Web-based dietary assessment in Norway: New methods and their validity

Anine Christine Medin

PhD Thesis

Department of Nutrition
Institute of Basic Medical Sciences
Faculty of Medicine
University of Oslo

Oslo 2018

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-8377-155-8

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Acknowledgements

The work presented in this thesis was carried out at the Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, in the period from December 2012 until July 2017. Funding was primarily obtained from the Institute of Basic Medical Sciences, University of Oslo through a four-year doctoral fellowship.

I would especially like to thank my principal supervisor Lene Frost Andersen for trusting in me from the very beginning. Your experience, high level of knowledge and ability to avoid micromanaging has made it a real pleasure being under your wing. Thank you for always finding the time to discuss my work, whenever I needed, despite your hectic schedule.

Likewise, I would like to thank my co-supervisor Monica Hauger Carlsen. I have appreciated your valuable help in practical KBS-issues and particularly our interesting discussions regarding methodological issues. You have the unique ability to be critical in the kindest way.

Thanks to Helene Astrup and Britt Marlene Kåsin for being an excellent observational team, and for providing invaluable practical assistance during several months of data collection. Thanks to Ida Sofie Kaasa and Helene Astrup (once more), for helping me to complete several hundred telephone interviews.

I also want to thank Catherine Hambley and John Speakman from the University of Aberdeen, Scotland, and Bjørge Herman Hansen and Ulf Ekelund from the Norwegian School of Sport Sciences for productive collaborations.

I want to thank Susi Strohmaier, for clarifying statistical issues and helping out with R, but most of all the good discussions we have had together with Lena Leder. It has been great having both of you around.

Being part of the Dietary Research Group and the Section of Nutritional Epidemiology, consisting of plenty of highly skilled, motivated and curious people, has been a privilege. I have had the pleasure to share my office with Mekdes and Mona during their postdoc periods. I have travelled to Australia with Sonja, who was my support team during my first oral presentation at an international congress. Thanks to Anne Lene, Mari, Nicolai, Anne Marte, Anette and Jannicke for interesting discussions. I also want to thank the rest of my many great colleagues over the years at the University of Oslo.

Thanks to my dear cousin and friend Tirill, for guidance on PhD issues; my grandfather Bjørn for being enthusiastic about the importance of methodological issues, and for being genuinely interested in my work. Thanks to the rest of my family and friends, who have shown their support.

Last but not least, I want to thank Øyvin, Balder and Alvin for their patience. I believe you have felt my PhD was a family project at times, especially during data collections.

Oslo, July 2017

Anine C. Medin

Table of contents

List of papers

The present thesis is based on the following publications:

Paper I — **Medin AC**, Astrup H, Kåsin BM, Andersen LF. Evaluation of a Web-Based Food Record for Children Using Direct Unobtrusive Lunch Observations: A Validation Study. J Med Internet Res 2015;17(12):e273

Paper II — **Medin AC**, Carlsen MH, Andersen LF. Associations between reported intakes of carotenoid-rich foods and concentrations of carotenoids in plasma: a validation study of a web-based food recall for children and adolescents. Public Health Nutr. 2016 Jun 27:1-11.

Paper III — **Medin AC**, Hansen BH, Astrup H, Ekelund U, Andersen LF. Validation of energy intake from a web-based food recall for children and adolescents. PLoS One. 2017 Jun 8;12(6):e0178921

Paper IV — **Medin AC**, Carlsen MH, Hambley C, Speakman JR, Strohmaier S, Andersen LF. The validity of a web-based food frequency questionnaire assessed by doubly labelled water and multiple 24-hour recalls.
Submitted to British Journal of Nutrition, 01-Jun-2017. Awaiting referee scores.

