009). An interesting finding was made in Ghana among five-year-old children; the relative validity of the 24HR was different in two villages that were selected to the study from the same region (Ferguson, Gibson & Opare-Obisaw 1994). The result emphasizes the importance of local validity studies.

Previous research suggests that getting acceptable estimates of mean intakes of energy and nutrients for young African children will be difficult when information is collected by the 24HR. Two out of three relevant studies found that food intake of young children was over-estimated (Olinto et al. 1995, Thakwalakwa et al. 2011). A comparable study from a high-income country supports these data (Fisher et al. 2008). It seems that methodological improvements to the 24HR for use in young children are needed. Unfortunately, only Dop et al. (1994) described sources of measurement error in their study making it difficult to target specific problems in measurement. Together, very few studies have been done to assess the relative validity of 24HR among young children, and only one has been done with i-24HR. Further studies are needed.

4 Aim of the study

The purpose of this methodological study was to assess the relative validity of an i-24HR used in a large randomized controlled intervention trial in Malawi. The study was carried out among 15-month-old rural children. The results of the i-24HR were compared to those of a more accurate method of dietary assessment, the WFR. Intakes of energy, protein, fat, iron, zinc and vitamin A were measured.

The study had two main objectives:

1. To assess, at the population level, whether average energy and nutrient intakes measured by i-24HR are different to those measured by the WFR.
2. To assess, at the individual level, whether the energy and nutrient intakes differ comparing the i-24HR and the WFR.

Secondary aims included:

1. To investigate whether memory lapses and incorrect portion size estimates are sources of measurement error.
2. To assess whether a differential measurement error might exist between the two study groups. The subgroups of this study consist of an intervention group that received a lipid-based nutrient supplement, and a control group that received no supplement.
3. To test whether correction values developed by Thakwalakwa et al. (2011) can be successfully used for adjusting values obtained by the i-24HR.

5 Subjects and methods

5.1 *Study design*

5.1.1 The main study

The present study was done for the purposes of a large clinical trial in Malawi. The trial is one of the many activities of a research collaboration called the International Lipid-Based Nutrient Supplements Project (www.ilins.org). The project aims to develop new LNSs and to investigate their efficacy in improving child and maternal health. The project has carried out trials in Burkina Faso, Ghana and Malawi.

The trial – within which the present study was conducted - is a randomized, controlled, single-blind trial that aims to identify the lowest daily dose and the most affordable formulation of LNS that is capable of preventing stunting (ClinicalTrials.gov identifier: NCT00945698). The primary efficacy outcome is the change in length-for-age Z-score. Participants are recruited at six months of age and they are supplemented with LNS until they reach the age of 18 months. A sample of 1920 infants was randomized into six groups receiving 10, 20 or 40 g of LNS that either contains milk or is milk-free. One group receives no supplement. The study is being carried out at two sites in Mangochi District, southern Malawi. The data collection period began in November 2009 and will be done by mid-2012. Dietary intakes from LNS and other complementary foods are being assessed with an i-24HR. Dietary assessment is made when participants are 9 and 15 months old. The trial is approved by the College of Medicine Research Ethics Committee (COMREC), Malawi and Pirkanmaa hospital district ethical board, Tampere, Finland.

5.1.2 The present study

The present study took place when the dietary assessment of the 9-month olds had been going on for some time in the main trial, and before the dietary assessments of the 15-month olds began. The i-24HR method was earlier pilot tested for use among 9-month old children. The dietary practices of 9-month old and 15-month old children are different. Hence, it was also necessary to pilot test the method for use among 15-month-old children. Thus, a cross-sectional pilot study was designed to take place in a sample of children outside the main study to see whether the i-24HR needed alterations. The initial planning of this study was done by the dietary assessment team of the main trial. Other activities - organizing recruitment, supervising data collection, recording and analyzing data - were the responsibility of the author of this report. Support and guidance were provided by the dietary assessment team during and after the data collection period.

The present study was done in four rural villages outside the catchment area of the main trial in Mangochi district. The villages were chosen because of their proximity to the Mpondasi health center. The recruitment sessions were held at the health center. The aim was to recruit a sample of approximately 60 children. To mimic the trial conditions, an intervention was included in the study. Half of the children were randomly assigned to receive LNS and the other half were assigned to follow their usual diet.

The children's food consumption was assessed using the i-24HR and the WFR. The i-24HR was done twice and the WFR was done once to estimate each child's food, beverage and LNS consumption. One of the two recalls was done before the WFR (i-24HR-I) and the other was done the day after WFR (i-24HR-II). This design was used so that dietary data were measured on the same day by the two different methods to eliminate the effect of day-to-day variability on inter-method comparisons. The first recall (i-24HR-I) was done to examine whether the WFR may have influenced the i-24HR-II results.

All dietary information was collected between June 17th 2010 and July 30th 2010, which corresponded to the post-harvest season in Malawi. The procedures for the study were considered as non-invasive and not harmful to the infant or the family. Hence, no separate ethical permission was sought from that of the main trial. The identity of the participants was kept confidential. All identifiable data are kept in a locked room at the premises of the main trial's study site in Mangochi. The study design, methodology and target population of this study are comparable with an earlier study on the relative validity of the i-24HR (Thakwalakwa et al. 2011). The exact protocols of the i-24HR method used by the two studies were a little different, subjects lived in slightly different types of environment and in the present study, an intervention was included.

Subjects and subgroups

The target population of this study consisted of healthy children aged 15 months living in rural villages of Chipeta, Misikatema, Saninkawa and Saiti kadzuwa in Mangochi district, southern Malawi. Participants were recruited with the help of health surveillance assistants (HSAs) who are a part of the formal Malawian health care system (Kadzandira & Chilowa 2001). Three HSAs who were working at the health center were hired to list children born in February, March and April 2009. The listing and enrolment took place on the third and fourth weeks of June 2010. Thus, at the time of enrolment participants were roughly from 13.5 to 16.5 months old. Each HSA was asked to invite approximately 20 children to participate in the study. HSAs listed the children by visiting households in the villages. They asked the date of birth of the child from the guardian, and explained the purpose of study. Interested caregivers were invited to gather at the local health center at a scheduled time to hear more about the study. Information was not collected about how many children the HSAs had listed and invited altogether.

Interested guardians came to the health center on the planned date. The recruitment session was held by RA who are fluent in the local Chichewa language. One of the RAs explained the details of the study at the same time to all the guardians. He explained the purpose of the study, practicalities, schedule, risks, information security, rights of the participant, and compensation for participation. A bar of soap was given at each visit as compensation for time and effort. Additionally, the control group would receive 2 kg of nutritious porridge after the completion of the study as the intervention group would receive a nutrient supplement for the duration of the study. The RA presented the information in a structured way using a letter of information translated into Chichewa as a memory aid (Annex 1, in English). After the information session, guardians decided whether they wanted their child to take part in the study or not, and indicated their decision verbally. At enrolment, all caregivers were mothers. No written consents were collected.

All mothers were willing to participate. Next, children were screened for chronic illness and peanut allergy, and tested for tolerance of LNS that was to be used by the participants in the intervention group. Each child was given a spoonful of LNS. They were followed for 15 minutes to ensure they had tolerated LNS. If a child would have developed symptoms, immediate medical attention would have been sought from the health center. If the child was not chronically ill, had no peanut allergy and tolerated LNS, he or she was eligible for randomization. Malaria was not an exclusion criterion.

The number of eligible infants was counted and a corresponding amount of lottery tickets of identical appearance was made. Half of the tickets had value 0 and the other half value 1 written on them, indicating assignment into control and intervention groups, respectively. Tickets were mixed up in a bowl and one additional ticket was put in so that even the last mother choosing a ticket would be able to choose a ticket out of two. Mothers chose the tickets one by one, waited until everyone had picked a ticket, and they opened the tickets at the same time. With this arrangement participants were allocated into one of two groups randomly. Participants that had been assigned to the control group were thanked for their decision to participate and reminded that they would be visited within roughly two weeks. At this point, snacks were served to all mothers and their children before the controls left for home and before giving further instructions to the guardians of children in the intervention group.

Once the snacks were finished, mothers of the participants in the intervention group were given instructions on how to use the supplement. A 5 g measuring spoon and a three-week supply of LNS were given to the mothers. More precisely, the mothers received 6 jars of LNS weighing 140 g and they were instructed to mix four 5 g spoonfuls of LNS with approximately 2 to 3 tablespoons of porridge, and feed the small amount of porridge at two separate occasions to the infant before feeding the remainder of the porridge meal. Twenty grams of type of LNS used in this study (Nutributter®, Nutriset S.A.S., France) corresponds to the recommended daily dose for a child. Mothers also received these instructions in writing. Furthermore, messages on how to use LNS were reinforced a week after the recruitment session because observations from the main study had shown that it was difficult to scoop out the correct amounts of LNS. The household of each child receiving the LNS was visited and instructions were discussed with the guardian. Participants of the control group were not visited because they did not consume the supplement. The recruitment session was held on two separate occasions. At both times, all mothers were willing to take part in the study, none of the children had peanut allergy and all children tolerated LNS.

Study protocol and visits

The study consisted of three dietary assessment visits. The study protocol is illustrated in Figure 4. The i-24HR-I was done prior to the WFR. The WFR and the i-24HR-II were done on subsequent days (Figure 3). Prior to any of the dietary visits, children of the intervention group consumed LNS for at least two weeks, and all participants were provided with equipment (i.e., picture charts, bowls and cups) required for carrying out the dietary visits. The earliest time the first dietary assessment visits were planned to take place was two weeks after recruitment.

Day	i-24HR-II	WFR
1	Cups, bowls and picture charts are delivered to respondents	
2	Respondents use the cups, bowls and fill the picture charts	A research assistant performs a weighed food record
3	A research assistant performs a recall interview on day 2	

Figure 3. Timing of interactive 24-hour recall and weighed food record visits. i-24HR-II, interactive 24-hour recall concerning the same day of food intake as WFR; WFR, weighed food record.

Use of the lipid-based nutrient supplement

The two-week period served as a period of habituation, for LNS consumption, that allowed the infant to get used to the taste and texture of the supplement, and allowed the guardian to incorporate the supplement into the usual diet of the child. Earlier observations on Malawian children have revealed that when LNS is consumed daily, more than one week and less than four weeks is needed to stabilize the feeding patterns (Flax et al. 2008). Flax et al. (2008) did not observe feeding practices during the interval between week one and four to precisely define when habituation took place. Thus, in this study, the participants received LNS for at least two weeks before assessing their dietary intakes. A period of two weeks was chosen arbitrarily based on the previous study and resource considerations. The type of LNS that the children consumed was commercially available LNS designed for the prevention of malnutrition (Nutributter®, Nutriset S.A.S., France). This supplement was not used by the main trial, but it was similar to the LNSs used in the main trial. Although the control group had no habituation period, their visits took place at the same time as the visits of intervention group.

Dietary assessment visits

Once the two-week habituation period was over, the cups, bowls and picture charts needed for the dietary assessment were delivered to the guardian two days before the first i-24HR interview. Guardians were asked to feed the child from the provided cups and bowls on the next day, and mark off the foods, when they were eaten, on the provided local food picture chart. When the picture chart was delivered to the guardian, its use was rehearsed on one blank chart to show the guardian how to use it, and another was given for use on the following day. Then, after two days, a RA returned to the household to do the i-24HR about the child's food consumption on the previous day.

The WFR was a 12-hour visit during which the food consumption of the child was observed starting from 6:00 hours in the morning lasting until 18:00 hours in the evening. Soon after randomization it was noted that many of the WFR visits were planned to be done three weeks after participants had been provided with LNS, i.e. the LNS would be finished before the WFR visit for some participants. Thus, the protocol was modified and the RAs doing the WFR visits were instructed to take a jar of LNS with them and give it to the guardian in the beginning of the day, if they were in the intervention group, to make sure that the participants in the intervention group had LNS available. This arrangement was justified because observations from the main study have shown that children usually finish their LNS ahead of time. Reasons for this are not known. It is possible that LNS is shared with other family members or larger daily portions are served than instructed. The day after the WFR visit, i-24HR-II was done for the same day of food intake as the WFR. The two visits were done by different RAs so that the RA doing the i-24HR-II would be blinded to direct observation. RAs located the households of the participants as geographic coordinates using a global positioning system receiver (eTrex H, Garmin Ltd, Taiwan).

Background characteristics data collection

Some questions about breastfeeding and the health status of the child were asked as a part of the i-24HR. The respondent was asked whether the child had been sick the previous day and whether it had affected the appetite of the child, and whether the child had breastfed. No information was collected on socio-economical status, food security or other background characteristics. The sex of the child, however, was determined after data collection, based on the names of the children. A Malawian researcher was able to conclude that out of the 68 participants, 38 were girls and 26 boys. The sex of four participants was left unsolved.

The participant flow is shown in Figure 4. Three participants were found to be too old or too young after randomization, all from the control group. The guardians were explained why their children could not be included in the study. Nevertheless, compensation gifts were given. One of the i-24HRs had to be omitted because the participant had not received the equipment needed for carrying out the i-24HR. Nine of the i-24HR visits had to be excluded because they were not done the day after WFR as planned. Two visits were missed due to incorrect information about the location of the household. Other reasons for missed visits included that the mother of the participant was working elsewhere, seeing relatives, or taking care of her husband at the hospital. One participant had moved away temporarily from the mother's home to the father's house. One participant was hospitalized. One mother was not willing to participate in the study when a RA came for the first visit. Forms for 13 visits were missing without explanation. In Figure 4, the term missed refers to both missed visits and forms that were missing. Altogether, 65 % of randomized subjects completed the i-24HR-II and the WFR while 50 % completed all three dietary visits.

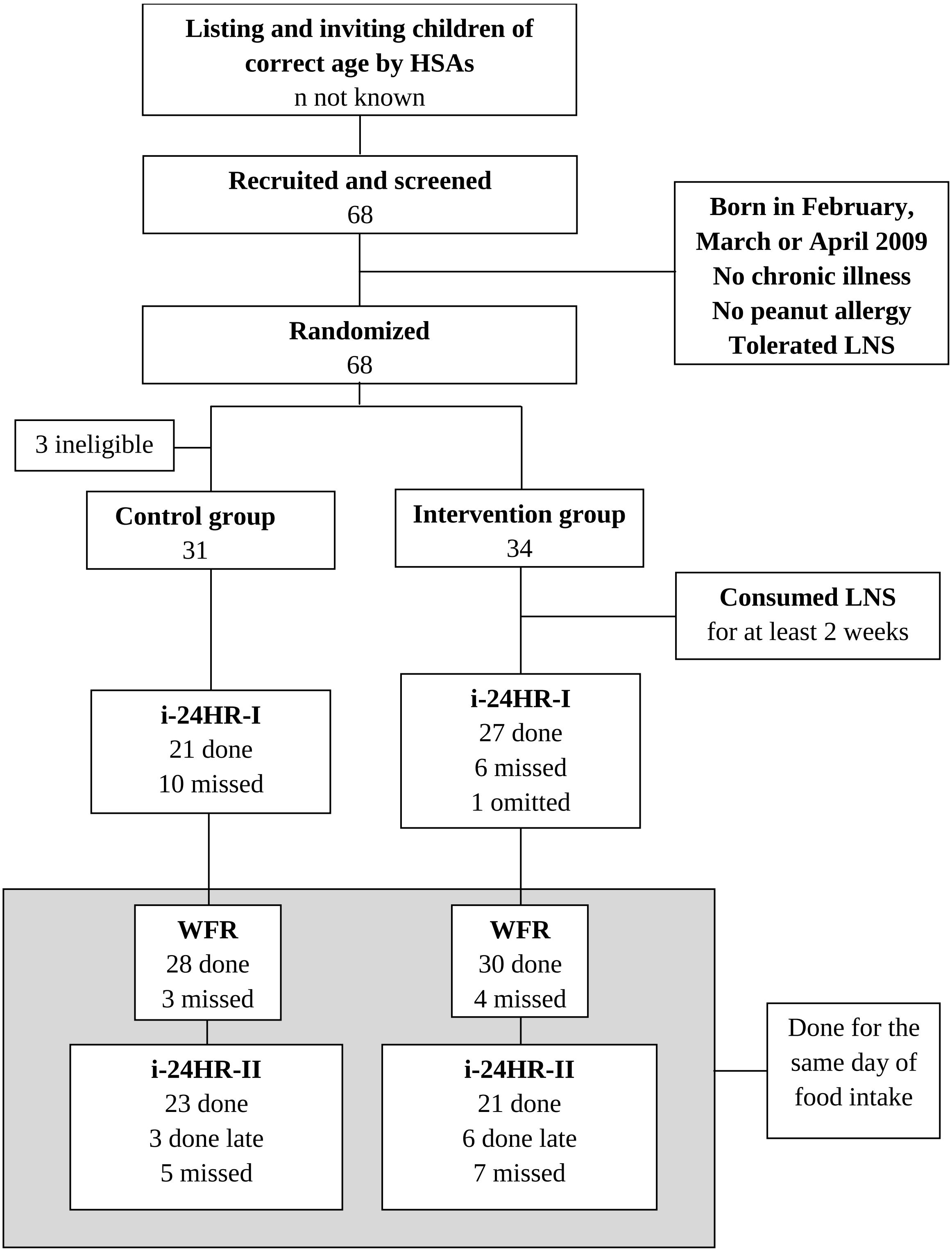

Figure 4. Study profile and participant flow. HSA, health surveillance assistant; LNS, lipid-based nutrient supplement; i-24HR-I, baseline interactive 24-hour recall; i-24HR-II, interactive 24-hour recall done on the same day as WFR; WFR, weighed food record.

5.2 *Dietary methods*

Information about each child's food consumption was collected with two different methods for the same day of food intake. The results of i-24HR (test method) were compared to those of a more accurate method, the WFR (comparison method). Breast milk intake was not measured. Ten male and female RAs with a minimum of secondary school level of education collected the data. All information was collected at the home or another convenient place for the respondent. Three RAs performed i-24HRs, six performed the WFRs and one RA participated in performing both i-24HRs and WFRs, but never for the same child. RAs doing the i-24HRs had four months of experience in doing corresponding interviews for guardians of nine month old Malawian infants within the main study.

The RAs had been trained to build rapport with respondents, quantify food and beverage intakes using food models, ensure completeness of the i-24HR, and to use the picture charts. Training included practice with estimating portion sizes and practice with the four phased i-24HR, including estimating portion size of LNS. Five RAs doing the WFRs had previous experience in collecting them and two were trained to perform them for this study. WFRs were done during the week, but Saturdays and Sundays were used if a visit had been missed. All forms used for data collection were checked for data completeness on average (mean, minimum-maximum) 4 days (0-12) after each visit. Suspicious and missing information were discussed with the RA that had collected the data, and if possible, information was completed. Finally, paper forms were scanned into electronic files.

5.2.1 Weighed food record

The WFR was chosen as the comparison method because it is the recommended reference dietary method for relative validity studies of the 24HR (Gibson 2005b). A RA went to the home of the participant at 6 am, and observed the food consumption of the child for the next 12 hours filling all necessary details to a special form developed for the WFR visit (Annex 2). All food and drinks that the child consumed over the 12-hour period were weighed and recorded. The RAs were instructed to follow the infant for the full duration of the day, including times when the mother was bringing the infant to work or when the infant was left in the care of a community member. A digital scale accurate within ±1 g and able to measure items of maximum 5 kg was used (Digital Kitchen Scale, Home Elegance, South Africa). The precision of each scale was monitored by weighing an object of known weight daily.

The guardians of the participants had been provided with a standard sized cup and bowl prior to the WFR. They were asked to use the utensils when feeding the infant. Using shared plates is a common practice in Malawi. Serving separate portions makes it possible to weigh foods eaten by the child. For each food eaten, information was recorded about who was feeding the child, and where and when the food was consumed. The ingredients and brand of the food were recorded as well. The served portion was weighed before the child was fed. Once the child had finished eating, the portion left over was weighed and recorded. Readings of the scale were not made known to the caregivers. In addition to recording what the child ate, recipes of foods that were cooked at the home were written down. The raw ingredients were weighed, and the weight of food after cooking recorded so that the proportion of each ingredient in the dish could be calculated.

5.2.2 Interactive 24-hour recall

Ferguson et al. (1995) have developed a modification of the traditional 24-hour recall to collect information on rural populations in developing countries. This i-24HR is a quantitative dietary method. The modifications help visualize foods eaten, reduce memory lapses, and help in estimating portion sizes. In this study, guardians of participants were asked to use standard sized bowls and cups when feeding the child, and mark off each eaten food on a picture chart the day before the i-24HR. The use of the i-24HR is discussed in more detail in a manual prepared by Gibson and Ferguson (2008).

Two days prior to the i-24HR each guardian was provided with a standard sized cup and bowl, and a picture chart of most common locally eaten food items. In the picture chart foods were categorized into 20 groups: nsima (maize-based staple food), porridge, sweet potato/cassava, ground nut flour, beans, fish, meat, egg, fruits, infant formulas, vegetables, pastries, beverages, milk, sugar, fat, sweets, insects, LNS, and other foods (Annex 3). Guardians were asked to feed their infant from the cup and bowl instead of their own utensils on the following day, and mark off each food eaten on the picture chart at the time of consumption.

The i-24HR is an interview that gradually collects information of foods eaten during the previous day, starting from the first food eaten and ending with the last food eaten. The i-

24HR proceeds in four phases whereby a structured form is filled (Annex 4). In the first phase the RA asked the respondent to list everything the child had eaten and drank the previous day. In the second phase more details were asked about the listed foods: time of day and place where the food was eaten, person feeding the infant, brand names of foods, ingredients of composite dishes, and preparation method of the foods.

The third phase of the i-24HR quantifies the amounts of food items consumed by weighing food models whose density is close to the density of the actual food eaten. RAs carried with them a collection of fresh and salted food models. New food models were prepared weekly. Food models included food items that are commonly eaten among this population during this season; thin, medium and thick porridge, nsima (maize-based staple food), sugar, milk powder, margarine, ground nut flour, LNS, cooked red beans, dry beans, fresh fish, maize puffs, salted pumpkin leaves, white sweet potato, and mandasi (African doughnut). The RA chose the appropriate food model as per instructions in a portion size guide developed to be used when measuring food intakes of 9-month olds in the main trial. The portion size guide included instructions on which food model to use for different foods and how to ask details about the foods.

To quantify the amount food eaten, the respondent was asked to serve the amount of food the infant ate the previous day using the same utensils that were used when feeding the infant. After the quantity of the served amount was recorded, the respondent was asked whether the infant had finished the portion. If not, the respondent was asked to show how much there was left-over by removing the amount eaten from the bowl or cup. The remaining portion was weighed. If there were no leftovers, the empty bowl was weighed. Some food items were measured using monetary units (e.g. package of biscuits) or actual units (e.g. boiled egg).

Once all foods were listed and their portion sizes estimated, the respondent was asked to provide the picture chart filled the day before. The RA examined the picture chart comparing it to the list of foods just recalled to see whether the two lists were identical. Discrepancies were discussed. The picture chart is based on food groups of ingredients or foods and does not reflect consumption of individual meals. For instance, a vegetable relish made out of pumpkin leaves, cooking oil and salt should result in marking off categories “vegetables” and “oils and fats” from the picture chart.

In case there was something marked on the picture chart that was not appearing on the recall, the respondent was asked whether the infant had in fact consumed the food item. Likewise, if there was something on the recall not appearing on the picture chart, the respondent was asked if the infant had consumed that food item. Corrections were made to the recall accordingly.

Finally, in the fourth phase of the i-24HR, the RA summarized the list of foods and drinks the respondent had provided and asked if that was an accurate representation of the child's food consumption of the previous day. If the respondent remembered an additional item, it was added to the recall with details. If the respondent suggested something was added by mistake, the food item was removed from the recall. Respondents were also asked whether feeding their child from a bowl and cup provided by the study had influenced the amount or type of foods their child ate.

5.3 *Data preparation and statistical analysis*

5.3.1 Data preparation

During weighed food record (WFR) visits, recipes of mixed dishes were recorded. Individual recipes of dishes with the same ingredients (e.g. dry kidney beans with tomato, cooking oil and salt) were compiled into average recipes. The number of individual recipes used for compiling each average recipe ranged from one (ground nut flour soup) to 23 (nsima, a maize-based staple food). These average recipes were used for computing raw ingredient intakes from mixed dishes reported in i-24HR-I and i-24HR-II. For WFR, raw ingredient intakes were calculated using the individual recipes related to that specific visit. Average recipes were used for the WFR data if the individual recipe was missing. This was the case if a neighbor had served food to the child, for instance.

Food intake from i-24HR-I, i-24HR-II and WFR were converted into grams consumed. In many cases lists of foods eaten during WFR and reported being eaten by i-24HR-II were not alike. Initially, the WFR conducted from 6:00 hours to 18:00 hours and the i-24HR-II were expected to correspond to each other. It gets dark before 18:00 hours and there is no electricity in every household in the villages. Moreover, electricity breaks are common. Thus, cooking was not expected to take place after 18:00 hours. However, it turned out that 22 out of 44 participants (50 %) had eaten foods after the WFR visit had ended. Twenty out of the 22 subjects (91 %) had consumed a warm meal. It was concluded that the WFR and the i-24HR concerning the same day of food intake were not comparable as such.

In order to compare the WFR and the i-24HR-II it was necessary to exclude food items eaten after 18:00 hours from the i-24HR-II. The respondents of the i-24HRs had reported the times of feeding as clock readings. A food item was considered as consumed after 18:00 hours if it was both reported as consumed after 18:00 hours and there was no corresponding food item on the WFR form. The i-24HR-II that covers foods consumed between 6:00 hours and 18:00 hours is referred to as the 12-hour interactive 24-hour recall (12-i-24HR). The lists of foods eaten by each subject as measured by the WFR and the i-24HR-II are provided in Annex 5. Because i-24HR-II was intended to be compared not just with WFR but also with i-24HR-I, amounts of all foods reported in i-24HR-II, including those consumed after 18:00 hours, were calculated into grams, too. Finally there were four sets of food intake data: i-24HR-I, i-24HR-II, WFR and 12-i-24HR (Figure 5).

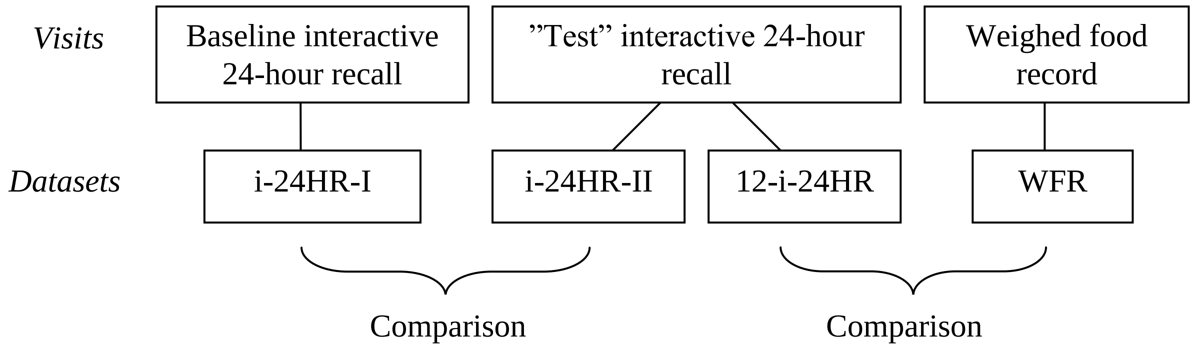

Figure 5. Four datasets and their purposes.
i-24HR-I, baseline interactive 24-hour recall; i-24HR-II, interactive 24-hour recall concerning the same day of food intake as WFR; 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours; WFR, weighed food record

Intakes of energy, protein, fat, iron, zinc and vitamin A were calculated using a computer program installed in Microsoft Excel 7.0. The program was linked to a food composition database, both developed for an earlier study (Ndekha et al. 2000). The program and composition table are the same that Thakwalakwa et al. (2011) used in their study. The food composition database comprises of 118 food items from the Kenyan food composition table (Ministry of health, 1993, Kenya), 64 items from the international Minilist nutrient database (World Food Dietary Assessment System, version 2.0., Office of Technology Licensing, University of California Berkeley, USA), and some foods analysed by the Malawi Bureau of Standards.

For this study, the food composition database was complemented with 16 foods. Nutrient composition of soy flour, doughnuts and maize snacks were retrieved from USDA National Nutrient Database for Standard Reference (U.S. Department of Agriculture, Agricultural Research Service, 2010, USDA National Nutrient Database for Standard Reference, Release 23). Nutrient compositions of 12 broths were retrieved from the Tanzanian food composition tables (Lukmanji et al. 2008). In addition, nutrient composition of the LNS used in this study was added. Per 100 g, the LNS had 540 kilocalories energy, 12.5 g protein, 35 g fat, 20 mg zinc, 45 mg iron and 2000 µg vitamin A.

5.3.2 Statistical analysis

Normality tests

The normality of intake distributions of energy, protein, fat, iron, zinc and vitamin A was assessed for the four sets of data (i-24HR-I, i-24HR-II, WFR and 12-i-24HR) using the Kolomogorov-Smirnov test, each separately for the intervention and control groups (Table 3).

Table 3. Normality of crude variables in four datasets.

	Normality of distribution[a]					
Dietary factor	Energy	Protein	Fat	Iron	Zinc	Vitamin A
i-24HR-I						
Intervention	yes	yes	yes	yes	yes	no
Control	yes	no	no	no	yes	no
i-24HR-II						
Intervention	no	no	no	yes	no	yes
Control	no	yes	yes	yes	yes	no
WFR						
Intervention	no	yes	yes	no	no	no
Control	yes	no	yes	no	no	no
12-i-24HR						
Intervention	no	no	yes	yes	no	yes
Control	yes	no	no	yes	no	no

[a] Yes = Kolomogorov-Smirnov test $p \geq 0.200$, no = Kolomogorov-Smirnov test $p < 0.200$

Log-transformations were made for the non-normally distributed variables. Log-transformations improved the normality assumptions of some variables. However, the normally distributed log-transformed variables did not perform better in tests and analyses assuming normality compared with crude variables analyzed with non-parametric methods (data not shown). For this reason, crude, non-normally distributed variables are described and results of statistical analyses that do not assume normality are presented.

Determining the influence of WFR on 12-i-24HR

The i-24HR-I and the i-24HR-II were compared to determine whether doing the i-24HR-II on the same day as a the WFR had an effect on the recalled estimates. The significance of the difference between intake values of the i-24HR-I and the i-24HR-II was evaluated by Wilcoxon's matched-pair signed-rank test.

Assessing the relative validity of 12-i-24HR

The 12-i-24HR and the WFR were compared to assess the relative validity of the 12-i-24HR. Statistical analyses consisted of group average comparisons, within subject comparisons and ranking of subjects. All analyses were done separately for the intervention and control groups. Median intakes obtained by the 12-i-24HR and the WFR were compared and the significance of the difference between them assessed using the Wilcoxon matched-pair signed-rank test.

The strength of relationship between the 12-i-24HR and the WFR was assessed by Spearman rank correlation coefficients. Subjects in intervention and control groups were also combined into one group and partial correlation analysis was done to see if statistical power would increase. The ability of the 12-i-24HR to separate subjects into classes of intake was evaluated by ranking the data into tertiles and by calculating the percentage of respondents classified into the same and opposite categories by the 12-i-24HR and the WFR. A graphical comparison of the 12-i-24HR and the WFR was done using the Bland-Altman analysis (Bland & Altman 1986). The Bland-Altman analysis is used to plot the individual differences between two methods against the mean level of intake, followed by the calculation of limits of agreement (mean difference ± two standard deviations of the difference). The Bland-Altman plot indicates if the measurement error is constant across all levels of intake. Outliers in the analyses based on energy and nutrient intakes (Wilcoxon matched-pair signed-rank test, Spearman rank correlation coefficients, tertile agreement and Bland-Altman plots) were explored to find out the likely reasons for disagreement between the 12-i-24HR and the WFR at the food level.

Testing the feasibility of correction values developed for an earlier study

Correction values developed by Thakwalakwa et al. (2011) were tested on the present data by multiplying the 12-i-24HR median intakes of energy and selected nutrients by the correction values. The corrected values were compared with estimates based on the WFR to determine whether the correction values would bring estimates based on the 12-i-24HR closer to the reference information (WFR).

Describing sources of measurement error

Portion size estimates, intrusion of food items to the 12-i-24HR and omission of food items from the 12-i-24HR were assessed as sources of measurement error according to a framework introduced by Smith and co-workers (2007). For each participant, the list of food items in the 12-i-24HR was compared to information obtained by the WFR, and each food item was labelled as a match, an intrusion or an omission. Then, the amounts of food items were classified into five categories of amounts; 1) corresponding amount from matches, 2) overreported amount from matches, 3) unreported amount from matches, 4) overreported amount from intrusions, and 5) unreported amount from omissions. Report rates, correspondence rates and inflation ratios were calculated from the five categories of amounts. The amounts were based on gram intakes of foods from 12 food categories.

An indicative assessment of the existence of differential measurement error was done by calculating the so called factor A for energy and each selected nutrient, as White (2003) has advised. Differential measurement error occurs when the measurement error differs between subjects in two groups (White 2003). Statistical analyses were performed using PASW Statistics 18.0 software (IBM SPSS Inc, NY, USA) and StataSE 10.0 (StataCorp, College Station, Texas, USA).

6 Results

6.1 *Characteristics*

6.1.1 Subjects

There were more girls (56 %) than boys (38 %) among the 68 recruited participants. Information about sex was not available for four (6 %) of the participants. Proportions of girls and boys remained almost the same for participants who were included in analysis (Table 4). The participants included in analysis were on average 16 months old (Table 4). The age on the date of WFR is reported.

Table 4. Background characteristics of 44 participants included in analysis.

Background characteristics	Intervention (n=21)[a]	Control (n=23)	All[a]
Age in months, mean (SD)	16.3 (0.7)	16.2 (0.7)	16.2 (0.7)
Sex, n (%)			
Female	14 (67)	11 (48)	25 (57)
Male	6 (29)	10 (43)	16 (37)
Not known	1 (5)	2 (9)	3 (7)
Village, n (%)			
Chipeta	4 (19)	3 (13)	7 (15)
Misikatema	10 (48)	8 (35)	18 (41)
Saninkawa	3 (14)	6 (26)	9 (21)
Saiti kadzuwa	4 (19)	6 (26)	10 (23)

[a]The sum of the percentages exceeds 100 % due to rounding.

Some questions about the health status and feeding patterns of the child were asked together with the i-24HR. Table 5 summarizes the results obtained from the i-24HR-II. Ninety-one percent of the children were breastfed and 86 % of the respondents considered that the appetite of the child had been usual the previous day.

Table 5. Health and feeding patterns on observation day in intervention and control groups.

Health and feeding patterns, n (%)	Intervention (n=21)	Control (n=23)	All (n=44)
Breastfed	18 (86)	22 (96)	40 (91)
Sick	1 (5)	4 (17)	5 (11)
Food intake usual	20 (95)	18 (78)	38 (86)
Feasting	1 (5)	1 (4)	2 (5)
Fasting	0 (0)	4 (17)	4 (9)

Food consumption

The diet of the subjects of this study consisted mainly of nsima (maize-based staple food), maize porridge, fish and vegetable relishes, and tea. Median intakes of foods recorded during the WFR are summarized in Table 6. Twelve food categories are presented. Animal-origin relishes entail dishes with fish, egg, chicken or beef as the main ingredient. Relishes with green leafy vegetables, okra, cabbage and beans are categorized as plant-origin relishes. All relishes consumed by the children had one main ingredient and some or all of the following ingredients: tomato, cooking oil, onion, salt. Broths include watery soups and the liquid part of relishes when the solids were not eaten. Most of the potato consumed by the children was sweet potato. One participant ate Irish potato. All consumed bread was white bread. African cake and African doughnuts were the two types of cakes consumed. LNS was consumed by 15 out of 21 participants in the intervention group. The main meals consumed by the children were produced at home. Most of the children were served snacks such as sweets, maize puffs and cakes bought from the local market. Meals were served typically two to three times per day. The morning meal consisted usually of maize porridge. A typical meal served during the day consisted of nsima and one type of relish, usually made out of fish.

Table 6. Frequency (proportion) of consumers and median (25th, 75th percentile) intakes of foods (g) between 6:00 and 18:00 hours according to weighed food record (n=44).

Food category	Consumers n (%)[a]	Median, g[b] (25th;75th percentile)
Porridge	40 (91)	164 (112;206)
Nsima, rice	42 (95)	83 (50;155)
Animal-origin relishes	14 (32)	34 (23;62)
Plant-origin relishes	24 (55)	35 (18;78)
Broths	5 (11)	27 (18;30)
Potato, cassava	14 (32)	61 (25;104)
Bread, cakes	14 (32)	21 (15;34)
Tea, soda, milk	30 (68)	105 (66;164)
Fruits	5 (11)	71 (14;155)
Sweets, biscuits	5 (11)	6 (5;27)
Puffs, popcorn, nuts	8 (73)	17 (8;30)
LNS [b]	15 (71)	16 (7; 26)

[a] Number of subjects who were observed consuming something from the food category once or more times during weighed food record. The percentage is calculated as the percentage of the total sample (n=44), except for LNS.
[b] Median intake calculated for subjects who consumed the food
[c] LNS was consumed by the subjects in the intervention group (n=21)
LNS, lipid-based nutrient supplement

6.1.2 Surrogate respondents

Mothers provided the information for 91 % and 87 % of i-24HR-I (n=34) and i-24HR-II (n=44), respectively. Taking together the remaining i-24HR-I and i-24HR-II visits, the rest of the information was provided by two grandmothers and one sister. Information was missing for five subjects. During the WFR no one was interviewed but the person feeding and cooking food for the child was followed. In 39 of the 44 WFRs (87 %), this person was the mother. The remaining five visits were covered by three aunts, one grandmother and one sister. Forty-nine and 51 % of recall interviews were performed by female and male RAs, respectively, taken together the number of all i-24HR-I and i-24HR-II visits.

Discrepancies

One respondent had ticked all foods in the picture chart and three respondents had not filled the chart at all. Discrepancies between the recalled foods and the foods marked on the picture chart were common. Of i-24HR-I and i-24HR-II, 9 % and 27 %, respectively, were fully in line with the picture chart. The most common reason for discrepancies was marking the picture chart fewer times than the ingredient was eaten if a food was eaten many times (e.g. cooking oil).

6.2 *Comparison of two interactive 24-hour recalls*

The two i-24HRs done before and after the WFR were compared to find out whether the WFR had an effect on the i-24HR concerning the same day of food intake as the WFR (i-24HR-II). A baseline i-24HR (i-24HR-I) was compared with i-24HR-II. None of the differences were statistically significant in the intervention group (Table 7). Ranges of intake between 25th and 75th percentiles were narrower in i-24HR-II than in i-24HR-I for energy and each reported nutrient.

Table 7. Intervention group: comparison of baseline interactive 24-hour recall and interactive 24-hour recall concerning the same day of food intake as weighed food record (n=18).

Dietary factor	i-24HR-I median (25th, 75th percentile)	i-24HR-II median (25th, 75th percentile)	Difference[a]	% difference[b]	P-value[c]
Energy, kJ	2313 (1599; 3045)	2052 (1800; 2539)	-261	-11	0.53
Protein, g	11.4 (8.4; 17.7)	11.3 (6.9; 12.9)	-0.1	-1	0.45
Fat, g	20.2 (9.4; 30.3)	15.4 (11.6; 20.3)	-4.8	-24	0.31
Iron, mg	12.3 (4.1; 18.4)	9.9 (5.9; 13.7)	-2.4	-20	0.50
Zinc, mg	5.8 (2.6; 8.2)	4.4 (3.2; 7.9)	-1.4	-24	0.71
Vitamin A, µg	619 (312; 1198)	643 (347; 1026)	24	4	0.74

[a] i-24-HR-II – i-24HR-I
[b] (i-24HR-II-i-24HR-I) / i-24HR-I × 100 for median intakes
[c] P-value of Wilcoxon matched-pair signed-rank test between i-24HR-I and i-24HR-II
WFR, weighed food record; i-24HR-I, baseline interactive 24-hour recall; i-24HR-II, interactive 24-hour recall concerning the same day of food intake as WFR.

For the control group, the intake of energy and each selected nutrient was lower in i-24HR-II than in i-24HR-I (Table 8). The difference was statistically significant for intake of protein and fat (Wilcoxon matched-pair signed-rank test $p < 0.05$), and the difference in energy intake showed borderline significance. Ranges of intake between 25th and 75th percentiles were narrower in i-24HR-II than in i-24HR-I for all other estimates except for protein.

Table 8. Control group: comparison of baseline interactive 24-hour recall and interactive 24-hour recall concerning the same day of food intake as weighed food record (n=16).