Abbreviations

24HR	24-hour recall
DBS	Dried blood spot
DLW	Doubly labelled water
EI	Energy intake
FFQ	Food frequency questionnaire
IOR	Interobserver reliability
KBS	Kostberegningssystem (food composition database and calculation system)
PDA	Personal digital assistant
REE	Resting energy expenditure
RFPM	The Remote Foods Photography Method
TADA	The Technology Assisted Dietary Assessment
TEE	Total energy expenditure
WebFFQ	The web-based food frequency questionnaire
WebFR	The web-based food record/recall

1 Introduction

1.1 The rationale for assessing dietary intake

Diet influences human disease. According to the Global Burden of Disease Study 2013, dietary risks accounted for more than 11 million deaths, and 241 million lost healthy lives (DALYs) globally in 2013 alone [1]. This insight comes from epidemiology, which has been described as *"the study of the distribution and determinants of health-related states or events (including disease), and the application of this study to the control of diseases and other health problems."* [2]. More specifically, it is nutritional epidemiology, in which the focus is on nutritional determinants [3], that has provided this understanding of dietary risks. To get further insights, and to prevent deaths, diseases, and health problems influenced or caused by nutritional determinants, there is an inevitable need to assess dietary intake.

1.2 A glimpse into the history of dietary assessment

Modern nutritional science first materialised in the late 1700's, under the so-called chemical revolution in France [4]; and it was not until the vitamin era during the first half of the 1900's, in which deficiencies in humans were given much attention, that nutritional science expanded massively [5]. The newly discovered nutrients and their content in common foods led to the possibility and interest in analysing dietary intake in humans; and with that, a demand for dietary assessment methods was created [6].

The first recognised written report on individual dietary assessment was published in 1936 in England [7], during dietary assessment methodology's early phase. Different forms of food records, 24-hour recalls (24HRs), and chemical food analyses of actually eaten foods were used in this period [6]. In 1947, Burke introduced the diet history method as a research tool developed to capture the average food intake for an individual [8], which was an advancement. Around this time, small-scale research studies dominated, and the diet history and lengthy paper-based food records were the principal methods in use [9]. The food frequency interview method, for assessment of the usual dietary intake, was introduced in 1962 [10]. It took until the early 1980's before paper-based food frequency questionnaires (FFQs) emerged and were used in large-scale epidemiological studies, like in the American Nurses' Health Study cohort

[11]. Such large observational studies became feasible and popular at the time, due to the previous decade's development in computer technology and statistical methods [6, 9].

Table 1. The main traditional dietary assessment methods based on self-reports [12].

Methodology	First made reference to	Brief description	Traditional form of administration	Information collected
Food record/ food diary	○ In 1936 by Widdowson (UK). A 7-day weighed food record [7].	○ Everything consumed is recorded in real time, for typically 3-7 consecutive days. ○ Weights can be used to increase the accuracy of portion sizes.	○ Self-administered. ○ Paper-format.	○ Short-term dietary intake of the whole diet. ○ Detailed and contextual information.
24-hour recall (24HR)	○ In 1938 by Burke (US). A 24HR interview [13].	○ Everything consumed over the past 24-hours are reported in great detail.	○ Interview administered. ○ Phone, or face-to-face.	○ Short-term dietary intake of the whole diet. ○ Detailed and contextual information.
Food frequency questionnaire (FFQ)	○ In 1962 by Stefanik and Trulson. The food frequency interview. ○ In the 1980's by several: E.g., Willet's paper-FFQ used in Nurses' Health Study [11].	○ The usual consumption of foods and beverages, found in a fixed list is specified by frequency, and often the amount. ○ It covers a defined period, usually the last 1-12 months.	○ Self-administered. ○ Paper-format.	○ Usual dietary intake assessed by a single administration. ○ Entire or parts of the diet.

The 1980's also became a decade of significant events related to biological markers of dietary intake. The work over the previous decades materialised into some important publications. Isaksson published his paper on 24-hour urinary nitrogen as a biological marker for total protein intake [14]. Schoeller introduced the doubly labelled water (DLW) method to assess total energy expenditure (TEE) in humans [15], whereas Plakké proposed that the composition of fatty acids in an individual's fat tissue reflected the composition of fatty acids in their habitual diet [16]. In the early1990's, the field of biological markers of dietary intake advanced further; Ziegler published a paper introducing a new method to analyse single carotenoids in blood, also found in vegetables and fruits, to assess their relation to lung cancer [17]. Around the year 2000, the principal methods based on self-reports were still food records, food recalls in addition to both long and short FFQs [12]. It could, therefore, appear as if the development of dietary assessment had stagnated. However, around this period, novel technological approaches started to emerge [9].