Dietary factor	i-24HR-I median (25th, 75th percentile)	i-24HR-II median (25th, 75th percentile)	Difference[a]	% difference[b]	P-value[c]
Energy, kJ	2215 (1398; 3611)	1597 (1145; 2684)	-618	-28	0.05
Protein, g	12.2 (8.4; 15.7)	9.0 (4.9; 13.3)	-3.2	-26	0.02[d]
Fat, g	14.2 (7.9; 27.6)	10.1 (6.5; 12.0)	-4.1	-29	0.04[d]
Iron, mg	3.3 (2.1; 5.2)	2.4 (1.6; 3.7)	-0.9	-27	0.10
Zinc, mg	2.2 (1.2; 3.3)	2.1 (1.1; 2.5)	-0.1	-5	0.57
Vitamin A, µg	198 (69; 460)	92 (11; 244)	-106	-54	0.33

[a] i-24-HR-II – i-24HR-I
[b] (i-24HR-II-i-24HR-I) / i-24HR-I × 100 for median intakes
[c] P-value of Wilcoxon matched-pair signed-rank test between i-24HR I and i-24HR II
[d] Significant at the 0.05-level
WFR, weighed food record; i-24HR-I, baseline interactive 24-hour recall; i-24HR-II, interactive 24-hour recall concerning the same day of food intake as WFR

6.3 *The relative validity of interactive 24-hour recall*

6.3.1 Extent of agreement between methods on a group basis

6.3.1.1 Wilcoxon matched-pair signed-rank test

Group median intakes of energy and nutrients measured by the 12-i-24HR and the WFR did not differ (Wilcoxon matched-pair signed-rank test $p > 0.05$), except protein intake (Wilcoxon matched-pair signed-rank test $p < 0.05$) (Table 9). When analyzed separately for the intervention and control group, the disagreement regarding protein intake was seen only in the intervention group (Annex 6). The likely cause of the disagreement was underestimated amounts of relishes with animal protein. Median differences between 12-i-24HR and WFR were < 20 %, except for vitamin A.

Table 9. Relative validity of the interactive 24-hour recall at the group level: median intakes of energy and selected nutrients between 6:00 hours and 18:00 hours measured by interactive 24-hour recall and weighed food record: comparison of median intakes and statistical significance of differences (n=44).

Dietary factor	WFR median (25^{th}, 75^{th} percentile)	12-i-24HR median (25^{th}, 75^{th} percentile)	Absolute difference[a]	% difference[b]	P-value[c]
Energy, kJ	1712 (1368; 2371)	1708 (1239; 2166)	-4	-0.2	0.38
Protein, g	9.4 (6.7; 13.2)	8.5 (6.0; 10.3)	-0.9	-10	0.019[d]
Fat, g	11.1 (6.2; 14.7)	10.2 (7.2; 15.0)	-0.9	-8	0.84
Iron, mg	3.6 (2.2; 6.3)	3.0 (2.0; 9.6)	-0.6	-17	0.82
Zinc, mg	3.1 (1.4; 5.1)	2.5 (1.3; 4.4)	-0.6	-19	0.12
Vitamin A, µ	129 (36; 663)	262 (40; 704)	133	103	0.90

[a] 12-i-24-HR – WFR
[b] (12-i-24HR-WFR) / WFR × 100 for group median intakes
[c] P-value of Wilcoxon matched-pair signed-rank test between WFR and 12-i-24HR
[d] Significant at the 0.05-level
WFR, weighed food record; 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours

6.3.1.2 Application of correction values developed by Thakwalakwa et al.

Correction values developed by Thakwalakwa and co-workers are applied on the median intake of energy and selected nutrients in Table 10. By comparing 12-i-24HR to predicted values, and the percentage differences presented in Table 10, it can be seen that values obtained by 12-i-24HR were closer to WFR than those adjusted with correction values. The only exception was vitamin A, for which the percentage difference was reduced from 103 % to 54 %.

Table 10. Correction values developed by Thakwalakwa et al. (2011) applied on the median intake estimates of energy and selected nutrients (n=44).

Dietary factor	Correction Value	12-i-24HR Median	Predicted value[a]	WFR median	% difference 1[b]	% difference 2[c]
Energy, kJ	0.86	1708	1469	1712	-0.2	-14
Protein, g	0.80	8.5	6.8	9.4	-10	-28
Fat, g	0.68	10.2	6.9	11.1	-8	-38
Iron, mg	0.69	3.0	2.1	3.6	-17	-42
Zinc, mg	0.72	2.5	1.8	3.1	-19	-42
Vitamin A, µg	0.76	262	199	129	103	54

[a] Correction value × 12-i-24HR
[b] (12-i-24HR-WFR) / WFR × 100 for group median intakes
[c] (Predicted value-WFR) / WFR × 100 for group median intakes
12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours; WFR, weighed food record

6.3.2 Extent of agreement between methods on an individual basis

6.3.2.1 Spearman rank correlation coefficient and partial correlation

In the intervention group Spearman rank correlation coefficient was at least 0.65 except for zinc (Table 11). The low coefficient for zinc intake may have been explained by four cases in which estimates on foods high in zinc (LNS and a local fish) were affected by reporting errors; amount of LNS was underestimated and fish relishes were reported as fish broth. When these four observations were removed, Spearman rank correlation coefficient for zinc rose to 0.60. Correlation coefficients had a range of 0.43 to 0.57 in the control group, all statistically significant. Correlation coefficient for intake of zinc measured by 12-i-24HR and WFR remained low (r = 0.26) when using partial correlation to combine the two subgroups. For energy and other selected nutrients than zinc, partial correlation ranged from 0.46 to 0.71.

Table 11. Relative validity of the interactive 24-hour recall at the individual level: correlation analysis: Spearman rank correlation coefficients between 12-hour interactive 24-hour recall and weighed food record in intervention and control groups, and partial correlation for all subjects.

	Spearman rank correlation coefficient		Partial correlation
	Intervention group	Control group	All
Dietary factor	n=21	n=23	n=44
Energy	0.65[a]	0.57[a]	0.65[a]
Protein	0.78[a]	0.56[a]	0.71[a]
Fat	0.74[a]	0.45[a]	0.56[a]
Iron	0.63[a]	0.45[a]	0.46[a]
Zinc	0.22	0.43[a]	0.26
Vitamin A	0.65[a]	0.50[a]	0.62[a]

[a] Significant at the 0.01 level

6.3.2.2 Cross-classification into tertiles

Forty-three to 62 % of subjects in the intervention group were correctly classified into the same third of intake for energy and five nutrients by the 12-i-24HR and the WFR (Table 12). For zinc, four out of 21 cases were grossly misclassified into the opposite third of intake. For protein, iron and vitamin A, one respondent was classified into the opposite tertile of intake. There were no grossly misclassified respondents for energy and fat intake. Likely reasons at the food level for the gross misclassifications included: intrusions (plain LNS), omissions (plain LNS), overestimated portion sizes (porridge, amount of LNS in porridge), and reporting relish as being consumed as broth (fish relish reported as fish broth in 12-i-24HR by two separate respondents).

The percentage of subject classified into the same and opposite categories of intake by the 12-i-24HR and the WFR ranged from 48 to 65 in the control group (Table 12). Not over two respondents were misclassified into the opposite tertile of intake. Likely reasons at the food level for the gross misclassifications included: intrusions (maize puffs, fish relish, papaya), omissions (relish with green leafy vegetables), underestimated portion sizes (several starchy staples in one 12-i-24HR underestimated), overestimated portion sizes (nsima, several starchy staples overestimated in one 12-i-24HR), mis-specified recipes (porridge made into milk but reported as made into water), and reporting fish relish as being consumed as broth.

Table 12. Relative validity of the interactive 24-hour recall in ranking individuals: frequency (percentage) of subjects classified into the same and into opposite nutrient intake tertiles by two dietary methods, the weighed food record and the 12-hour interactive 24-hour recall (n=44).

	Intervention group (n=21)		Control group (n=23)	
Dietary factor	Same tertile n (%)	Opposite tertile n (%)	Same tertile n (%)	Opposite tertile n (%)
Energy	12 (57)	0 (0)	13 (57)	2 (9)
Protein	11 (52)	1 (5)	12 (52)	1 (4)
Fat	10 (48)	0 (0)	11 (48)	2 (9)
Iron	13 (62)	1 (5)	12 (52)	1 (4)
Zinc	9 (43)	4 (19)	13 (57)	2 (9)
Vitamin A	12 (57)	1 (5)	15 (65)	2 (9)

6.3.2.3 Bland-Altman analysis

Bland-Altman plots showed that the mean difference in intake for energy and selected nutrients were very close to the zero bias line, apart from protein intake in the intervention group (Figures 6-17). Excluding protein intake in the intervention group, the observations were spread evenly on the both sides of the zero bias line representing mean difference in intake indicating that under- and overestimating energy and nutrient intakes were approximately equally common. Some nutrients seemed to have larger differences between 12-i-24HR and WFR at high levels of intake relative to low levels of intake. None of the observations deviated extremely from the limits of agreement. Seventeen out of 264 observations (6 variables × 44 participants) fell outside the limits of agreement. These 17 observations were attributable to 11 individuals' food intake. Not differentiating which nutrient is in question, reasons for observations being further than two standard deviations from the mean difference in intake included: intrusions (a portion of porridge containing LNS (n=3), cake (n=2), papaya (n=1)); omissions (plain LNS (n=4)); misestimated portion sizes (amount of sweet potato underestimated (n=2), amount of fish relish underestimated (n=2), amount of nsima overestimated (n=1)); and relish reported being consumed as broth (fish relish reported as fish broth (n=1); and discrepancy between tea recipe used in WFR and average recipe used in 12-i-24HR (n=1).

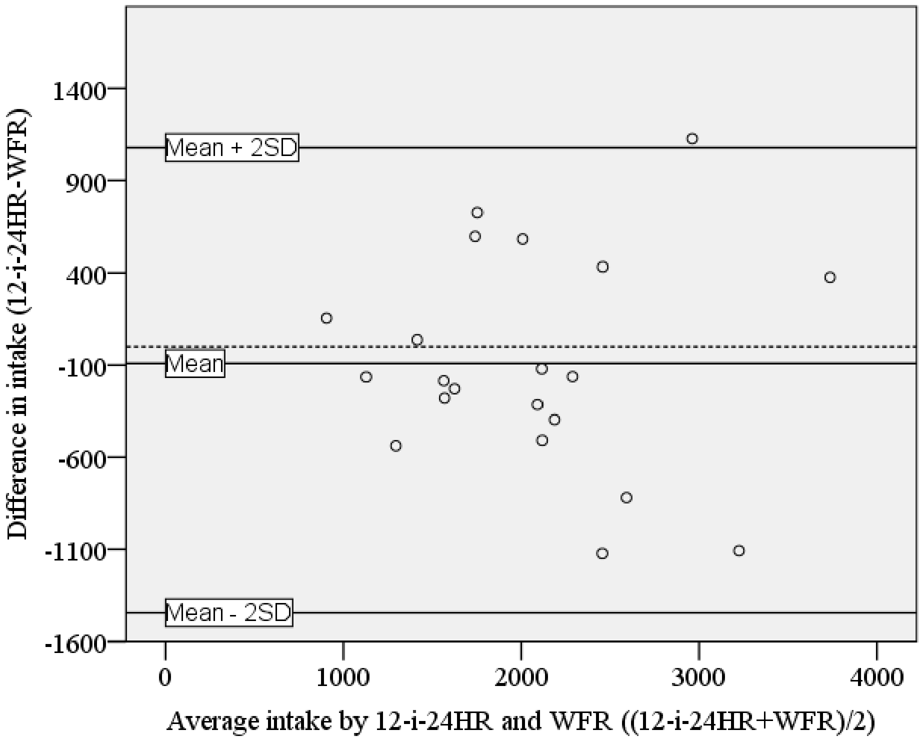

Figure 6. Energy intake (kJ): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21). Dotted line represents zero bias; solid lines represent ±2 SD from the mean. 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours; WFR, weighed food record; SD, standard deviation

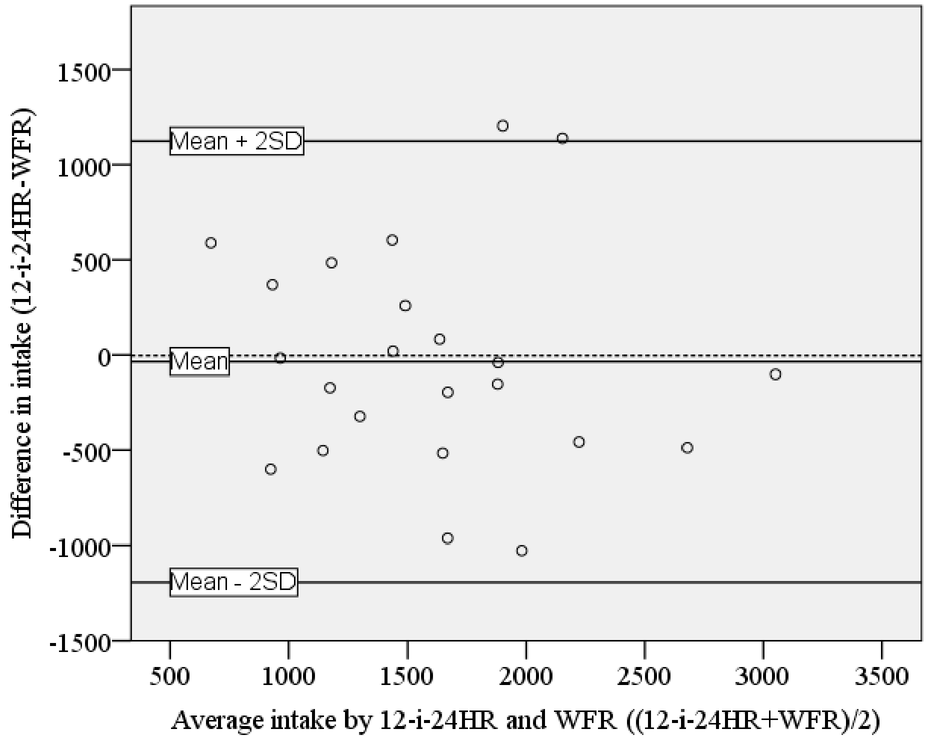

Figure 7. Energy intake (kJ): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=23).

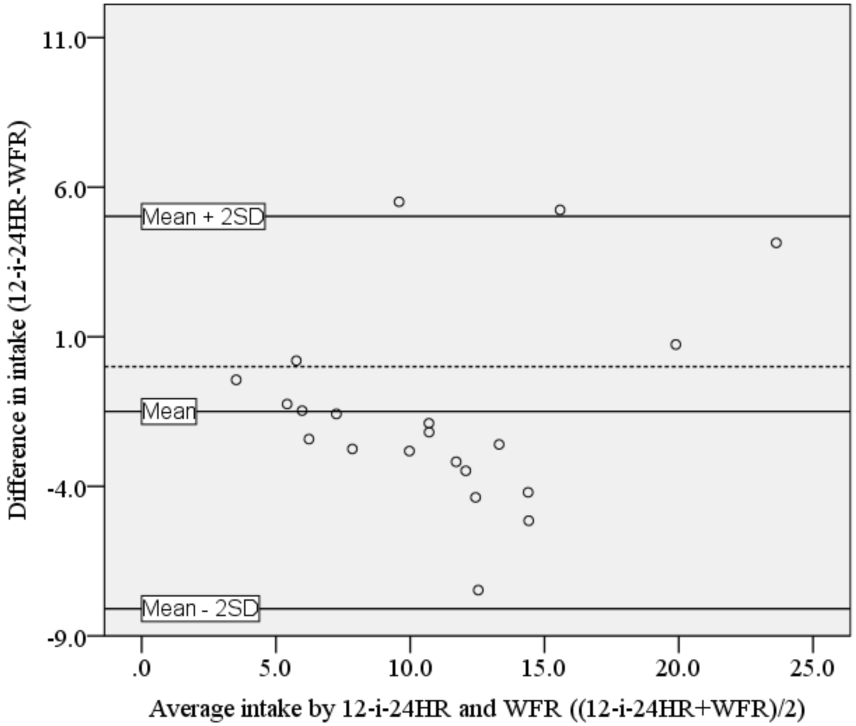

Figure 8. Protein intake (g): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21).

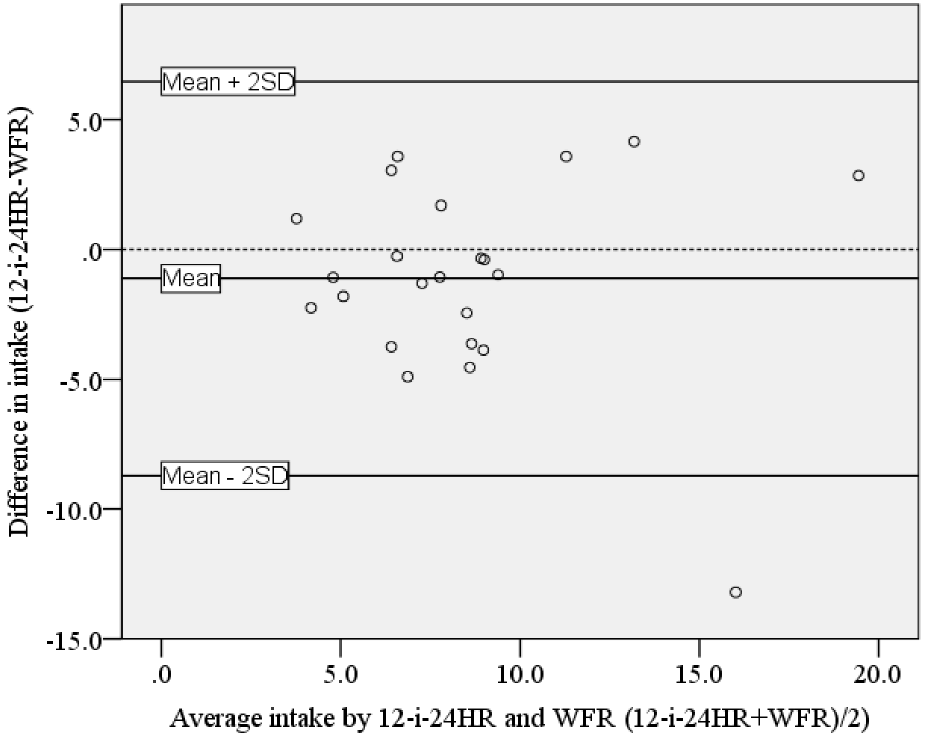

Figure 9. Protein intake (g): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=23).

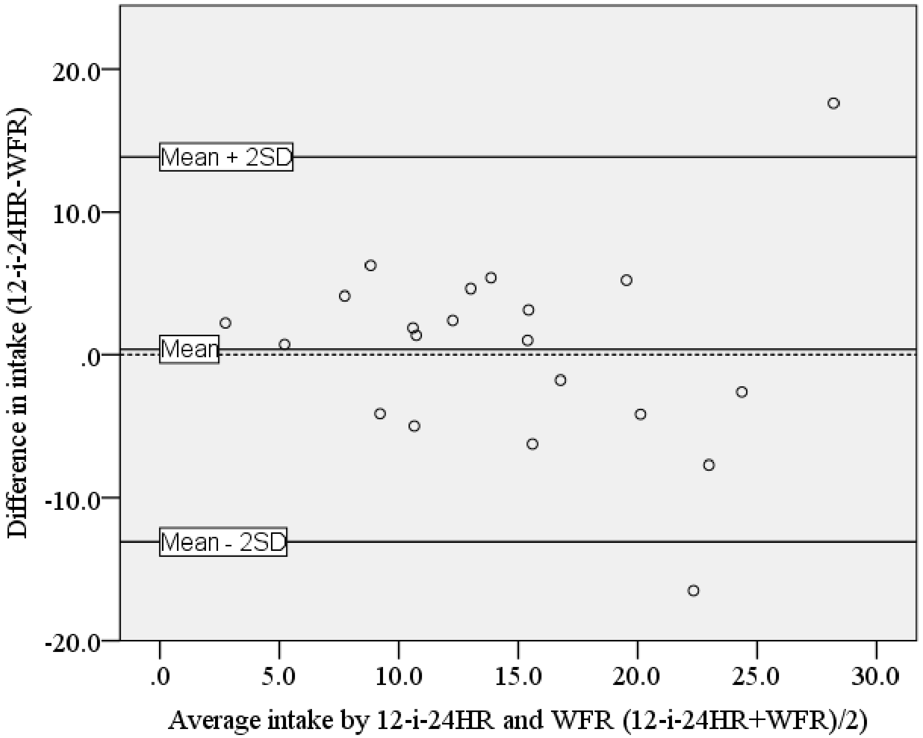

Figure 10. Fat intake (g): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21).

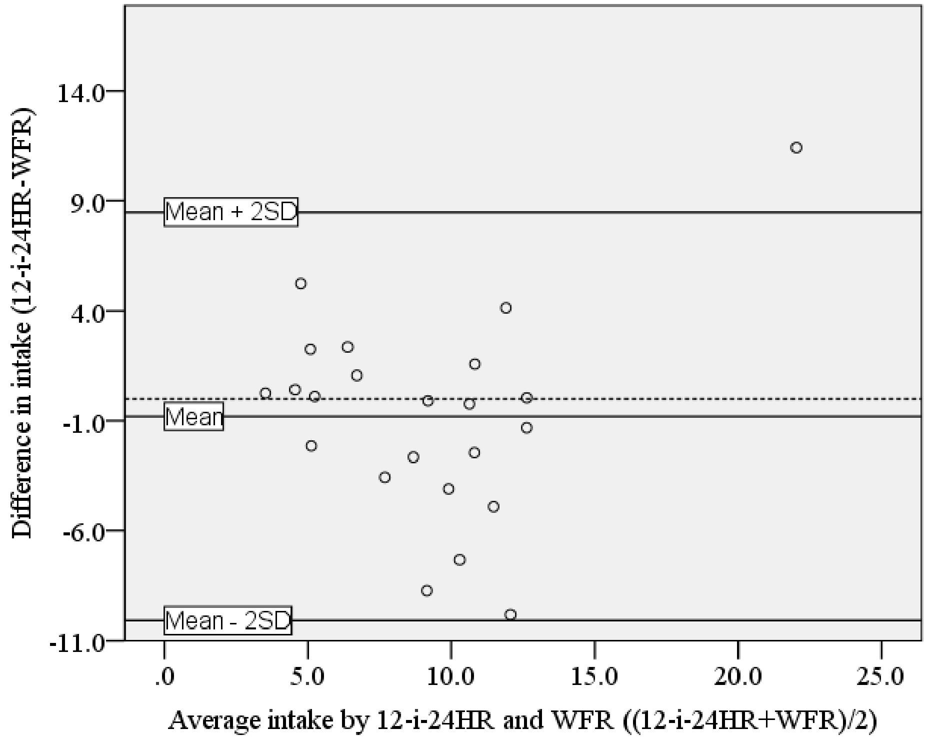

Figure 11. Fat intake (g): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=23).

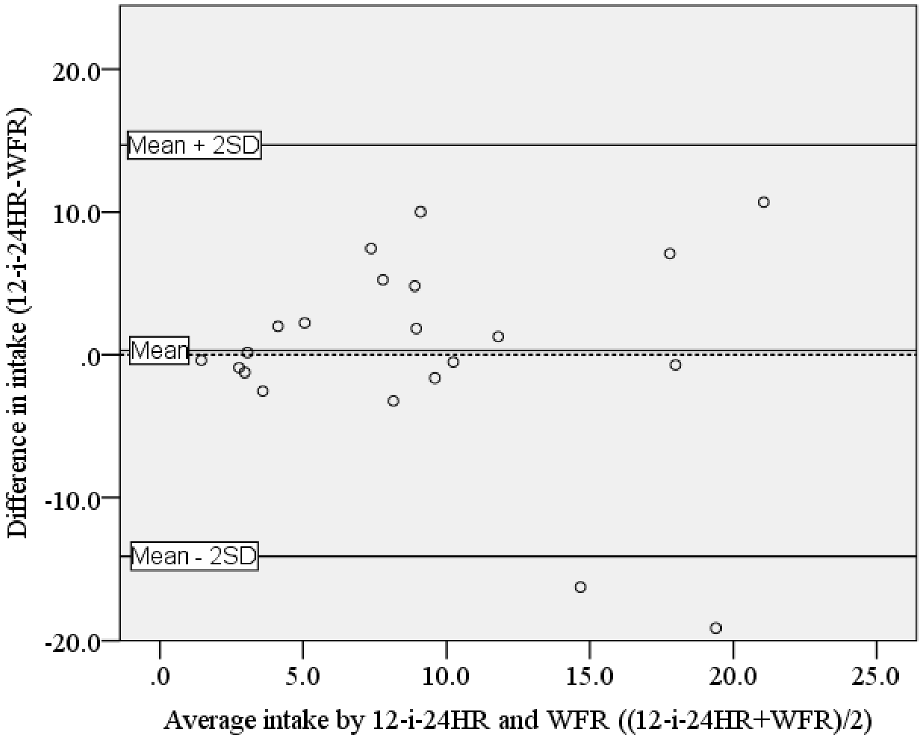

Figure 12. Iron intake (mg): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21).

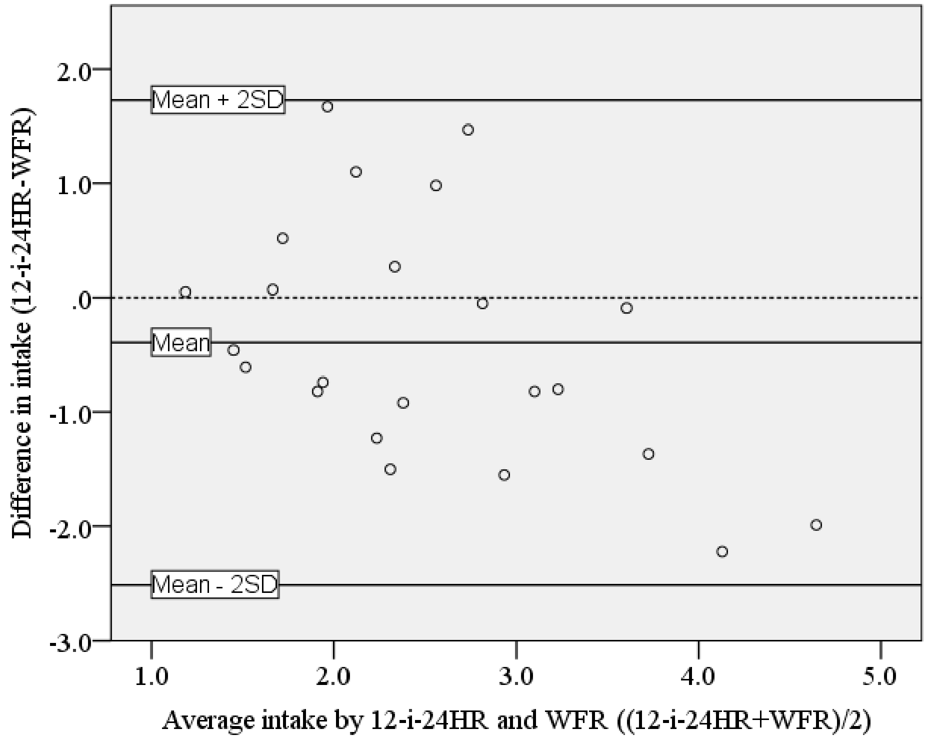

Figure 13. Iron intake (mg): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=23).

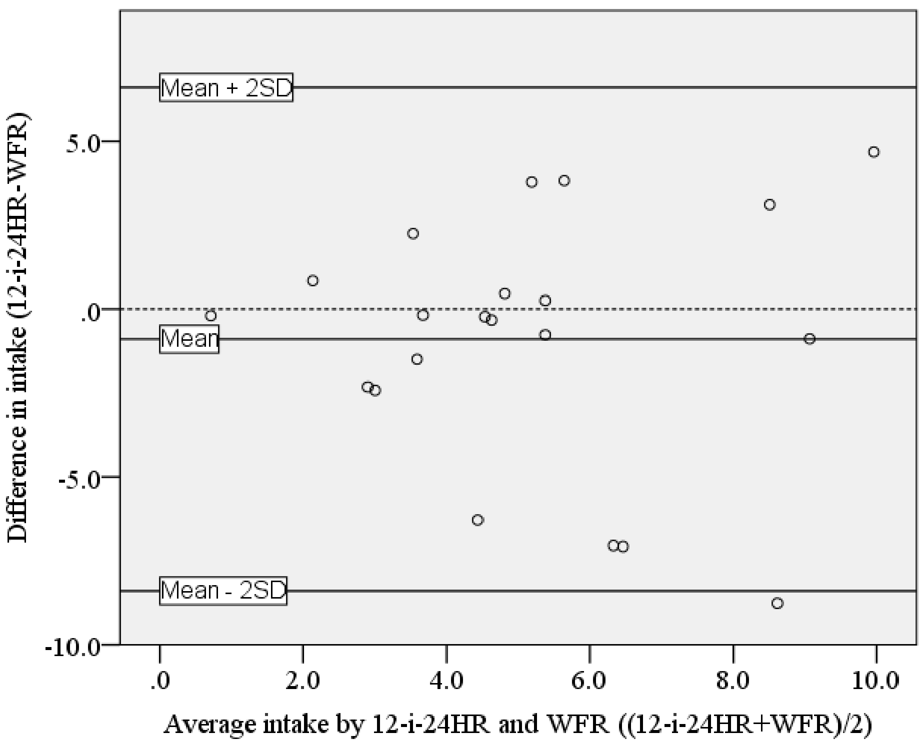

Figure 14. Zinc intake (mg): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21).

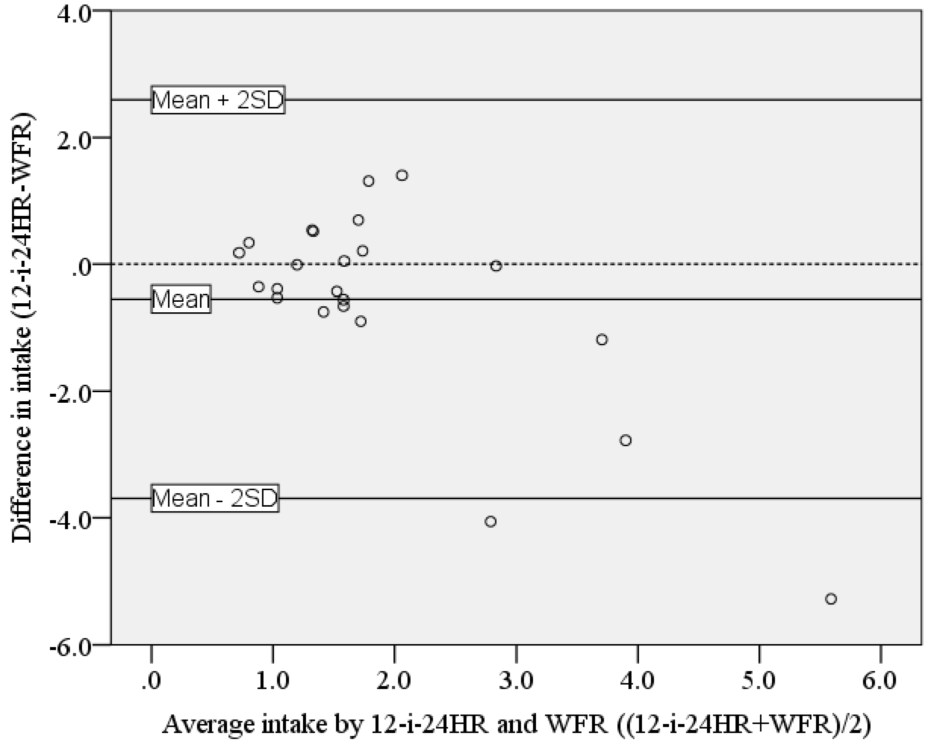

Figure 15. Zinc intake (mg): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=23).

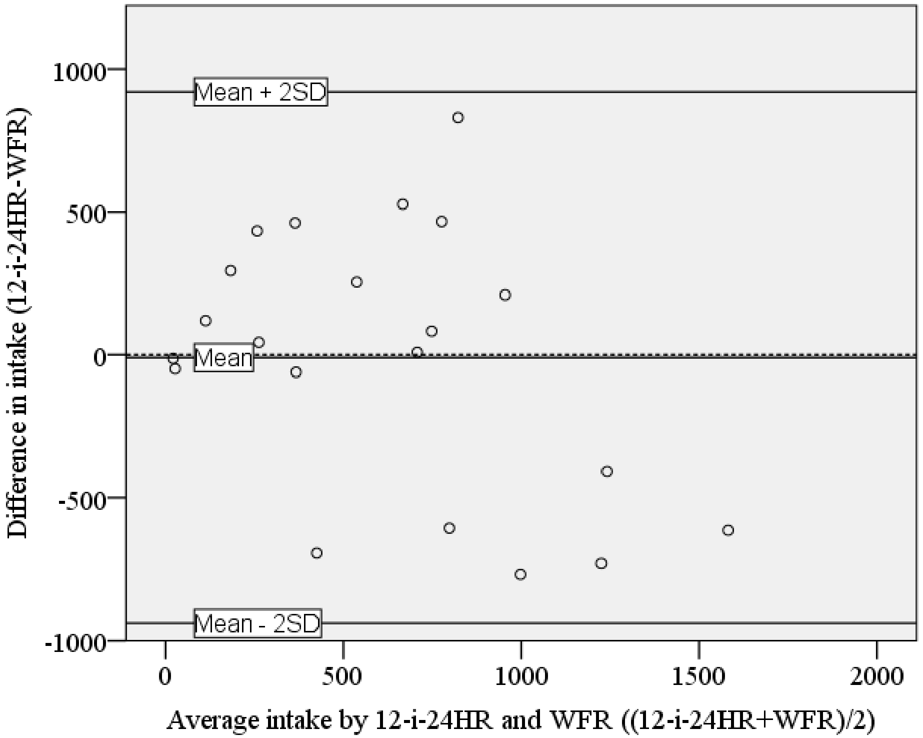

Figure 16. Vitamin A intake (µg): intervention group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=21).

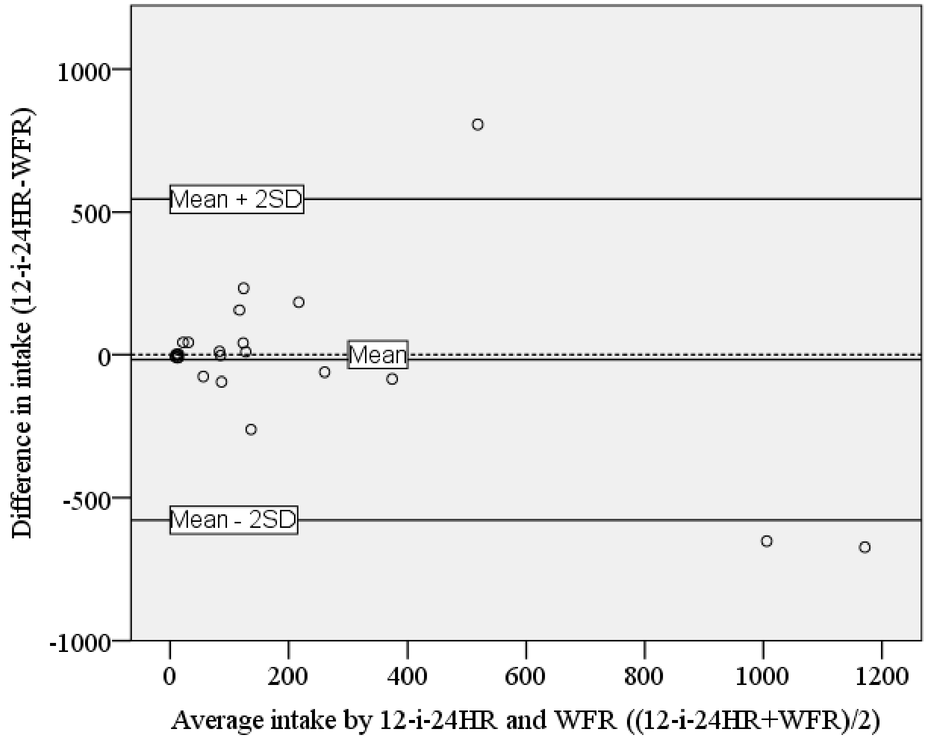

Figure 17. Vitamin A intake (µg): control group. Bland-Altman plot showing differences against mean daily intake estimated by WFR and 12-i-24HR (n=

6.3.3 Sources of measurement error

6.3.3.1 Intrusion and omission of food items, and portion size estimates

Foods divided into 12 categories are broken down into five categories of amounts in Table 14. Corresponding amounts from matches constituted only a part of reported amounts. Intrusions, omissions and incorrectly estimated portion sizes from matches were all sources of measurement error. Intrusions were more common than omissions (Table 13).

Table 13. Number of participants who were observed eating something from a food category once or more times during WFR, and number of respondents who omitted or added one or more food items from the 12-i-24HR.

	n		
Food category (g)	WFR[a]	Intrusions[b]	Omissions[c]
Porridge	40	3	1
Nsima, rice	42	2	0
Animal-origin relishes	14	2	0
Plant-origin relishes	24	6	4
Broths	5	6	0
Potato, cassava	14	5	0
Bread, cakes	14	13	5
Tea, soda, milk	30	7	1
Fruits	5	7	1
Sweets, biscuits	5	3	1
Puffs, popcorn, nuts	8	4	1
LNS	15	3	3

[a] Number of subjects who were observed consuming something from the food category once or more times during WFR
[b] Number of respondents who omitted one or more food items from the 12-i-24HR
[c] Number of respondents who added one or more food items to the 12-i-24HR
WFR, weighed food record; 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hoursLNS, lipid-based nutrient supplement
For terminology on reporting-error sensitive analysis, see Table 1 (page 27).

Table 14. Amounts (g) of food items classified into 12 food categories according to 12-hour interactive 24-hour recall, weighed food record, and five categories of amounts as median (25th, 75th percentile) intakes (n=44)[a].

Food category (g)	n[b]	12-i-24HR	WFR	Overreported g from intrusions (n)	Overreported g from matches	Corresponding g from matches	Unreported g from matches	Unreported g from omissions
Porridge	40	140 (94;192)	164 (112;206)	78 (73;276)	40 (11;66)	119 (89;175)	44 (23;83)	183 (183;183)
Nsima, rice	42	66 (45;118)	83 (50;155)	20 (14;25)	33 (18;55)	59 (33;93)	40 (17;57)	-
Animal-origin relishes	14	21 (15;33)	34 (23;62)	27 (20;33)	9 (3;26)	19 (14;31)	19 (13;51)	-
Plant-origin relishes	24	32 (17;65)	35 (18;78)	35 (15;66)	10 (3;18)	20 (16;54)	24 (8;56)	21 (13;66)
Broths	5	24 (20;37)	27 (18;30)	21 (7;28)	6 (5;7)	18 (17;29)	8 (3;12)	-
Potato, cassava	14	32 (19;62)	61 (25;104)	20 (15;43)	7 (1;35)	32 (23;54)	31 (26;45)	-
Bread, cakes	14	34 (23;41)	21 (15;34)	27 (17;36)	4 (2;30)	18 (13;33)	10 (6;42)	19 (6;23)
Tea, soda, milk	30	106 (67;184)	105 (66;164)	103 (27;185)	29 (19;62)	73 (55;118)	42 (18;79)	69 (69;69)
Fruits	5	54 (32;135)	71 (14;155)	50 (32;140)	24 (5;41)	24 (14;114)	29 (11;47)	72 (72;72)
Sweets, biscuits	5	17 (11;28)	6 (5;27)	15 (12;18)	5 (2;7)	9 (5;30)	6 (6;6)	5 (5;5)
Puffs, popcorn, nuts	8	10 (3;27)	17 (8;30)	13 (3;15)	7 (4;13)	10 (3;21)	11 (2;18)	13 (13;13)
LNS	15	19 (12;27)	16 (7;26)	17 (14;23)	14 (9;17)	10 (3;19)	7 (5;15)	14 (8;60)

[a] See Table 1 and Figure 2 for explanation
[b] Number of subjects who were observed consuming something from the food category once or more times during weighed food record
12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours, WFR, weighed food record; LNS, lipid-based nutrient supplement
For terminology on reporting-error sensitive analysis, see Table 1 (page 27).

Table 15 shows report rates, correspondence rates and inflation ratios for the 12 categories of foods. Report rates varied from 63 % for potato and cassava to 168 % for bread and cakes. Correspondence rates were lower than report rates. Median correspondence rate was lowest for animal-origin relishes (58 %) and highest for broths, fruits, and sweets and biscuits (100 %). Inflation ratios showed that food categories most affected by intrusions and overreported amounts from matches included plant-origin relishes, broths, bread and cakes, fruits, salty snacks (puffs, popcorn, nuts), and LNS. For three food categories with the highest number of consumers (nsima, porridge and tea, soda and milk), median correspondence rates ranged from 74 % to 94 %. Median inflation ratios were 0 % for these food categories.

Table 15. Report rates, correspondence rates and inflation ratios for amounts (g) of foods, given as median % (25th %, 75th % percentile) (n=44).

Food category	n[a]	Report rate[b] %	Correspondence rate[c] %	Inflation ratio[d] %
Porridge	39	94 (71;118)	84 (65;100)	0 (0;40)
Nsima, rice	42	82 (56;116)	74 (52;100)	0 (0;28)
Animal-origin relishes	14	65 (30;121)	58 (30;100)	0 (0;21)
Plant-origin relishes	24	89 (46;144)	79 (39;100)	9 (0;63)
Broths	5	125 (84;178)	100 (73;100)	25 (11;78)
Potato, cassava	14	63 (57;138)	62 (57;100)	0 (0;38)
Bread, cakes	14	168 (99;248)	74 (59;100)	45 (0;156)
Tea, soda, milk	30	94 (62;151)	94 (62;100)	0 (0;51)
Fruits	5	126 (39;280)	100 (38;100)	26 (1;180)
Sweets, biscuits	5	103 (75;386)	100 (75;100)	0 (0;193)
Puffs, popcorn, nuts	8	97 (83;129)	83 (33;100)	28 (0;50)
LNS	15	138 (69;236)	87 (64;100)	42 (0;186)

[a] Number of subjects who were observed consuming something from the food category once or more times during weighed food record visit
[b] (Reported amount/Reference amount) × 100
[c] (Corresponding amount/Reference amount) × 100
[d](Overreported amount/Reference amount) × 100
Reported amount = amount obtained by 12-i-24HR
Reference amount = amount obtained by WFR
Corresponding amount = for matches, overlap between reported g and reference g
Overreported amount = overreported amount from intrusions plus overreported amount from matches
Match = food item that is both in 12-i-24HR and WFR
Intrusion = food item appearing in 12-i-24HR but not observed in WFR
Overreported amount from intrusions = g from intrusions
Overreported amount from matches = for matches for which reported g > reference g, absolute differences between reported g and reference g
12-i-24HR = interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours
WFR = weighed food record
LNS = lipid-based nutrient supplement

6.3.3.2 Description of differential measurement error

A description of differential measurement error in energy and nutrient estimates based on 12-i-24HR are presented in Table 16. Differential measurement error was quantified by calculating factor A for energy and selected nutrients (White 2003). Values of factor A suggested that 12-i-24HR underestimated (0 < Factor A < 1) the true median difference in intake between intervention and control groups for energy and protein, and overestimated it (Factor A > 1) for fat, iron and vitamin A. Intake of zinc was not affected by differential measurement error (Factor A = 1). Overall, intake of energy, fat, iron and vitamin A seemed to be the more affected by differential measurement error than protein and zinc.