1.3 Dietary assessment today

1.3.1 New technology and methods

Technology is now everywhere, and it is widely recognised that humanity is on the edge of a technological revolution, believed to profoundly alter everything from civil society to private and public sectors, including academia [18]. Dietary assessment methods are not left unaffected by this, which is reflected in the increasing heterogeneity in methodology and use of technology described extensively in a review by Illner et al. in 2012 [19], and in several other recent review articles [20-24]. Nevertheless, the conventional food record, 24HR and FFQ, described in Table 1, are not abandoned, and are still in use and form the basis for many of the new tools. An overview of new approaches assessing dietary intake is displayed in Table 2.

Assessment of past intake

The simplest adaptation of the traditional methods is a straightforward digitalization of paper-based questionnaires, without any use of additional images or interactive features. Such computer- or web-based dietary assessment tools may be troubled with many of the fundamental challenges as paper-based tools [19] but can offer advantages for researchers due to reduced resources needed for data handling, and flexibility regarding when to collect data [25]. The possibilities of built-in error checks securing completeness and consistency of the web-based questionnaires are additionally clear advantages. By incorporating images for portion size estimates, like what is done in the GraFFQ [26], an additional refinement is added that may increase accuracy.

Several self-administered web-based 24HR platforms have been developed for different age groups, based on the principles of the 24HR methodology [24]. There was initially a shift from traditional 24HRs interviews, to interview-assisted software on computers, like the American Automated Multiple-Pass Method (AMPM) developed in the late 1990's [27], and finally to self- administered platforms, like the American ASA24 first used in 2006 [28]. By avoiding the costly and inconvenient interviews, the use of multiple self-administered web-based 24HRs is becoming a real alternative to the FFQ for use in large-scale nutrition studies [29].

Table 2. New approaches[a] to assess dietary intake

Approach	Assessment of past intake		Real-time assessment (ambulatory assessment)	
Brief description	○ Assessment of self-reported diet in recent or distant past. ○ Self-administrated reports (or interview-assisted). ○ Possibilities of audio and image functionalities. ○ Can be interactive. ○ Memory dependent.		○ Assessment of self-reports or objective data, during normal daily activities. ○ Self-administrated (active) or automatic recording (passive). ○ Considerable method heterogeneity. ○ Various portable devices can be used (e.g. PDAs, smart phones, cameras, sensors, scanners). ○ Instant, memory independent.	
Information collected	○ No real-time data collection. ○ Short- or long-term past dietary intake.		○ Real time data in the natural environment using portable devices. ○ Short-term dietary intake ○ Rich contextual info possible to collect (e.g. time, GPS).	
Selected examples, grouped by category	**Computer-based** Not connected to the Internet. ○ **24HR software, interview-assisted** • SCRAN24 [30], EPIC-Soft [31], AMPM 24HR [27]. ○ **24HR software, self-administered** • FIRSSt [32], YANA-C [33], IMM [34]. ○ **Self-administered questionnaires** • FFQs [35, 36], DHQs [37-39].	**Web-based** Connected to the Internet. ○ **24HR web-based, self-administered** •ASA24 [28], ASA24-Kids-2012 [40], CANAA-W [41], SNAP [42], FIRSSt4 [43], PAC24 [44], INTAKE24 [45], myfood24 [46], Compl-eat [47]. ○ **Self-administered questionnaires** • FFQs [26, 48-52], DHQs [53-55]. ○ **Record/recall-hybrids**[b] • WebFR[c] [56], WebDASC [57]. ○ **Miscellaneous** • 24HR FL [58], Foodbook24[d] [59].	**Self-administered** Active assessment. ○ **Image based food records for handheld devices** Images captured of dietary intake on the go. Can be complemented using voice recording, text, etc. •TADA [60], RFPM [61], FRapp [62], Wellnavi [63], Nutricam [64]. ○ **Other food records for handheld devices** Can be complemented by capturing images or adding comments as memory aids. •For PDAs [65]. •Apps for smartphones: MMM [66], Mobile Phone App [67] e-CA [68]. •Spoken dietary records [69]. ○**Wearable camera** Audiovisual recordings. Must be switched on before eating. • Micro-camera [70]. ○ **Scan tech** Used as stand-alone tools or combined with e.g. PDAs. • Handheld scanners [71, 72].	**Automated** Passive assessment. ○ **Wearable sensor tech** Eating sound recognition, jaw and hand motion sensors. •AutoDietary [73], AIM [74] ○ **Wearable camera** Photos captured automatically: •Microsoft SenseCam[e] [75]. ○ **Wearable computer** Video, photos, sound, etc. captured automatically: •eButton [76].