Table 16. Factor A as a description of differential measurement error in median intakes of energy and selected nutrients between the intervention group (n=21) and the control group (n=23).

	Energy	Protein	Fat	Iron	Zinc	Vitamin A
Factor A[a]	0.73	0.96	2.09	1.74	1.00	1.63

[a] the ratio of 12-i-24HR median difference in energy and nutrient intake between subjects in intervention and control groups to the WFR median difference in respective intakes, or:

$$Factor\ A = 1 + \frac{a-b}{c-d}$$

where

a = 12-i-24HR median intake minus WFR median intake for the intervention group

b = 12-i-24HR median intake minus WFR median intake for the control group

c = WFR median intake for the intervention group

d = WFR median intake for the control group

12-i-24HR, an interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours; WFR, weighed food record

7 Discussion of results

7.1 *Aims of the study and main findings*

The aim of the present study was to assess the relative validity of the 12-i-24HR, as compared with WFRs, in rural Malawian children aged 15 months. The relative validity of the 12-i-24HR was good at the group level, and moderate in estimating the intake of individuals. The relative validity of the 12-i-24HR in a group of children consuming LNS was affected by difficulties in assessing the consumption of animal-origin relishes and LNS accurately. Analyzing food consumption as a source of error showed that the 12-i-24HR performed better in assessing the intake of staple foods than in assessing foods that were consumed less frequently.

7.2 *Sample vs. target population*

7.2.1 Internal validity

There was a flaw in the study design whereby all participants did not go through the exact same protocol. The i-24HR-I was done as a baseline measurement before the WFR and the i-24HR-II. For the intervention group, 16 out of 21 completed the i-24HR-I and for the control group, 18 out of 23 completed all visits. Thus, some of the respondents were familiar with the method when the relative validity of the i-24HR-II was assessed while some respondents were not. The sample was, hence, not homogenous in this sense. This issue was not taken into account in the statistical analyses. Nutrient analyses also suggest that the i-24HR-I improved respondent's abilities in performing well in the i-24HR-II; intake of energy and nutrients were lower, and distribution in observations narrower in the i-24HR-II than in the i-24HR-I. Furthermore, picture charts of the i-24HR-II were more in line with the oral information compared with the i-24HR-I. These findings may also reflect true variability in food intake or the fact that the presence of RAs during the WFR made respondents more aware of their children's food intake. The differences in intake of energy and nutrients between i-24HR-I and i-24HR-II were more marked in the control group than in the intervention group, but this difference may have been attributable to an uneven distribution of certain characteristics in the study groups, due to the small number of participants. A stronger study design would have been to perform half of the i-24HR-Is before the WFR and half after the i-24HR-II, in order to remove the learning affect.

Chance was not involved in choosing the sample. Individuals were invited to participate in the study by HSAs. Hence, the sample may be biased. It is, however, difficult to come up with a reason that would have led to favoring some individuals over others. The control group represented the target population as such, i.e. children living in the chosen villages, and the intervention group was intended to vaguely represent the participants of the main trial, which it did. Participants of this study consumed LNS for a short period of time whereas the subjects of the main trial consume LNS for many months. The sample size of this study was determined by expenses of data collection, not by a calculation of sample size required to yield a certain power for statistical tests. Inter-group comparisons are limited by the small sample size.

In the present study 44 out of 68 randomized subjects (65 %) completed both the WFR and the i-24HR-II. Characteristics of respondents and those that dropped out can be different. It may be that those who withdrew from the study found it more difficult to provide the desired information than those who completed the entire study. Other researchers have reported higher response rates. Thakwalakwa et al. (2011) stated that the number of subjects was 169 in the beginning of their study. The reader is not informed about a nonresponse rate, implying that the response rate was 100 %. A very high response rate is reported also by Flax et al. (2010) whose dietary study entailed recruitment and one observational visit. For Flax and co-workers, 280 individuals were invited to participate and 176 were randomized. Out of randomized subjects 170 (97 %) completed the observational visit.

7.2.2 External validity

Due to the nature of the sample selection, the furthest generalization that can be made about the results dealing with the control group of this study extends to the villages in which the subjects of this study lived. Results dealing with the intervention group may be useful for trials that use products like LNS. However, even though other studies cannot fully justify using the i-24HR solely based on the results of this study, these results suggest the i-24HR may be a relatively valid method for assessing dietary intakes of 15-month old children in poor rural areas of Mangochi district during the post-harvest season, depending on the objective of a given study.

The reasons these results can likely be generalized to other 15-month old children living in rural areas of Malawi are that children being studied in this study were like their rural Malawian peers in some respects. The breastfeeding rate of this study (91 %) compared favorably to the national estimate of 16-month olds (92 %) (NSO Malawi and ORC Macro 2005). The diet of the children consisted of similar components as described by others carrying out research among young children in rural Malawi (Vaahtera et al. 2001, Flax et al. 2008, Hotz & Gibson 2008). Daily energy intake from complementary foods, as measured by the i-24HR-II, was 10 % and 30 % below the estimated requirements in the intervention and control groups, respectively (Dewey & Brown 2003). The median energy intake in the control group was comparable to a previous study of 16-month-old children from an area nearby (Thakwalakwa et al. 2011). Furthermore, the level of maternal education in rural Mangochi is probably not very different from other rural areas of Malawi. Primary education was the highest level of education for 20 % of women aged five years and above living in rural areas of Mangochi district, compared with an average of 25 % for rural women in the entire country (NSO Malawi 2008). Less than one percentage of women in rural Mangochi had secondary education while the average for all rural women in Malawi was 1.5 %.

7.3 *Relative validity of interactive 24-hour recall in this study*

7.3.1 Results of the present study

Results reported here showed that the 12-i-24HR estimated the average intakes of energy and nutrients for the group on the day of measurement to within a range of -0.2 % (energy) to -19 % (zinc) of their corresponding WFR intake estimates. Intake of vitamin A was affected by random errors; median difference between 12-i-24HR and WFR was 103 %. Yet, Wilcoxon matched-pair signed-rank test failed to detect a systematic bias. Protein intake in the intervention group was the only nutrient whose median intake was affected by a systematic bias. However, the degree of disagreement regarding protein intake in the intervention group was not extremely high since the percentage difference for protein was relatively low (-17 %) and the significance of the difference was close to the 0.05 limit ($p = 0.046$). Partial correlation coefficients for all subjects ranged from poor to good. For the control group, Spearman correlation coefficients were moderate. For the intervention group, Spearman correlation coefficients ranged from poor to good, depending on the nutrient. According to tertile classification agreement, the 12-i-24HR estimated individual intakes well in the control group, and moderately in the intervention group. Reporting weighed kappa values for tertile agreement classification is recommended to include agreement that can be accounted for by chance (Masson et al. 2002). In this study, however, the weighed kappa values were not calculated due to the small number of participants in the two subgroups. Inaccurate estimates of LNS and animal-origin relishes may have caused the discrepancies in the intervention group, although the small sample size may have also been a contributing factor.

7.3.2 Comparison with previous studies and recommendations

Previous relative validation studies have not supported the idea that the 24HR or the i-24HR can be used for assessing dietary intakes of young children in developing countries (Olinto et al. 1995, Thakwalakwa et al. 2011), except for one study that assessed intakes of energy and macronutrients (Dop et al. 1994). In the present study, median differences between 12-i-24HR and WFR in energy, protein and fat intakes were comparable with Dop et al. (1994) who conducted their study among 45 Senegalese weanlings (< 10 % vs. < 11 %). The i-24HR used in the present study had lower percentage differences between methods in assessing the intake of iron (-17 %) and zinc (-19 %) compared with Thakwalakwa et al. (2011) who reported percentage differences of 39 % and 28 % for iron and zinc, respectively. Thakwalakwa et al. had a lower percentage difference for vitamin A than the present study (34 % vs. 103 %) but in the present analysis the difference seemed to be caused by to random errors whereas for Thakwalakwa, there was a systematic bias. The sample size of the present study was one fourth of Thakwalakwa's sample.

In the present study, Spearman correlation coefficients in the intervention group were above the recommended level of 0.5 (see chapter 3.3.5), except for zinc (Masson et al. 2002). For the control group, Spearman correlation coefficients ranged from 0.43 to 0.57. Partial correlation for all subjects varied from 0.26 to 0.71. Dop et al. (1994) reported Spearman rank correlation coefficients of 0.6 or above for energy and macronutrients. In the present study intake of carbohydrates was not assessed, but partial correlation coefficients for energy, protein and fat ranged from 0.56 (fat) to 0.71 (protein). Thakwalakwa et al. (2011) reported intraclass correlation coefficients ranging from 0.42 (iron and zinc) to 0.83 (vitamin A) in their study where the i-24HR systematically overestimated intakes. Olinto et al. (1995) did not report correlation coefficients. In the current study, the control group passed the requirement suggested for tertile agreement (see chapter 3.3.5) whereas the intervention group did not (Masson et al. 2002). Those studies that are comparable to the present study have not used tertile agreement classification, or Bland-Altman plots.

7.4 *Secondary objectives of the study*

7.4.1 Memory lapses and inaccuracies in portion size estimates

A framework introduced by Smith and co-workers (2007) was used in this study in order to assess memory lapses and inaccuracies in portion size estimates thoroughly, and this analysis was complemented by examining outliers found in the aggregate nutrient analyses as well as making general observations during data handling. The reporting-error sensitive analysis showed that measurements made with 12-i-24HR were affected by omission and addition of food items, and inaccuracies in portion size estimates. The number and amount of staple foods (nsima, porridge) were well recalled. Adding food items to 12-i-24HR was more common than forgetting to report food items. It is possible that RAs performing the WFR have missed some food items while eating themselves, for instance.

A number of other errors related to reporting food consumption were noted during data handling and while examining outliers in nutrient analyses. Some respondents reported relishes as being consumed as broths, i.e. only the liquid part of the relish. In the present study, the amount of fish relish consumed was done using nsima for the flesh part of fish, fresh fish for whole fish relishes, and water for the soup of relish. For LNS, there were discrepancies in reporting the way of adding LNS to the porridge. LNS was either reported as added to the pot while cooking the porridge, or as added directly to the plate of the child. Both of these techniques were observed in the WFRs, but it was common that a different technique was reported in the 12-i-24HR compared with the respective WFR. One could assume that if the LNS has been added to a large pot while cooking the porridge, the amount of LNS in one portion of porridge may be smaller than if the LNS is added directly to the plate of the child and mixed with only that portion. Furthermore, mis-specified recipes, i.e. discrepancy between ingredients of mixed dishes reported in 12-i-24HR and in respective WFR, were another source of reporting error.

These observations should be taken into consideration when developing the i-24HR method further, at least if used in a similar population and similar setting as the present study. The current study strongly supports the use of actual foods as food models for staple foods. For relishes, RAs should be very neutral when asking about how the relish was consumed, and quantify the amount of fish relish using an actual fish relish as the food model. For LNS, RAs need to be as neutral as possible when asking about LNS consumption. It may be helpful if RAs are aware that such a reporting error may exist. In the present study, respondents were instructed to add the LNS on the plate of the child, and mix it with a small amount of porridge. Hence, respondents may find it difficult to report having used another technique than instructed. Getting good estimates of LNS consumption is especially important because, although consumed in small amounts, it is a very nutrient-dense product.

Dop et al. (1994) have conducted the only comparable study that has also reported data at the food level. They assessed the extent of reporting errors, but using different techniques. Dop et al. (1994) concluded that in their study, fish was the most often omitted food and that portion size estimates of rice and other foods from the household common pot measured as "handfuls" were the food categories most affected by reporting errors. In the present study estimates of fish relishes were attenuated by reporting errors, however, not by memory lapses but by underestimated portion sizes as there were no observations in the category "unreported amount from omissions".

Contradictory to the study of Dop et al. (1994), estimates on consumption of staple foods (nsima and rice, porridge) were good in the present study. Both report and correspondence rates were high and inflation ratio low for nsima and rice, and porridge. Some outliers in the nutrient analyses were caused by misreported consumption of these staple foods, but staple foods were also the most frequently consumed foods. This result suggests that using actual foods as food models is superior to using "handfuls" as a measure of quantity.

7.4.2 Differential measurement error

An important question, when evaluating the relative validity of a dietary assessment method for estimating inter-group differences in energy or nutrient intakes between an intervention and control group, is whether there is a differential bias. In the current study, there was particular interest in this question because the use of LNS in the intervention but not in the control group could introduce a differential bias. Differential measurement error was described as factor A in the present study. Factor A varied between 0.73 and 2.09 for energy and the selected nutrients. Not many studies have reported factor A. White (2003) presented an example of factor A based on a nested case-control study of breast-cancer where diet was assessed using two questionnaires about the same period of time. The questionnaire being tested was filled retrospectively in 1989 and the reference method was a questionnaire filled prospectively in 1986. The initial study was conducted by Giovannucci et al (1993). In White's example, factor A for fibre intake was 0.2.

The results noted here imply that especially fat, iron and vitamin A intake could have been affected by differential measurement error (factor A > 1), in the direction of overestimating the true mean difference between intervention and control groups. On the other hand, factor A for energy and protein pointed at the direction of underestimation. These inconsistencies indicate there was not a differential bias related to reported LNS intakes because inaccurate estimates of LNS consumption would bias fat, energy, protein and micronutrient intakes in the same direction. It should be noted, however, that calculations of factor A reported here did not include estimates on statistical significance. Hence, they should only be interpreted as descriptive information. Furthermore, as the assessments of relative validity were not adjusted for any confounding variables, it might be that variability in the values of factor A actually reflect variation in the subjects' and respondent's characteristics.

7.4.3 Correction values developed by Thakwalakwa et al.

Thakwalakwa's (2011) correction values were not helpful in this study since crude estimates based on the 12-i-24HR were in agreement with the WFR at the group level. This study does not support using the correction values among rural Malawian children aged 15 months. However, the exact i-24HR protocols used in this study and in Thakwalakwa's study were different. The method used in the present study was slightly more developed. In both studies, picture charts were provided two days before i-24HR, and foods and drinks were listed before asking more detailed information during i-24HR. However, Thakwalakwa et al. did not provide cups or bowls for feeding the children prior to the i-24HR, and they may have used fewer salted food models than what was used in the present study. The current study utilized 16 fresh and salted food models while Thakwalakwa et al. reported using "some" salted food models. Furthermore, Thakwalakwa et al. quantified foods using household measures while in the current study portion sizes were estimated by direct weighing of food models. The differences in i-24HR protocols may explain differences in the results. On the other hand, the relative validity of a dietary method is affected by the population it is applied to. Hence, differences in the respondents' characteristics may also have resulted in the noted differences in the relative validity of the i-24HRs.

7.5 *Strengths and limitations*

The study design by itself is the most important limitation in studies that use the WFR as the reference method. A day during which a RA is present at the subject's home does not represent a usual day in the life of the family. Performing the WFR may have made respondents more aware of their children's eating habits or even altered household dietary behavior. The problem of measuring food intake of free-living individuals is, however, common to all dietary studies. Furthermore, the i-24HR creates disruption to the everyday life of the family because participants eat from special cups and bowls during the day of interest, and a picture chart is filled. Statistically, agreement between the i-24HR and the WFR can be overestimated if performing the WFR or using cups and bowls and filling a picture chart sensitized respondents to the types and quantities of foods consumed by the child. Another unavoidable limitation in this study was the use of surrogate respondents. Surrogate respondents may bring an element of error to measurement because they do not necessarily observe everything that the index subjects eat.

The WFR covered the light hours and left out food consumption after 18:00 hours. Hence assessing the relative validity of the 12-i-24HR extends only to the light hours. An i-24HR may perform differently during the night if respondents' visual impression for foods differs according to the time of the day. However, it is likely that most eating occasions take place during the light hours and dark hours are of less importance. Previous studies have used similar approaches as the present study in terms of duration of the WFR (Gewa, Murphy & Neumann 2009, Thakwalakwa et al. 2011). A limitation of the current study is that the relative validity of the 12-i-24HR for meals consumed in the dark was not assessed. Furthermore, the relative validity of the i-24HR was done during only one season of the year.

To determine how generalizable the results of a study are, it is necessary to collect information about confounding factors. This study did not collect any information about the subjects' families' socio-economical status, education, food security situation, access to health care or other relevant variables. In the present study, access to healthcare could be considered as good because the mothers of the participants agreed to come to the local health center for the recruitment session. Some researchers have measured the access to health care as the physical distance to the local health center (Kulmala et al. 2000).

In the present study, the i-24HR and WFR were repeated only once and conclusions were made about the degree of accuracy of the nutrient estimates obtained by the i-24HR. From the perspective of an epidemiological study, the degree of relative validity of the chosen dietary method is just one piece of information that researchers need to know when designing their study. The study design depends also on the intended use of dietary data and on the reproducibility of the method. Determining the mean nutrient intake of a group may not require as much efforts as does estimating usual intake of individuals.

There are several strengths to this investigation. Even though the number of subjects was relatively small, a number of significant results emerged. Furthermore, analyzing the food consumption data supported the aggregate nutrient analyses. The use of average recipes collected from the households participating in the study probably reduced errors in measurement when the alternative would have been to use national or regional recipes for calculating raw ingredient intakes from mixed dishes. Compared with previous studies, the use of actual foods as food models was an important strength. The local knowledge held by the members of the research group was a major benefit as the team has carried out child health-related research projects in Mangochi Disctrict for almost 15 years. The RAs also had considerable experience in performing dietary recalls prior to this study.

8 Conclusions

Implications for epidemiological research

Improving complementary feeding practices has a crucial role in preventing undernutrition. Assessing energy and nutrient intakes is needed for research and programmatic purposes. Given the low feasibility of using labor-intensive methods such as the WFR at the population level, the possibility of using less burdensome methods like i-24HR is advantageous. The results of the present investigation suggest that the i-24HR is a relatively valid method for estimating the average energy and nutrient intakes of rural 15-month old Malawian children at a group level, and it can be used to moderately estimate individual intakes and rank children into tertiles of energy and nutrient intakes; at least for the nutrients examined in the current study. To improve the i-24HR for use with the studied population, researchers should consider using fish relish as an actual food model. If LNS is used, RAs should pay special attention to asking about how LNS has been served to the children.

However, before using the i-24HR, these results suggest researchers should under-take pilot studies to adapt the i-24HR to new settings where information on actual intakes is of interest. In particular this will provide information on ways to improve portion size estimates and identify foods that may likely be omitted. In addition, the results of this study showed that randomized controlled trials should be aware of the possibility that a differential measurement error may exist between study groups, blurring the associations between diet and outcomes of interest.

Future research needs

The relative validity of i-24HR needs to be re-assessed, using a larger sample size, in new settings, for different age groups of infants and young children and at different seasons of the year to generalize its relative validity. Further, it would be very useful to conduct a study where the relative validity of i-24HR and the traditional 24HR were assessed in parallel against a 24-hour period of WFR and biomarkers. The assessment would ideally include a large number of nutrients. Information on socioeconomic factors and other issues related to food intake should be collected. This type of study would allow us to determine how important the modifications included in the protocol of i-24HR actually are, which is important information given the additional respondent burden, time and resources required to use the picture chart and deliver the cups and bowls.

9 References

Alasuutari P. Social theory and human reality. London: Sage; 2004.

Alemaheyehu AA, Abebe Y, Gibson RS. A 24-h recall does not provide a valid estimate of absolute nutrient intakes for rural women in southern Ethiopia. Nutrition 2011;27:919-24.

Andersen LF, Lande B, Arsky GH, Trygg K. Validation of a semi-quantitative food frequency questionnaire used among 12-month-old Norwegian infants. Eur J Clin Nutr 2003;57:881-8.

Andersen LF, Lande B, Trygg K, Hay G. Validation of a semi-quantitative food frequency questionnaire used among 2-year-old Norwegian infants. Public Health Nutr 2004;7:757-64.

Arthur PG, Hartmann PE, Smith M. Measurement of the milk intake of breastfed infants. J Pediatr Gastroenterol Nutr 1987;6:758–63.

Aubel J, Touré I, Diagne M. Senegalese grandmothers promote improved maternal and child nutrition practices: the guardians of tradition are not averse to change. Soc Sci Med 2004;59:945-59.

Aziz SA, Hussein L. Evaluation of vitamin B12 status in Egypt IV: food consumption patterns among lactating mothers and their impact on the intake of the vitamin. Int J Food Sci Nutr 2005;56:455-62.

Bailey KD. Methods of social research. New York: Free Press; 1978.

Basch CE, Shea S, Arliss R, Contento IR, Rips J, Gutin B et al. Validation of mothers' reports of dietary intake by four to seven year-old children. Am J Public Health 1990;80:1314-7.

Bates CJ, Thurnham DI, Bingham SA, Margetts BM, Nelson M. Biochemical markers of nutrient intake. In: Margetts BM, Nelson M, editors. Design concepts in nutritional epidemiology. 2nd ed. New York : Oxford University Press; 1997.

Beaton GH, Burema J, Ritenbaugh C. Errors in interpretation of dietary assessment. Am J Clin Nutr 1997;65:1100S-7S.

Bellach B, Kohlmeier L. Energy adjustment does not control for differential recall bias in nutritional epidemiology. J Clin Epidemiol 1998;51:393-8.

Bezner Kerr R. Contested knowledge and disputed practice: Maize and groundnut seeds and child feeding in northern Malawi. (PhD Dissertation), Ithaca, NY, USA: Department of Development Sociology, Cornell University; 2006.

Bezner Kerr R, Dakishoni L, Shumba L, Msachi R, Chirwa M. "We grandmothers know plenty": complementary feeding and the multifaceted role of grandmothers in Malawi. Soc Sci Med 2008;66:1095-105.

Bhutta ZA, Ahmed T, Black RE, Cousens S, Dewey K, Giugliani E et al. What works? Interventions for maternal and child undernutrition and survival. Lancet 2008;371:417-40.

Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M et al. Maternal and child undernutrition: global and regional exposures and health consequences. Lancet 2008;371:243-59.

Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.

Block G. A review of validations of dietary assessment methods. Am J Epidemiol 1982;115:492-505.

Briefel RR, McDowell MA, Alaimo K, Caughman CR, Bischof AL, Carroll MD et al. Total energy intake of the US population: the third National Health and Nutrition Examination Survey, 1988-1991. Am J Clin Nutr 1995;62:1072S–80S.

Brown KH, Dewey KG, Allen L. Complementary feeding of young children in developing countries: a review of current scientific knowledge. Geneva: World Health Organization; 1998.

Bryce J, Coitinho D, Darnton-Hill I, Pelletier D, Pinstrup-Andersen P. Maternal and child undernutrition: effective action at national level. Lancet 2008;371:510-26.

Burema J, Van Staveren WA, Feunekes GI. Guidelines for reports on validation studies. Eur J Clin Nutr 1995;49:932-33.

Buzzard M. 24-hour dietary recall and food record methods. In: Willett WC. Nutritional epidemiology. 2nd ed. New York: Oxford University Press;1998.

Carruth BR, Skinner JD. Feeding behaviors and other motor development in healthy children (2-24 months). J Am Coll Nutr 2002;21:88-96.

Carruth BR, Ziegler PJ, Gordon A, Hendricks K. Developmental milestones and self-feeding behaviors in infants and toddlers. J Am Diet Assoc 2004:104;S51-S56.

Cassidy CM. Walk a mile in my shoes: culturally sensitive food-habit research. Am J Clin Nutr 1994;59:190S-7S.

Collins S, Sadler K. Outpatient care for severely malnourished children in emergency relief programmes: a retrospective cohort study. Lancet 2002;360:1824-30.

da Costa TH, Haisma H, Wells JC, Mander AP, Whitehead RG, Bluck LJ. How much human milk do infants consume? Data from 12 countries using a standardized stable isotope methodology. J Nutr. 2010;140:2227-32.

Coward WA, Cole TJ, Sawyer MB, Prentice AM. Breast-milk intake measurement in mixed-fed infants by administration of deuterium oxide to their mothers. Hum Nutr Clin Nutr 1982;36:141-8.

Dewey KG, Brown KH. Update on technical issues concerning complementary feeding of young children in developing countries and implications for intervention programs. Food Nutr Bull 2003;24:5-28.

Diop el HI, Dossou NI, Ndour MM, Briend A, Wade S. Comparison of the efficacy of a solid ready-to-use food and a liquid, milk-based diet for the rehabilitation of severely malnourished children: a randomized trial. Am J Clin Nutr 2003;78:302-7.

Dop MC, Milan C, Milan C, N'Diaye AM. The 24-hour recall for Senegalese weanlings: a validation exercise. Eur J Clin Nutr 1994;48:643-53.

Eck LH, Klesges RC, Hanson CL. Recall of a child's intake from one meal: are parents accurate? J Am Diet Assoc 1989;89:784-9.

Ferguson EL, Gadowsky SL, Huddle JM, Cullinan TR, Lehrfeld J, Gibson RS. An interactive 24-h recall technique for assessing the adequacy of trace mineral intakes of rural Malawian women; its advantages and limitations. Eur J Clin Nutr 1995;49:565-78.

Ferguson EL, Gibson RS & Opare-Obisaw C. The relative validity of the repeated 24 h recall for estimating energy and selected nutrient intakes of rural Ghanaian children. Eur J Clin Nutr 1994;48:241-52.

Ferguson EL, Gibson RS, Opare-Obisaw C, Osei-Opare F, Lamba C, Ounpuu S. Seasonal food consumption patterns and dietary diversity of rural preschool Ghanaian and Malawian children. Ecol Food Nutr 1993;29:219-34.

Ferguson EL, Gibson RS, Ounpuu S, Sabry JH. The validity of the 24 hour recall for estimating the energy and selected nutrient intakes of a group of rural Malawian preschool children. Ecol Food Nutr 1989;23:273-85.

Fisher JO, Butte NF, Mendoza PM, Wilson TA, Hodges EA, Reidy KC et al. Overestimation of infant and toddler energy intake by 24-h recall compared with weighed food records. Am J Clin Nutr 2008;88:407-15.

Flax VL, Ashorn U, Phuka JC, Maleta K, Manary MJ, Ashorn P.. Feeding patterns of underweight children in rural Malawi given supplementary fortified spread at home. Matern Child Nutr 2008;4:65-73.

Forrestal SG. Energy intake misreporting among children and adolescents: a literature review. Matern Child Nutr 2011;7:112-27.

Fouts HN, Brookshire RA. Who feeds children? A child's-eye-view of caregiver feeding patterns among the Aka foragers in Congo. Soc Sci Med 2009;69:285-92.

Fowler FJ, Mangione TW. Standardized Survey-Interviewing: Minimizing Interviewer-related Error. Thousand Oaks, CA: Sage Publications; 1990.

Gardiner HW, Kosmitzki C. Lives across cultures: Cross-cultural human development. Boston: Allyn & Bacon; 2005.

Gewa CA, Murphy SP, Neumann CG. Out-of-home food intake is often omitted from mothers' recalls of school children's intake in rural Kenya. J Nutr 2007;137:2154-9.

Gewa CA, Murphy SP, Neumann CG. A comparison of weighed and recalled intakes for schoolchildren and mothers in rural Kenya. Public Health Nutr 2009;12:1197-204.

Gibson RS. Nutritional assessment systems. In: Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press; 2005a.

Gibson RS. Validity in dietary assessment methods. In: Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press; 2005b.

Gibson RS. Measuring food consumption of individuals. In: Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press: 2005c.

Gibson RS. Measurement errors in dietary assessment. In Gibson RS. Principles of nutritional assessment. 2nd ed. New York: Oxford University Press: 2005d.

Gibson RS, Bailey KB, Gibbs M, Ferguson EL. A review of phytate, iron, zinc, and calcium concentrations in plant-based complementary foods used in low-income countries and implications for bioavailability. Food Nutr Bull 2010;31:S134-46.

Gibson RS, Ferguson EL. An interactive 24-hour recall for assessing the adequacy of iron and zinc intakes in developing countries. Washington, DC: International Life Sciences Institute; 2008.

Gibson RS, Ferguson EL, Lehrfeld J. Complementary foods for infant feeding in developing countries: their nutrient adequacy and improvement. Eur J Clin Nutr 1998;52:764-70.

Gibson RS, Huddle JM. Suboptimal zinc status in pregnant Malawian women: its association with low intakes of poorly available zinc, frequent reproductive cycling, and malaria. Am J Clin Nutr 1998;67:702-9.

Giovannucci E, Stampfer MJ, Colditz GA, Manson JE, Rosner BA, Longnecker M et al. A comparison of prospective and retrospective assessments of diet in the study of breast cancer. Am J Epidemiol 1993;137:502–11.

Guldan GS, Zeitlin MF, Beiser AS, Super CM, Gershoff SN, Datta S. Maternal education and child feeding practices in rural Bangladesh. Soc Sci Med 1993;36:925–35.

Harrison GG. Proceedings of the workshop on food-consumption surveys in developing countries: methodological considerations in descriptive food-consumption surveys in developing countries. Food Nutr Bull 2004;25:415-9.

Harrison GG, Galal OM, Ibrahim N, Khorshid A, Stormer A, Leslie J et al. Underreporting of food intake by dietary recall is not universal: a comparison of data from Egyptian and American women. J Nutr 2000;130:2049-54.

Harrington KF, Kohler CL, McLure LA, Franklin FA. Fourth graders' reports of fruit and vegetable intake at school lunch: does treatment assignment affect accuracy? J Am Diet Assoc 2009;109:36-44.

Hebert JR, Clemow L, Pbert L, Ockene IS, Ockene JK. Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. Int J Epidemiol 1995;24:389-98.

Heitmann BL, Lissner L. Dietary underreporting by obese individuals--is it specific or non-specific? BMJ 1995;311:986-9.

Hotz C, Gibson RS. Complementary feeding practices and dietary intakes from complementary foods amongst weanlings in rural Malawi. Eur J Clin Nutr 2001;55:841-9.

IAEA (International Atomic Energy Agency). Stable isotope technique to assess intake of human milk in breastfed infants. Vienna: International Atomic Energy Agency; 2010.

International lipid-based nutrient supplements project. www.ilins.org. [2011 Jun 23].

Kaaks RJ. Biochemical markers as additional measurements in studies of the accuracy of dietary questionnaire measurements: conceptual issues. Am J Clin Nutr l997;65:l232S-9S.

Kadzandira JM, Chilowa WR. The role of health surveillance assistants (HSAs) in the delivery of health services and immunization in Malawi. Lilongwe: University of Malawi; 2001.

Kant AK. Dietary patterns and health outcomes. J Am Diet Assoc 2004;104:615-35.

Kayongo-Male D & Onyango P. The sociology of the African family. London: Longran:1984.

Kersting M, Alexy U, Sichert-Hellert W, Manz F, Schöch G. Measured consumption of commercial infant food products in German infants: results from the DONALD study. J Pediatr Gastroenterol Nutr 1998;27:547–52.

Kigutha HN. Assessment of dietary intake in rural communities in Africa: experiences in Kenya. Am J Clin Nutr 1997;65:1168S-72S.

Kipnis V, Midthune D, Freedman L, Bingham S, Day NE, Riboli E et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutr 2002;5:915-23.

Klesges RC, Brown G, Frank CG. Validation of the 24-hour dietary recall in preschool children. J Am Diet Assoc 1987;87:1383-5.

Klesges RC, Eck LH, Ray JW. Who underreports dietary intake in a dietary recall? Evidence from the Second National Health and Nutrition Examination Survey. J Consult Clin Psychol 1995;63:438-44.

Kohlmeier L, Bellach B. Exposure assessment error and its handling in nutritional epidemiology. Annu Rev Public Health 1995;16:43-59.

Kruger R ja Gericke GJ. A qualitative exploration of rural feeding and weaning practices, knowledge and attitudes on nutrition. Public Health Nutr 2003;6(2):217-23.

Lafay L, Mennen L, Basdevant A, Charles MA, Borys JM, Eschwège E et al. Does energy intake underreporting involve all kinds of food or only specific food items? Results from the Fleurbaix Laventie Ville Sante (FLVS) study. Int J Obes Relat Metab Disord 2000;24:1500-6.

Lauer JA, Betrán AP, Victora CG, de Onís M, Barros AJ.. Breastfeeding patterns and exposure to suboptimal breastfeeding among children in developing countries: review and analysis of nationally representative surveys. BMC Med 2004;2:26.

Livingstone MBE, Robson PJ. Measurement of dietary intake in children. Proc Nutr Soc 2000;59:279-93.

Livingstone MBE, Robson PJ and Wallace JMW. Issues in dietary intake assessment of children and adolescents. Br J Nutr 2004; 92:S213–S222

Loria CM, McDowell MA, Johnson CL, Woteki CE. Nutrient data for Mexican-American foods: are current data adequate? J Am Diet Assoc 1991;91:919-22.

Lukmanji Z, Hertzmark, E, Mlingi N, Assey V, Ndossi G, Fawzi W. Tanzania food composition tables. 1st edition. Dar es Salaam: Muhimbili University of Health and Allied Sciences, Tanzania Food and Nutrition Centre & Harvard School of Public Health: 2008.

Macintyre S, Anderson A. Socio-demographic variables and psycho-social variables. In: Margetts BM, Nelson M (editors) Design concepts in nutritional epidemiology. 2nd ed. New York: Oxford University Press; 1997.

Madden JP, Goodman SJ, Guthrie HA. Validity of the 24-hr. recall. Analysis of data obtained from elderly subjects. J Am Diet Assoc 1976;68:143-7

Manary MJ, Ndekha MJ, Ashorn P, Maleta K, Briend A. Home based therapy for severe malnutrition with ready-to-use food. Arch Dis Child 2004;89:557-61.

Margetts BM, Thompson RL. Validation of dietary intake estimation. Eur J Clin Nutr 1995;49:934.

Masson LF, McNeill G, Tomany JO, Simpson JA, Peace HS, Wei L, Grubb DA et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire : use of correlation coefficients and the kappa statistic. Public Health Nutr 2002;6:313-21.

Maurer J, Taren DL, Teixeira PJ, Thomson CA, Lohman TG, Going SB et al. The psychosocial and behavioral characteristics related to energy misreporting. Nutr Rev 2006;64:53-66.

Miller TM, Abdel-Maksoud MF, Crane LA, Marcus AC, Byers TE.. Effects of social approval bias on self-reported fruit and vegetable consumption: a randomized controlled trial. Nutr J 2008;7:18.

Mitchikpe CE, Dossa RA, Ategbo EA, Van Raaij JM, Kok FJ. Seasonal variation in food pattern but not in energy and nutrient intakes of rural Beninese school-aged children. Public Health Nutr 2009;12:414-22.

Natarajan L, Pu M, Fan J, Levine RA, Patterson RE, Thomson CA et al. Measurement error of dietary self-report in intervention trials. Am J Epidemiol 2010;172:819-27.

National Statistical Office (NSO) of Malawi. 2008 Population and housing census report. http://www.nso.malawi.net. [2010 Oct 18].

National Statistical Office (NSO) of Malawi and ORC Macro. Malawi Demographic and Health Survey 2004. Calverton, Maryland; NSO and ORC Macro: 2005.

Ndekha M, Kulmala T, Vaahtera, M, Cullinana T, Salinae M, Ashorn P. Seasonal variation in the dietary sources of energy for pregnant women in Lungwena, rural Malawi. Ecol Food Nutr 2000;38:605-22.

Nelson M. The validation of dietary assessment. In: Margetts BM, Nelson M (editors). Design concepts in nutritional epidemiology. 2nd ed. New York: Oxford University Press; 1997.

Nyambose J, Koski KG, Tucker KL. High intra/interindividual variance ratios for energy and nutrient intakes of pregnant women in rural Malawi show that many days are required to estimate usual intake. J Nutr 2002;132:1313-8.

Olinto MT, Victora CG, Barros FC, Gigante DP. Twenty-four-hour recall overestimates the dietary intake of malnourished children. J Nutr 1995;125:880-4.

Onuorah CE, Ayo JA. Food taboos and their nutritional implications on developing nations like Nigeria - a review. Nutr Food Sci 2003;33 (5) 235-240.

PAHO (Pan American Health Organization). Guiding principles for complementary feeding of the breastfed child. Washington, DC: Pan American Health Organization; 2003.

Papalia DE, Olds SW, Feldman DR. Human development. 10th ed. New York: The Mgraw-Hill Companies Inc; 2007.

Paulhus DL, Duncan JH, Yik MSM. Patterns of shyness in East-Asian and European-heritage students. J Res Pers 2002;36:442-62.

Paxton A, Baxter SD, Fleming P, Ammerman A. Validation of the school lunch recall questionnaire to capture school lunch intake of third- to fifth-grade students. J Am Diet Assoc 2011;111:419-24.

Pelto GH, Levitt E, Thairu L. Improving feeding practices: current patterns, common constraints, and the design of interventions. Food Nutr Bull 2003;24:45-82.

Phuka JC, Maleta K, Thakwalakwa C, Cheung YB, Briend A, Manary MJ et al. Complementary feeding with fortified spread and incidence of severe stunting in 6- to 18-month-old rural Malawians. Arch Pediatr Adolesc Med 2008;162:619-26.

Piwoz EG, Creed de Kanashiro H, Lopez de Romana G, Black RE, Brown KH. Potential for misclassification of infants' usual feeding practices using 24-hour dietary assessment methods. J Nutr 1995;125:57-65.

Ruel MT, Brown KH, Caulfield LE. Moving towards with complementary feeding: Indicators and research priorities. International Food Policy Research Institute (IFPRI) Discussion Paper 146. Food Nutr Bull 2003;24:289-90.

Rutishauer IHE. Food intake studies in pre-school children in developing countries: problems of measurement and evaluation. Hum Nutr 1973;27:253-261. (ref. Ferguson et al. 1989; 1994)

Saha KK, Frongillo EA, Alam DS, Arifeen SE, Persson LÅ, Rasmussen KM. Household food security is associated with infant feeding practices in rural Bangladesh. J Nutr 2008;138:1383–90.

Sandige H, Ndekha MJ, Briend A, Ashorn P, Manary MJ. Home-based treatment of malnourished Malawian children with locally produced or imported ready-to-use food. J Pediatr Gastroenterol Nutr 2004;39:141-6.

Savenije OEM, Brand PLP. Accuracy and precision of test weighing to assess milk in newborn infants. Arch. Dis. Child. Fetal Neonatal Ed. 2006:91;F330–F332.

Scagliusi FB, Ferriolli E, Lancha AH Jr. Underreporting of energy intake in developing nations. Nutr Rev 2006;64:319-30.

Sharma S, Mbanya JC, Cruickshank K, Cade J, Tanya AK, Cao X et al. Nutritional composition of commonly consumed composite dishes from the Central Province of Cameroon. Int J Food Sci Nutr 2007;58:475-85.

Simoons FJ. Eat not this flesh: Food avoidances from prehistory to the present. 2nd ed. London: The University of Wisconsin Press; 1994.

Slimani N, Ferrari P, Ocké M, Welch A, Boeing H, Liere M et al. Standardization of the 24-hour diet recall calibration method used in the European prospective investigation into cancer and nutrition (EPIC): general concepts and preliminary results. Eur J Clin Nutr 2000;54:900-17.

Smith AF, Domel Baxter S, Hardin JW, Nichols MD. Conventional analyses of data from dietary validation studies may misestimate reporting accuracy: illustration from a study of the effect of interview modality on children's reporting accuracy. Public Health Nutr 2007;10:1247-56.

Solomons NW, Valdés-Ramos R. Dietary assessment tools for developing countries for use in multi-centric, collaborative tools. Public Health Nutr 2002;5:955-968.

Thakwalakwa CM, Kuusipalo HM, Maleta KM et al. The validity of a structured interactive 24h recall in estimating energy and nutrient intakes in 15-month old rural Malawian children. Matern Child Nutr 2011 Feb ebub ahead of print.

Thompson FE, Byers T. Dietary assessment resource manual. J Nutr 1994;124:2245S-2317S.

Thürigen D, Spiegelman D, Blettner M, Heuer C, Brenner H. Measurement error correction using validation data: a review of methods and their applicability in case-control studies. Stat Methods Med Res 2000;9:447-74.

UN (United Nations). A-RES-55-2 General Assembly Resolution 55/2: United Nations Millennium Declaration. New York: United Nations; 2000.

USDA (US Department of Agriculture). Continuing survey of food intake by individuals, 1994–96, 1998. Washington, DC: US Department of Agriculture; 2000.

Vaahtera M, Kulmala T, Hietanen A, Ndekha M, Cullinan T, Salin ML et al. Breastfeeding and complementary feeding practices in rural Malawi. Acta Pediatr 2001;90:328-32.