PDA, Personal digital assistant; App, application software; Tech, technology; DHQ, diet history questionnaire; FFQ, food frequency questionnaire; 24HR FL, 24-hour food list; 24HR, 24-hour recall.
[a] Biomarkers of dietary exposure are omitted from this overview.
[b] Food records/diaries in which recordings are not done in real-time.
[c] Included in the current thesis.
[d] Combination of 24HR and FFQ.
[e] Used in combination with traditional food records or 24HRs [77, 78].

Some new technology-based methods, defined as food records or diaries in Table 2, are categorised as retrospective dietary assessment tools because they do not allow recordings in real-time.

The web-based food record/recall (WebFR) validated in this thesis is a typical example of this. It was constructed as a classic record or diary in which recordings are done for the duration of several consecutive days. However, users do not record in real-time, but rather in the evening each recording day. The method is therefore per definition an assessment of past intake, bearing similarities to a recall, covering the previous hours only, in contrast to the 24HRs, in which the participant must recall the dietary intake the previous day.

Real-time or ambulatory assessment

Several methodologically new dietary assessment methods fit under the umbrella term ambulatory assessment. The term ambulatory assessment origins from the field of psychology, and covers a range of real-time assessments conducted in real-life, aided by computer-assisted technology, believed to minimise recall bias due to the nature of the momentary data collection [79]. The data collection using ambulatory assessment typically involves multiple assessments for an individual over a defined period, using methods listed in Table 2. Among the various assessment methods, a few automated, sensor-based tools, like the AutoDietary [73], using eating sound recognition, have been developed. The diverse group of self-administered event-based records are, however, more common. They range from portable food records on personal digital assistants (PDAs), with food lists and images of portion sizes [65], to image-based smartphone applications like the Technology Assisted Dietary Assessment (TADA) [60], in which the users first and foremost actively captures images of their eating occasions.

Mixed method approach

The ambulatory assessment captures real-time data, and thus any recall methodology is per definition excluded from this category. Mixed methods have, however, been developed. For example, using a combination of a web-based 24HR, and a wearable camera to automatically capture real-time images the day previous to the 24HR [77, 80], thus using features from ambulatory assessment as an aid to improve the 24HR.

1.3.2 Biomarkers of dietary exposure

Several definitions of biological markers, often called biomarkers, have been established in the literature [81]. The Biomarkers Definitions Working Group has described a biomarker as: *"A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."* [82]. Biomarkers may serve as substitute endpoints, to predict coming events, whereas past events are determined by the biomarkers of exposure [83].

The term nutritional biomarker also covers a variety of objective biological measurements, including markers of future disease or pre-clinical disease, nutritional status or dietary exposure [84, 85], displayed in Figure 1. However, the term is often used exclusively for biomarkers of dietary exposure [86-88], that is, as objective indicators of past dietary intake.