Victora CG, Adair L, Fall C, Hallal PC, Martorell R, Richter L et al. Maternal and child undernutrition: consequences for adult health and human capital. Lancet 2008;371:340-56.

Victora CG, de Onis M, Hallal PC, Blössner M, Shrimpton R. Worldwide timing of growth faltering: revisiting implications for interventions. Pediatrics 2010;125:e473-80.

White E. Design and interpretation of studies of differential exposure measurement error. Am J Epidemiol 2003;157:380-7.

White E, Armstrong BK, Saracci R. Exposure measurement error and its effects. In: White E, Armstrong BK, Saracci R. Principles of exposure measurement in epidemiology: collecting, evaluating, and improving measures of disease risk factors. 2nd ed. New York: Oxford University Press; 2008.

Wolff R. Meaning of food. Trop Geogr Med 1965;17:45-51

WHO (World Health Organization). Indicators for assessing breastfeeding practices. Geneva: World Health Organization; 1991.

WHO (World Health Organization). Global Strategy for Infant and Young Child Feeding. Geneva: World Health Organization; 2003.

WHO (World Health Organization). Indicators for assessing infant and young child feeding practices: Part I Measurement. Geneva: World Health Organization ; 2008.

WHO (World Health Organization). Indicators for assessing infant and young child feeding practices: Part II Measurement. Geneva: World Health Organization; 2010.

Willett WC. Correction for the effects of measurement error. In: Willett WC. Nutritional epidemiology. 2nd ed. New York: Oxford University Press;1998.

Winkvist A, Persson V, Hartini TNS. Underreporting of energy intake is less common among pregnant women in Indonesia. Public Health Nutr 2002;5:523-9.

10 Annexes

Annex 1

iLINS DOSE STUDY – PILOT STUDY OF 24-HOUR RECALL AND MINI VALIDATION USING WEIGHED RECORDS

LETTER OF INFORMATION FOR DIETARY INTAKE PILOT STUDY

Dear Mother, Father or Caregiver,

What is the reason for doing this study?

We are inviting you to participate in a pilot study through the University of Malawi, College of Medicine. This is a pilot (practice) study of part of a larger project, iLiNS, happening in the Mangochi and Namwera areas. The purpose of this study is to know about the effects of vitamins and minerals on the well-being and growth of infants and children. We hope that the results of the study will help us to find better ways to give infants and children in Malawi and other parts of the world the nutrients they need in order to grow well and remain healthy.

This product we are using in the study is called chiponde. It is made from soy or peanuts, milk powder, vegetable oil and vitamins. We would like to test whether this chiponde can help improve infant growth and health.

The purpose of this smaller study is to find out more about what your infant is eating using two different ways of asking about normal food intake (a weighed record and a 24-hour dietary recall). We will also be observing how your child (if your child is randomly selected to take the chiponde) is eating the chiponde.

You have been asked to participate because you have an infant who is between the age of 15-17 months and you live in the catchment area of the Mpondasi Health Centre.

What will happen if I agree to participate?

If you agree to participate, your child will be randomly assigned to one of two study groups – the control group, which will receive no additional study food, and the intervention group which will receive 40g of Lipid Based Spread every day for about three weeks. The reason for the control and the intervention group is so we can compare the food intake of infants without the LNS and with the LNS. During the pilot study, you will be visited by our research study assistants at your home (or wherever your food preparation takes place). There will be four visits in total – the first visit is to deliver equipment you will need to complete the study, it should only take about 15 minutes. The second visit is for just a short time – no more than 30 minutes, where the research assistant will do a short interview asking you to mention what your infant consumed the day before. The third visit is where, if you and your family agree, the research assistant will stay with you in your home. At the last visit, the research assistant will return for approximately half an hour to perform an interview.

In total, you will be asked to:

1. Be available for 3 weeks
2. You will be visited 4 times, one visit to deliver equipment, two visits lasting 30 minute each and one visit we would like to spend the whole day with you in your home. Our interviewers are well trained and will strive to minimize the disruption this causes by having a guest for the full day.

Page 1 of 4

3. If you are in the intervention group, you will be asked to feed your baby LNS for about 3 weeks.

What is the Schedule of the Study?

DAY 1-14 – LNS GROUP – This is where you will be asked to feed your baby the supplement for two weeks before we begin the visits. Otherwise you will continue feeding your baby normally.

Visit 1 – ALL PARTICIPANTS – The day before visit 1, a research assistant will visit your home quickly to deliver two items. 1) A Pictorial Chart – which will be used to mark off foods from various food groups that your infant consumed and 2) Bowls and Cups – which will be used to help you estimate how much your infant has eaten or drank by placing the food or drink in the bowl or cup before consumption.

Visit 2 – 24-hour recall from the previous day – The research assistant will visit you in your home or in a convenient place to do an interview about the type and amount of food your infant consumed the day before. The research assistant will use food models, pictures, and other ways of helping you to estimate how much your infant consumed of each type of food and drink. This should only last about 30 minutes.

Visit 3 – Weighed Record of Food Intake – The research assistant will come to your home at the beginning of the day before your infant has taken anything to eat or drink.

- We will ask to weigh everything before you give your child to eat or drink and we will weight what is left over. The purpose of this is so we have a good estimation of the amount of food and drink your baby is consuming.
- We will write down and weigh all the ingredients you use to prepare the food given to the child, so we can calculate how much of each ingredient the baby eats when the baby eats the cooked recipe.
- We will ask you to use the bowl and cup to use to serve food and drinks to your baby – this makes it easier for you to picture how much food the baby took.

Visit 4 – 24-hour recall from the previous day – The research assistant will visit you in your home or in a convenient place to do an interview about the type and amount of food your infant consumed the day before during the weighed record. This should only last about 30 minutes.

What will I benefit from participating?

There are no direct benefits from participating in this study. You will be contributing to our knowledge of how infants and children are fed and how they eat. This will contribute to a better understanding of nutrition, which may lead to helping more children attain better health. To compensate for your time and energy you invested during the study, you will get a bar of soap at each visit. The control group will be provided with a two-week supply of whole maize and ground nut flour phala after the study is over.

iLINS DOSE STUDY – PILOT STUDY OF 24-HOUR RECALL AND MINI VALIDATION USING WEIGHED RECORDS

Will my child be at risk if I participate in the study?

No. The study food is a well-test commercial product that is not expected to cause any adverse events. It has a good track record for being a safe and tolerable source of nutrition for infants. There are no risks to any member of your family by participating in this study.

Will my information be kept private?

Your privacy will be respected. All research materials in which you are identified will be kept confidential and will not be made publicly available. The questionnaires obtained from you will be labelled with a numbering system that identifies only your study number, but does not give away your identity. If the results of the study are published, your name will not be used and no information that discloses your identity will be published.

This information will be kept securely at the University of Malawi.

How long is the study? Will I be compensated for my time?

If you choose to participate in this study, you will be asked that a research participant come to your home for 1 full day, and 2 30-minute sessions. The only costs associated with this study are your time and inconvenience. We realize this requires a large amount of time to complete this study, so to compensate for your time your family will be provided with a bar of soap at each visit.

What will happen if I choose not to participate?
Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on your future care.

What if I decide to change my mind?
You can choose to stop the study at any point during the study without any effect on your care at any health facility.

Final Notes:

If you are participating in another nutrition study at this time, please inform our study team right away.

If, during the course of this study, new information becomes available that may relate to your willingness to continue to participate, this information will be provided to you by the study coordinators.

iLINS DOSE STUDY – PILOT STUDY OF 24-HOUR RECALL AND MINI VALIDATION USING WEIGHED RECORDS

Questions

We will be happy to answer any questions you have about this study. Please take time to discuss this study with your family members and friends to make the most informed decision possible about participating. If you have a question, concern or problem, you may contact the study Nutritionist (Chiza Kumwenda at 0888 378 093).

Mangochi, ________________, 2010

____

Chiza Kumwenda
Study Nutritionist, iLiNS Dose Project
0888 378 093

____

Olivia Jelenje
iLiNS Project Manager, iLiNS Dose Project
0888 851 642

____

Mr. Atusaye Mbisa
Deputy District Environmental Health Officer
0888 895 034

Annex 2

iLiNS Dose Trial — Participant Code: |__|__|__|__| {ChildNumber}

Form 23e. Individual Recipe Record, English

1. VISIT INFORMATION

1.1. Number (code) of visit: {NumberVisit}................|___|___|___|

1.2. Date of visit: {DateVisit}..........|___|___| |___|___| 20 |___|___|
dd mm yy

1.3. Respondent (relationship to the participating child): {Respondent} ☐ [1] Mother ☐ [2] Father ☐ [66] Other 1.3.1. Explain:____________ {SpecRespondent}

1.4. Scale Code: {CodeScale}................................|__|__|

1.5. Known weight object: {Scaleknown}................|__|__|

1.6. Observed weight of object: {Scaleobserved}.... |__|__|

2.0 Recipe name {NutRecName} ______			**2.0** Recipe name {NutRecName} ______		
2.1 Recipe number {NutRecNum}........ |__|__|__|			**2.1** Recipe number {NutRecNum}........ |__|__|__|		
3.0 Weight of Cooking Container (pot) {NutWtCookCon}........|__|__|__|__|g			**3.0** Weight of Cooking Container (pot) {NutWtCookCon}........|__|__|__|__|g		
4.0 Raw Ingredients: *(please list)*	*(please record the weight in grams of each raw ingredient + bowl/plate)*	*(please record the weight in grams of plate after the ingredient has been added to the recipe)*	**3.0** Raw Ingredients: *(please list)*	*(please record the weight in grams of each raw ingredient + bowl/plate)*	*(please record the weight in grams of plate after the ingredient has been added to the recipe)*
4.1.	|__|__|__|__|g	|__|__|__|__|g	4.1.	|__|__|__|__|g	|__|__|__|__|g
4.2.	|__|__|__|__|g	|__|__|__|__|g	4.2.	|__|__|__|__|g	|__|__|__|__|g
4.3.	|__|__|__|__|g	|__|__|__|__|g	4.3.	|__|__|__|__|g	|__|__|__|__|g
4.4.	|__|__|__|__|g	|__|__|__|__|g	4.4.	|__|__|__|__|g	|__|__|__|__|g
4.5.	|__|__|__|__|g	|__|__|__|__|g	4.5.	|__|__|__|__|g	|__|__|__|__|g
4.6.	|__|__|__|__|g	|__|__|__|__|g	4.6.	|__|__|__|__|g	|__|__|__|__|g
4.7.	|__|__|__|__|g	|__|__|__|__|g	4.7.	|__|__|__|__|g	|__|__|__|__|g
4.8.	|__|__|__|__|g	|__|__|__|__|g	4.8.	|__|__|__|__|g	|__|__|__|__|g
4.9.	|__|__|__|__|g	|__|__|__|__|g	4.9.	|__|__|__|__|g	|__|__|__|__|g
4.10.	|__|__|__|__|g	|__|__|__|__|g	4.10.	|__|__|__|__|g	|__|__|__|__|g
5.0 Time Cooking Started: hh/mm Time Cooking Finished: hh/mm	|__|__|:|__|__| |__|__|:|__|__|		**5.0** Time Cooking Started: hh/mm Time Cooking Finished: hh/mm	|__|__|:|__|__| |__|__|:|__|__|	
6.0 Weight after cooking: *(please weigh the cooking container together with the cooked food)*	|__|__|__|__|g		**6.0** Weight after cooking: *(please weigh the cooking container together with the cooked food)*	|__|__|__|__|g	
7.0 Volume after cooking: *(Please use a small calibrated container and measure a specific volume of food. Please only use the containers designated to you)*	**7.1** Capacity (in ml) of calibrated tool **7.2** Wt of empty calibrated tool **7.3** Wt of calibrated tool + food	|__|__|__|ml |__|__|__|g |__|__|__|g	**7.0** Volume after cooking: *(Please use a small calibrated container and measure a specific volume of food. Please only use the containers designated to you)*	**7.1** Capacity (in ml) of calibrated tool **7.2** Wt of empty calibrated tool **7.3** Wt of calibrated tool + food	|__|__|__|ml |__|__|__|g |__|__|__|g

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1 {NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__|__| ______	i.2. |__|__|__|__| ______	i.3. |__|__|__|__| ______	i.4. |__|__|__|__| ______
ii.Date	______	______	______	______

Page 1 of 2

iLiNS Dose Trial
Form 23e. Weighed Dietary Record, English

Participant Code: |__|__|__|__| {ChildNumber}

1. VISIT INFORMATION

1.1. Number (code) of visit: {NumberVisit}................|__|__|__|
1.2. Date of visit: {DateVisit}..........|__|__| |__|__| 20 |__|__|
dd mm yy
1.3. Respondent (relationship to the participating child): {Respondent}
☐ [1] Mother
☐ [2] Father
☐ [66] Other 1.3.1.
Explain:____________ {SpecRespondent}

1.4. Scale Code: {CodeScale}.. |__|__|
1.5. Known weight object:{Scaleknown}............................ |__|__|
1.6. Observed weight of object: {Scaleobserved}............... |__|__|
1.7. Wt of LNS Pot at Beginning of Day {NutWtLNSB4}... |__|__|__|g
1.8. Wt of LNS Pot at End of Day {NutWtLNSEnd}........... |__|__|__|g

2.0 Weighed Dietary Record

Time	Place Eaten	Who Fed this Food to Infant?	Food and Beverage	Ingredients, Brand	Food Preparation Method	Recipe No.	Amount Served + bowl/cup? (g)	Amount left over + bowl/cup (g)	Amount Eaten (g)
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			
						|__|__|			

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1{NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__| ____________	i.2. |__|__|__| ____________	i.3. |__|__|__| ____________	i.4. |__|__|__| ____________
ii.Date	____________	____________	____________	____________

Annex 3

iLiNS Dose Trial **Form 23b. Pictorial Chart, Chewa** (version: 2010-03-24):   Participant Code: |__|__|__|__| {ChildNumber}/ Visit Code: |__|__|__| {VisitNumber}

Food Group – Pictures and Words	Check each time its eaten (✓)	Food Group – Pictures and Words	Check each time its eaten (✓)
Nsima 		**Zamasamba** 	
Phala 		**Mandasi/ Chitumbua/Chikondamoyo** 	
Mbatata/Chinangwa 		**Zakumwa /Thobwa** 	
Nsinjiro/Mtedza 		**Zochokela ku Mkaka** 	
Za Nyemba 		**Shuga** 	
Nsomba 		**Mafuta ophikira/ Majarine** 	
Nyama/Nkhuku/ Mbewa 		**Maswiti** 	
Madzila 		**Ngumbi** 	
Zipatso 		**Chiponde** 	
Mkaka wa mwana monga Lactogen kapena NAN kapena Aspen 		**Zina** 	

Annex 4

iLiNS Dose Trial Participant Code: |__|__|__|__| {ChildNumber}

Form 23. 24-hour Interactive Dietary Recall Visit, English (version: 2010-06-21)

1. VISIT INFORMATION

1.1. Number (code) of visit: {NumberVisit}..........................|__|__|__| *(to be filled by iLiNS data monitor)*

1.2. Date of visit: {DateVisit}..|__|__| |__|__| 20 |__|__|
dd mo yy

1.3. Respondent (relationship to the participating child): {Respondent}
☐ [1] Mother
☐ [2] Father
☐ [66] Other 1.3.1. Explain:____________ {SpecRespondent}

1.4. Day Food Eaten {DayofWeek}.. ________________

1.5. Scale Code: {CodeScale}.. |__|__|

1.6. Known weight object: {Scaleknown}................................ |__|__|

1.7. Observed weight of object: {Scaleobserved}..................... |__|__|

I would like to ask you some questions about [NAME] and how and what he/she is eating. First I need to know whether yesterday was a normal day of eating patterns was.

2. DIETARY PATTERNS IN PAST 24-HOURS

2.1 Has your child eaten or drank anything at all in the past 24-hours? {NutIntakepast}.................
☐ [1] Yes
☐ [0] No

2.2 Was the food intake usual this past 24-hours? {NutIntakeUsual} ....
☐ [1] Yes
☐ [0] No

2.3 How was the food intake unusual? {NutUnusual}........................... ....
☐ [1] Feasting (> usual)
☐ [2] Fasting (< usual)
☐ [77] Not applicable
☐ [66] Other 2.3.1. Specify. ____________ {unusualspec}

2.4. Has your child been sick in the past 24-hours? {NutChildSick}...
☐ [1] Yes
☐ [0] No
☐ [99] Not Sure

2.5 What type of illness has your child experienced? {NutSickType}..
☐ [1] Diarrhea
☐ [2] Vomiting
☐ [3] Fever
☐ [66] Other 2.4.2. Specify. ________________ {specifysick}
☐ [77] Not applicable

2.6. Did sickness affect appetite? {NutSickAppet}..
☐ [1] Yes
☐ [0] No
☐ [77] Not applicable

2.6.1. How did it affect appetite? {NutAppHow}..
☐ [1] Increased
☐ [0] Decreased
☐ [77] Not applicable
☐ [99] Not Sure

2.7. Did your child breast feed yesterday? {NutBreastFed}..
☐ [1] Yes
☐ [0] No

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1{NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__|	i.2. |__|__|__|	i.3. |__|__|__|	i.4. |__|__|__|
ii.Date				

"I am now going to start the 24-hour dietary recall. Could you please tell me everything your child ate and drank yesterday including main meals, snacks, things shared by friends or other members of the community and anything eaten during the night. Remember by starting from the first thing your infant ate or drank in the morning and continuing through the day until the last thing he or she ate in the evening or night. It may also help you to remember if you think about activities you did yesterday starting in the morning until the end of the day.

Page 1 / 5

3. INTERACTIVE 24-HOUR RECALL (page 1 of 2) (Initial Wt of LNS Pot if Listed: |__|__|__|g)

Time	Place Eaten	Who Fed this Food to your Infant?	Food or Drink	Description / Brand / Ingredients / Additions	Food Preparation Method	Amount Served	Amount Left Over (wt of empty bowl/cup)	Units	Food Model Used	Normal Frequency Consumed
										Week

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1 {NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__| ________________	i.2. |__|__|__| ________________	i.3. |__|__|__| ________________	i.4. |__|__|__| ________________
ii.Date	________________	________________	________________	________________

Participant Code: |__|__|__|__| {ChildNumber}

3. INTERACTIVE 24-HOUR RECALL (page 2 of 2)

Time	Place Eaten	Who Fed this Food to your Infant?	Food or Drink	Description / Brand / Ingredients / Additions	Food Preparation Method	Amount Served	Amount Left Over	Units	Food Model Used	Normal Frequency Consumed Week

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1{NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__| ________________	i.2. |__|__|__| ________________	i.3. |__|__|__| ____________	i.4. |__|__|__| ________________
ii.Date	________________	________________	________________	________________

4.0 QUESTIONS ABOUT PICTORIAL CHART

4.1 ***After scanning through the pictorial chart, did you note any differences between the dietary recall and the pictorial chart?***

☐ [1] There was something on the **pictorial chart** NOT on **24-hr recall**. {NutPicRecall}
☐ [2] There was something on the **24-hr recall** NOT on **pictorial chart** {NutRecallPic}
☐ [3] No Difference {NutNoDiff}

4.1.1 ***When there is something on the pictorial chart, which is not appearing on the recall....*** Can you recall during the day yesterday whether this food which is recorded on the pictorial chart was actually eaten?

Food / Drink {NutListPic} (please write)	**What Time?** (hh:mm) {NutTimePic}	**Actually eaten?** {NutConfPic}
4.1.1.1	4.1.1.2|__|__|:|__|__|	4.1.1.3. ☐ [1] Yes ☐ [0] No
4.1.1.4	4.1.1.5|__|__|:|__|__|	4.1.1.6 ☐ [1] Yes ☐ [0] No
4.1.1.7.	4.1.1.8.|__|__|:|__|__|	4.1.1.9. ☐ [1] Yes ☐ [0] No
4.1.1.10.	4.1.1.11.|__|__|:|__|__|	4.1.1.12. ☐ [1] Yes ☐ [0] No
4.1.1.13.	4.1.1.14.|__|__|:|__|__|	4.1.1.15. ☐ [1] Yes ☐ [0] No

4.1.1 If eaten, please add to recall!

4.1.2 **When there is something on the recall, which is not appearing on the pictorial chart.....** *Can you recall during the day yesterday whether this food was actually eaten? Is the type and amount listed on the 24-hour recall correct?*

Food / Drink {NutListRecall} (please write)	**What Time?** (hh:mm) {NutTimeRecall}	**Actually eaten?** {NutConfRecall}
4.1.2.1.	4.1.2.2.|__|__|:|__|__|	4.1.2.3. ☐ [1] Yes ☐ [0] No
4.1.2.4.	4.1.2.5.|__|__|:|__|__|	4.1.2.6. ☐ [1] Yes ☐ [0] No
4.1.2.7.	4.1.2.8.|__|__|:|__|__|	4.1.2.9. ☐ [1] Yes ☐ [0] No
4.1.2.10.	4.1.2.11.|__|__|:|__|__|	4.1.2.12. ☐ [1] Yes ☐ [0] No

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1 {NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |__|__|__|__| ________________	i.2. |__|__|__|__| ________________	i.3. |__|__|__|__| ____________	i.4. |__|__|__|__| ________________
ii.Date	________________	________________	________________	________________

SUMMARY OF RECALL AND CONCLUDING QUESTIONS

Summarize the 24-hour recall including the items listed below:

- ✓ *Meals*
- ✓ *Snacks*
- ✓ *Study food if mentioned.*

5.1 Is this a good reflection of what your infant had to eat and drink yesterday? {NutSummary}

☐ [1] Yes
☐ [0] No
(If no, please summarize again and ask the caregiver to provide information about the foods or beverages eaten, which were not included in the recall.)

5.2 Did feeding your child from a separate bowl and cup (provided by the study) change the type or amount of food your infant consumed? {NutCupBowl}

☐ [1] Yes
☐ [0] No
☐ [99] Not Sure

5.3 Since you have listed LNS in your recall, is it possible for us to have a look at the pot and weigh the pot for our records?

5.3.1. Weight of Pot at Delivery (g) {NutLNSWtDeliv}	**5.3.2. Weight of Pot (s) in Progress (g) {NutLNSWtPotPro}**
|___|___|___|g	5.3.2.1. |___|___|___|g 5.3.2.2. |___|___|___|g *(only record if more than one in progress)* 5.3.2.3. |___|___|___|g
5.3.3. Number of Empty Pots (0-4) {NutLNSPotEmpty}	**5.3.4. Number of Full Pots & Pots in Progress (0-4){NutLNSPotsPro}**
|__|__|	|__|__|

	Collector {NutCollector}	Monitor {NutMonitor}	Entry_1{NutEntry_1}	Entry_2 {NutEntry_2}
i.Code / Sign.	i.1. |___|___|___| ________________	i.2. |___|___|___| ________________	i.3. |___|___|___| ________________	i.4. |___|___|___| ________________
ii.Date	________________	________________	________________	________________

Annex 5

Foods written in *italic* are considered as eaten after 18:00 hours.

Participant 1		
1. Banana	7:00	-
2. Porridge with LNS	8:00	7:00
3. Nsima	13:30	12:00
4. Fish relish	13:30	12:00
5. Porridge	16:13	-
6. Banana	-	14:00
7. Juice	-	16:00
8. Sweets	-	15:00
9. Porridge	-	18:00
10. Nsima	-	*20:00*
11. Fish broth	-	*20:00*
12. Mandasi	-	*20:00*
13. Masau	-	*20:00*

Participant 2		
1. Porridge	7:35	7:00
2. Papaya	-	10:00
3. Nsima	12:53	12:00
4. Relish with vegetables	12:53	12:00
5. Bread	-	15:00
6. Nsima	-	*19:00*
7. Fish relish	-	*19:00*

Participant 3		
1. Porridge	8:35	7:30
2. Tea	13:29	10:00
3. Sweet potato	13:29	10:00
4. Bread	-	10:00
5. Nsima	14:10	13:00
6. Fish relish	14:10	13:00
7. Goat meat relish	14:10	13:00
8. Vegetable relish	-	13:00
9. Biscuit	-	15:00
10. Tea	-	16:00
11. Sweet potato	-	16:00
12. Mandasi	-	16:00
13. Nsima	-	*19:00*
14. Fish relish	-	*19:00*
15. Relish with vegetables	-	*19:00*

Participant 4		
1. Porridge	7:20	7:00
2. Juice	9:46	-
3. Tea	10:36	10:00
4. Mandasi	10:36	10:00
5. Nsima	13:20	13:00
6. Relish with beans	13:20	-
7. Fish relish	-	13:00
8. Porridge	13:48	15:00
9. Mandasi	14:41	-
10. LNS	15:48	-
11. Nsima	-	*19:00*
12. Fish broth	-	*19:00*

Participant 5		
1. Porridge	7:55	7:00
2. Tea	-	8:00
3. Sweet potato	-	8:00
4. Biscuits	8:51	9:00
5. Milk	-	11:00
6. Nsima	13:11	13:00
7. Relish with vegetables	13:11	13:00
8. Mandasi	-	15:00
9. LNS	15:46	-
10. Nsima	17:54	17:00
11. Relish with vegetables	17:54	17:00

Participant 6		
1. Porridge	8:35	9:00
2. Tea	9:06	10:00
3. Sweet potato	9:06	10:00
4. Nsima	13:10	13:00
5. Relish with vegetables	13:10	13:00
6. Mandasi	-	18:00

Participant 7		
1. Tea	8:05	7:00
2. Bread	8:05	7:00
3. Papaya	10:15	9:00
4. Nsima	14:20	14:00
5. Fish relish	14:20	14:00
6. Fish relish	14:20	14:00

Participant 8		
1. Tea	-	7:00
2. Porridge	7:35	10:00
3. Sweet potato	9:02	7:00
4. Tea	9:02	-
5. Nsima	12:08	12:30
6. Fish relish	12:08	12:30
7. Nsima	-	*19:20*
8. Chicken broth	-	*19:20*

Participant 9		
1. Porridge with LNS	7:54	7:00
2. Tea	8:40	9:00
3. African cake	8:40	9:00
4. Nsima	14:00	13:30
5. Fish relish	14:00	13:30

Participant 10		
1. Porridge with LNS	7:11	7:00
2. Tea	9:44	8:30
3. Bread	9:44	8:30
4. Popcorn	12:26	15:00
5. Nsima	13:19	12:30
6. Relish with beans	13:19	12:30

Participant 11		
1. Porridge	7:19	8:00
2. Tea	8:30	9:00
3. Mandasi	-	9:00
4. Tea	-	12:00
5. Nsima	13:39	15:00
6. Broth	13:39	15:00
7. Porridge	-	18:00

Participant 12		
1. Porridge	9:01	8:00
2. Tea	-	9:00
3. Juice	11:43	13:30
4. Puffs	11:43	-
5. Banana	11:53	11:00
6. Tangerine	-	11:00
7. Nsima	14:04	13:00
8. Relish with beans	14:04	13:00
9. Relish with cabbage	14:04	13:00

10. Porridge	16:17	16:00
11. Tangerine	-	17:00
12. Nsima	-	*19:00*
13. Egg	-	*19:00*
14. Relish with vegetables	-	*19:00*

Participant 13		
1. Tea	-	6:00
2. Mandasi	-	8:00
3. Porridge with LNS	9:30	7:00
4. Mandasi	11:23	11:00
5. Porridge with LNS	-	12:00
6. Maize	11:50	-
7. Nsima	14:00	13:00
9. Relish with cabbage	-	13:00
10. Porridge with LNS	14:50	16:00
11. Mandasi	-	15:00
12. Fish relish	-	*19:00*

Participant 14		
1. Tea	8:03	8:30
2. Porridge	9:43	7:00
3. Banana	10:08	-
4. Banana	10:49	11:00
5. Mandasi	-	11:20
6. Puffs	12:31	14:20
7. Fish relish	14:10	12:00
8. Nsima	14:10	12:00
9. Banana	15:12	14:00
10. Banana	15:17	-
11. Mandasi	16:45	16:00
12. Cassava	-	16:00
13. Porridge	-	18:00
14. Cassava	-	*19:00*
15. Nsima	-	*20:00*
16. Fish relish	-	*20:00*

Participant 15		
1. Sweet potato	9:37	9:00
2. Tea	9:37	9:00
3. Nsima	14:05	13:00
4. Relish with vegetables	14:05	13:00
5. LNS	15:08	15:00
6. Sweet potato	16:59	17:00

7. Tea	16:59	17:00

Participant 16		
1. Porridge	8:45	7:00
2. Tea	9:00	8:00
3. Bread	-	8:00
4. Nsima	12:30	12:00
5. Fish relish	12:30	12:00
6. Bisuits	-	17:00
7. Nsima	-	*20:00*
8. Fish relish	-	*20:00*
9. Fanta	-	*23:00*

Participant 17		
1. Porridge	8:15	8:00
2. Puffs	8:40	8:15
3. Tea	9:00	10:00
4. Sweet potato	9:00	10:00
5. Nsima	14:15	13:50
6. Fish relish	14:15	13:50
7. Milk	15:50	16:00
8. Nsima	-	*19:00*
9. Fish relish	-	*19:00*

Participant 18		
1. Tea	-	9:00
2. Mandasi	-	9:00
3. Porridge with LNS	10:30	-
4. Bread	11:10	-
5. Sweet potato	11:10	9:00
6. Fish relish	14:52	13:00
7. Nsima	14:52	13:00
8. Porridge with LNS	17:58	18:00

Participant 19		
1. Porridge	8:05	8:00
2. Rice	10:15	10:00
3. Nsima	13:10	12:00
4. Relish with vegetables	13:10	12:00
5. Biscuit	15:00	15:00
6. Porridge	16:02	16:00

Participant 20		
1. Porridge with LNS	8:17	7:00

2. Tea	10:04	8:00
3. Sweet potato	10:04	8:00
4. Nsima	14:40	13:00
5. Relish with vegetables	14:40	13:30
6. Fish relish	-	13:00
7. Nsima	-	*19:00*
8. Relish with vegetables	-	*19:00*
9. Fish relish	-	*19:00*

Participant 21		
1. Porridge with LNS	8:40	7:00
2. Tea	9:19	8:00
3. Sweet potato	9:19	8:00
4. Nsima	12:00	12:00
5. Fish relish	12:00	12:00
6. Nsima	-	*19:00*
7. Fish relish	-	*19:00*

Participant 22		
1. Tea	9:03	8:00
2. Irish potato	9:03	8:00
3. Nsima	12:17	12:00
4. Fish relish	12:17	12:00
5. Porridge	16:20	16:00
6. Nsima	-	*19:00*
7. Fish relish	-	*19:00*

Participant 23		
1. Porridge	8:47	7:00
2. Tea	9:04	8:00
3. Bread	9:04	-
4. Nsima	13:43	14:00
5. Relish with beans	13:43	14:00
6. Nsima	-	*20:00*
7. Beans	-	*20:00*

Participant 24		
1. Puffs	7:30	7:00
2. Popcorn	7:35	no time
3. Tea	8:50	7:00
4. Sweet potato	8:50	7:00
5. Sweet potato	9:20	no time
6. Rice	12:43	14:00
7. Ground nuts	12:45	14:00

8. Porridge with LNS	17:05	no time

Participant 25		
1. Puffs	-	8:00
2. Porridge with LNS	8:37	8:00
3. African cake	-	9:00
4. Porridge	10:47	10:00
5. Tea	10:50	10:00
6. Nsima	14:48	19:10
7. Relish with cabbage	14:48	19:10
8. Porridge with LNS	16:00	16:20

Participant 26		
1. Porridge with LNS	8:14	8:00
2. Nsima	13:54	12:00
3. Fish relish	13:54	12:00
4. Relish with vegetables	-	12:00
5. Porridge with LNS	17:11	19:00

Participant 27		
1. Porridge	8:30	7:00
2. Rice	10:40	9:00
3. Tea	10:40	9:00
4. Beef relish	-	9:00
5. Relish with vegetables	-	9:00
6. Nsima	14:51	14:00
7. Relish with vegetables	14:51	14:00

Participant 28		
1. Porridge	8:48	7:00
2. Tea	10:44	8:00
3. Sweet potato	10:44	8:00
4. Nsima	14:30	13:00
5. Relish with cabbage	14:30	13:00

Participant 29		
1. Porridge	8:24	6:00
2. Rice	12:25	12:00
3. Tomato soup	12:28	12:00
4. Nsima	14:40	15:00
5. Relish with vegetables	14:40	-
6. Fish relish	-	15:00

Participant 30		
1. Sweet potato	-	6:00
2. Porridge	8:15	6:00
3. Sweets	10:50	-
4. African cake	-	10:00
5. Tea	11:22	10:00
6. Bread	11:22	10:00
7. Relish with vegetables	14:04	13:00
8. Nsima	14:04	13:00
9. Banana	-	16:00

Participant 31		
1. Porridge	8:45	8:00
2. Nsima	-	13:00
3. Relish with vegetables	-	13:00
4. Egg	-	13:00
5. Porridge	15:15	16:00
6. Nsima	-	*19:00*
7. Fish relish	-	*19:00*
8. Relish with beans	-	*19:00*

Participant 32		
1. Porridge	8:15	6:00
2. Tea	10:07	8:00
3. Sweet potato	10:07	8:00
4. Nsima	13:50	12:00
5. Relish with vegetables	13:50	12:00
6. Fish relish	-	12:00

Participant 33		
1. Mandasi	7:57	-
2. Tea	8:34	8:00
3. Mandasi	8:34	8:00
4. Nsima	-	11:00
5. Relish with vegetables	-	11:00
6. Tomato	12:04	-
7. Nsima	13:36	13:00
8. Fish relish	13:36	13:00
9. Relish with vegetables	13:36	13:00
10. Papaya	-	16:00
11. Porridge	17:37	18:00
12. Nsima	-	*19:30*
13. Fish relish	-	*19:30*

Participant 34		
1. Porridge	8:31	8:00
2. Tea	-	11:00
3. Bread	10:00	11:00
4. Fish relish	13:15	15:00
5. Nsima	13:15	15:00
6. Fish relish	18:10	18:00
7. Nsima	18:10	18:00

Participant 35		
1. Porridge	8:02	8:00
2. Cassava	8:02	10:00
3. Tea	10:30	10:00
4. Sweets	-	10:00
5. Nsima	12:51	12:00
6. Fish relish	12:51	12:00
7. Puffs	-	12:00
8. Papaya	-	13:00
9. Nsima	-	18:00
10. Fish relish	-	18:00

Participant 36		
1. Porridge	8:15	6:00
2. Sweet potato	-	6:00
3. Sweets	10:50	12:00
4. Tea	11:22	10:00
5. Bread	11:22	10:00
6. Mandasi	-	10:00
7. Green leafy vegetables	14:04	13:00
8. Nsima	14:04	13:00
9. Banana	-	16:00
10. Papaya	-	17:00
11. Sweets	-	17:00
12. Banana	-	*19:00*
13. Nsima	-	*19:00*
14. Fish relish	-	*19:00*
15. Relish with vegetables	-	*19:00*

Participant 37		
1. Porridge with LNS	8:04	7:00
2. Puffs	-	8:00
3. Banana	14:00	11:00
4. Nsima	12:30	12:00
5. Relish with beans	12:30	12:00

6. Relish with cabbage	12:30	12:00
7. Puffs	14:20	14:00
8. Tea	16:50	16:00
9. Sweet potato	16:50	16:00
10. Relish with beans	-	18:00
11. Relish with cabbage	-	18:00

Participant 38		
1. Porridge	8:35	8:00
2. Tea	9:00	9:00
3. Tea	10:00	11:00
4. Nsima	13:35	12:00
5. Relish with vegetables	13:35	12:00
6. Tea	-	13:00
7. Eggs	-	18:00
8. Rice	-	*18:30*

Participant 39		
1. Tea	9:58	8:00
2. Sweet potato	9:58	8:00
3. Nsima	12:50	12:00
4. Fish relish	12:50	12:00
5. LNS	-	15:00
6. Nsima	-	*19:00*
7. Relish with vegetables	-	*19:00*

Participant 40		
1. Tea	7:30	6:00
2. Bread	7:30	6:00
3. Porridge with LNS	9:40	8:00
4. Bread	11:30	-
5. Nsima	14:20	13:00
6. Relish with beans	14:20	13:00
7. Papaya	-	16:00
8. Nsima	-	*19:00*
9. Relish with beans	-	*19:00*

Participant 41		
1. African cake	9:14	8:00
2. Porridge	10:01	7:00
3. Papaya	11:23	9:00
4. Nsima	12:53	14:00
5. Fish relish	17:11	14:00
6. Ground nuts	17:01	9:30

7. Papaya	-	18:00
8. Nsima	-	*20:00*
9. Fish relish	-	*20:00*

Participant 42		
1. Porridge	9:31	7:00
2. Puffs	10:54	-
3. Nsima	12:34	12:30
4. Relish with beans	12:34	12:30
5. Relish with vegetables	12:34	12:30
6. Nsima	16:34	16:00
7. Relish with beans	16:34	16:00

Participant 43		
1. Porridge	8:00	9:00
2. Puffs	9:04	9:30
3. Nsima	13:40	13:00
4. Relish with beans	13:40	13:00
5. Porridge	16:45	16:00
6. Puffs	-	18:00

Participant 44		
1. Tea	7:30	8:00
2. Bun	-	8:00
3. Porridge	9:40	8:00
4. Nsima	14:20	13:00
5. Relish with beans	14:20	13:00
6. Papaya	-	16:00
7. Nsima	-	*19:00*
8. Relish with beans	-	*19:00*
9. Bread	-	*21:00*

Annex 6

Wilcoxon matched-pair signed-rank test separately for intervention and control groups

Group median intakes of energy and nutrients measured by 12-i-24HR and WFR were similar in the intervention group (Wilcoxon matched-pair signed-rank test $p > 0.05$), except protein intake (Wilcoxon matched-pair signed-rank test $p < 0.05$) (Table 1). Disagreement regarding protein intake may have been caused by underestimated amounts of relishes with animal protein (fish, egg, chicken, and beef). Eight participants in the intervention group consumed a portion of a relish with animal protein and each of the portions was underestimated. For the eight participants, these relishes accounted for 35 % of their protein intake as measured by WFR. The ranges of intake between 25th and 75th percentile were narrower for each estimate based on 12-i-12HR compared with WFR. Estimates based on 12-i-24HR were within a range of -6 % (zinc) to 51 % (iron) of their corresponding WFR intake estimates.

Table 1. Intervention group: relative validity of the interactive 24-hour recall at the group level: median intakes of energy and selected nutrients between 6:00 hours and 18:00 hours measured by interactive 24-hour recall and weighed food record: comparison of median intakes and statistical significance of differences (n=21).

Dietary factor	WFR median (25th, 75th percentile)	12-i-24HR median (25th, 75th percentile)	Absolute difference[a]	% difference[b]	P-value[c]
Energy, kJ	2175 (1503; 2390)	1988 (1453; 2253)	-187	-9	0.48
Protein, g	11.8 (7.1; 15.4)	9.8 (6.2; 12.2)	-2.0	-17	0.05[d]
Fat, g	13.2 (9.9; 19.1)	14.1 (10.6; 17.5)	0.9	7	0.73
Iron, mg	6.5 (3.6; 12.7)	9.8 (4.1; 11.9)	3.3	51	0.46
Zinc, mg	4.7 (3.7; 7.6)	4.4 (2.7; 6.3)	-0.3	-6	0.39
Vitamin A, µg	411 (93; 975)	614 (307; 970)	203	49	0.90

[a] 12-i-24-HR – WFR
[b] (12-i-24HR-WFR) / WFR × 100 for group median intakes
[c] P-value of Wilcoxon matched-pair signed-rank test between WFR and 12-i-24HR
[d] Significant at the 0.05-level
WFR, weighed food record; 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours

For the control group, 12-i-24HR and WFR gave similar estimates for energy and each selected nutrient (Wilcoxon matched-pair signed-rank test $p > 0.05$) (Table 2). The ranges of intake between 25th and 75th percentile were not systematically narrower or wider for estimates based on 12-i-12HR compared with WFR. In the control group, four participants consumed a protein-rich relish of animal origin. For the four cases, amount of three portions were overestimated and the quantity of one underestimated. Estimates based on 12-i-24HR were within a range of 0.5 % (energy) to -26 % (fat) of their corresponding WFR intake estimates.

Table 2. Control group: relative validity of the interactive 24-hour recall at the group level: median intakes of energy and selected nutrients between 6.00 hours and 18.00 hours measured by interactive 12-hour recall and weighed food record: comparison of mean intakes and statistical significance of differences (n=23).

Dietary factor	WFR median (25th, 75th percentile)	12-i-24HR median (25th, 75th percentile)	Absolute difference[a]	% difference[b]	P-value[c]
Energy, kJ	1460 (1223; 1956)	1468 (1115; 1862)	8	0.5	0.58
Protein, g	9.1 (6.0; 10.5)	7.2 (4.5; 8.8)	-1.9	-21	0.18
Fat, g	10.0 (5.2; 13.3)	7.4 (5.9; 10.5)	-2.6	-26	0.45
Iron, mg	2.3 (1.7; 3.6)	2.5 (1.6; 3.0)	0.2	9	0.11
Zinc, mg	1.6 (1.1; 2.2)	1.3 (1.0; 2.4)	-0.3	-19	0.18
Vitamin A, µg	86 (13; 134)	84 (12; 235)	-2	-2	0.81

[a] 12-i-24-HR – WFR
[b] (12-i-24HR-WFR) / WFR × 100 for group median intakes
[c] P-value of Wilcoxon matched-pair signed-rank test between WFR and 12-i-24HR
WFR, weighed food record; 12-i-24HR, interactive 24-hour recall concerning the same day of food intake as WFR and covering a time period from 6:00 hours to 18:00 hours