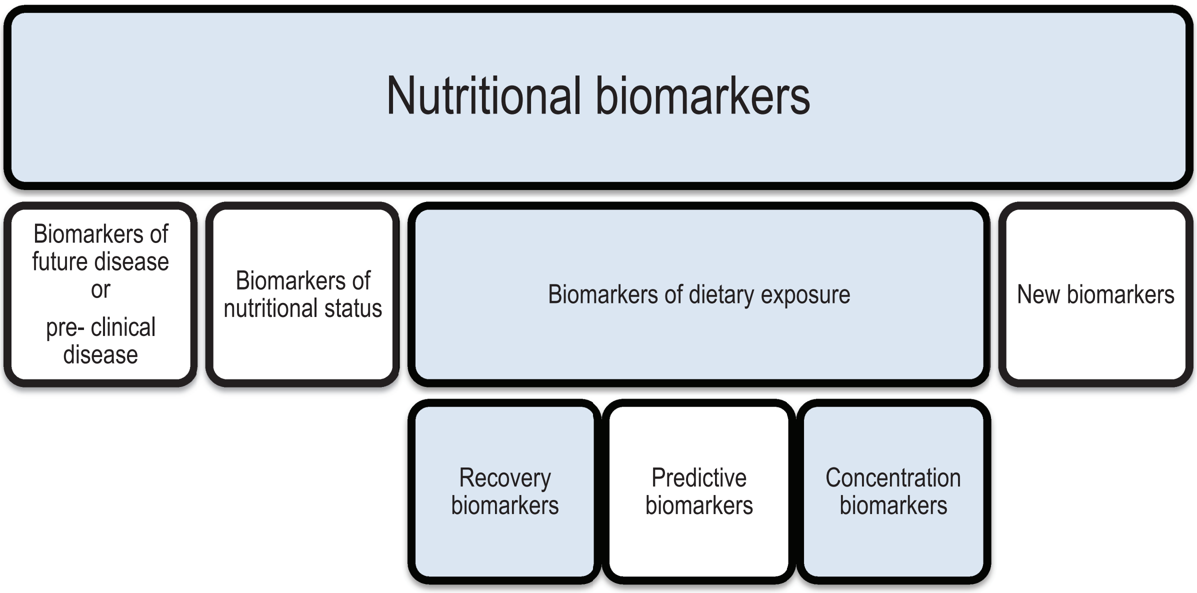


Figure 1. An overview of classification of nutritional biomarkers. The light blue coloured boxes are the types of biomarkers used in the current thesis.

Biomarkers of dietary exposure may be used to assess dietary intake through ranking or quantification, alone or as a supplement to other dietary assessment methods [84]. They are usually classified into three different groups: The recovery biomarkers, predictive biomarkers and concentration biomarkers [86].

Recovery biomarkers

For the very few existing recovery biomarkers, estimates of the absolute intake level are possible to obtain, due to a metabolic balance between intake and excretion over a particular period, in individuals in homeostatic balance [86]. The only known recovery biomarkers are DLW for energy [15], 24-hour urinary nitrogen for protein [89] and 24-hour urinary sodium [90] and potassium [91], reflecting the sodium and potassium intake, respectively. In Paper IV, the method of DLW was used.

<u>The DLW method</u>

The DLW method is a technique based on isotopes, used to estimate the TEE in humans in a natural setting [92]. Isotopes are forms of the same atoms with nearly identical properties; they have the same number of protons, only differing in neutron number, resulting in a slightly different weight. Protium (^{1}H), deuterium (^{2}H) and tritium (^{3}H) are all examples of isotopes of hydrogen, with zero, one and two neutrons in their nucleuses, respectively. In the DLW method, water (H_2O) is labelled by replacing a proportion of the most common isotopes of oxygen (^{16}O) and hydrogen (^{1}H) with the detectable, stable isotopes oxygen-18 (^{18}O) and deuterium (^{2}H), naturally present at very low concentrations [93]. The labelled water is typically administered orally in humans [93]; subsequently, isotopes are equilibrated in the body pool and washed gradually out of the body normally over a period of 4-14 days [92]. ^{2}H is primarily lost through H_2O (urine, sweat, etc.), and ^{18}O is lost through both H_2O and CO_2 (respiration); Thus, the wash-out rate between the two isotopes in e.g. urine is used to estimate the amount of ^{18}O that escaped the body pool through respiration as CO_2 [93]. Ultimately, this provides an estimate of the CO_2 production over the measurement period, which together with an estimate of the respiratory exchange ratio and equations, are used to calculate TEE [93]. The estimated TEE can subsequently be used as a marker for energy intake (EI) in weight-stable individuals.

Predictive biomarkers

Predictive biomarkers also show high correlations with intake, similar to the recovery biomarkers; however, their overall recovery is incomplete [86]. Thus estimates of the absolute intakes cannot be obtained. The only known predictive biomarkers today are 24-h urine sucrose and fructose, reflecting sugar consumption [94].

Concentration biomarkers

Most biomarkers of dietary exposure are concentration biomarkers, from which estimates of absolute intake cannot be obtained [95]. However, because the dietary intake (exposure) is associated with the concentration of these biomarkers [86], it is possible to rank individuals according to intake, i.e. to differentiate between low and high consumers [84]. Examples of the many concentration biomarkers include fatty acids in either adipose tissue, erythrocytes, or in plasma- or serum compartments, that reflect long-, medium- and short-term intake of specific fatty acids, respectively [84]; or carotenoids in blood, which are markers for carotenoid-rich foods [96]. In Paper II, the latter was used as a biomarker of exposure.

Carotenoids

Humans and animals cannot synthesise the natural pigments called carotenoids, in contrast to plants and microorganisms [97]. Consequently, all carotenoids detected in humans' blood, as in Paper II, can only originate from dietary intake. More than 700 carotenoids are identified [96], out of which about 50 have been identified to be absorbed and metabolised in humans [98]. Out of these, just a few (β-carotene, α-carotene, β-cryptoxanthin, lycopene, lutein and zeaxanthin) are both found in humans' diet, and also in a significant concentration in their blood [96]. Most of the dietary carotenoid intake in the western world originates from consumption of fruits and vegetables (80-90%) [96]. A dose-response relationship has been observed between the consumption of fruits and vegetables, and concentrations of these previously mentioned carotenoids in plasma, in controlled feeding studies [99-101]. Measuring carotenoids in plasma is therefore used as an objective indicator of the true intake of fruits and vegetables. However, different fruits and vegetables have a highly variable content of carotenoids [102]. For this reason, concentrations of carotenoids in plasma may be a more valuable marker for selected fruits and vegetables, rich in these particular carotenoids, than a general marker for the total intake of fruits and vegetables. Therefore, in Paper II, variables for carotenoid-rich foods were created, for comparison with the concentration of carotenoids in blood.

New biomarkers

The food metabolome

Metabolomics is the study of the metabolome, which is made up of small molecules called metabolites [103]. A part of the human metabolome is the food metabolome, which consists of several thousand metabolites originating from the diet, through digestion, absorption and metabolising of foods [88]. Urine is a much-used specimen for identifying typically new biomarkers, as non-nutrients (or nutrients in excess) or their secondary metabolites are excreted in the urine, reflecting the intake over the past hours [103]. O'Gorman listed several newfound putative biomarkers, derived from the food metabolome, that are supposed to reflect the intake of specific foods: Salmon, broccoli, whole grain wheat cereals, raspberry, cruciferous vegetables, citrus fruits, coffee, onions and red meat [104]. However, only a very few of these metabolites, like the proline betaine that reflects the consumption of citrus fruits, have been extensively validated [105]. The lack of long-term biomarkers is also an apparent limitation [105]. Nevertheless, it is expected that numerous more will be discovered in the years to come, as this is still a large unexploited area of research [88].

Stable isotopes

Stable isotopes were used already in the 1930's, in studies of metabolism [106], and still is [107]. Moreover, the previously described DLW method, usually classified as a recovery biomarker, is in fact, a stable isotope based technique. The use of stable isotopes ratios has recently been suggested as an approach to identify new biomarkers of dietary exposure at the atomic level, for use in epidemiology [108]. The idea comes from the studies of archaeology, palaeontology, and ecology, and has, for instance, been used to study the diet of our ancestors and extinct animals [109, 110]. The natural and consistent variation of stable isotopes between different foods is also captured in human tissues, reflecting the dietary intake [107]. The stable isotopes ratios of both carbon and nitrogen, in the forms of $^{13}C/^{12}C$ and $^{15}N/^{14}N$, are for instance suggested as biomarkers of fish protein intake [111]. They may be used as both short- and long-term biomarkers, depending on the rate of elemental turnover of the tissue, from which the sample is drawn (e.g. hair does not undergo any elemental turnover) [108]. The possibility to obtain information on long-term dietary exposure makes the stable isotopes especially interesting for nutritional epidemiology.

1.3.3 Other objective measurements of dietary intake

Direct observations of dietary intake

Direct observation is used to provide information about numerous activities and behaviours, without the need of depending on participants' ability or willingness to answer questions [112]. In contrast to indirect observations, in which outcomes of an activity or behaviour are observed (i.e. children's plate waste after their school lunch), direct observations capture information about activities or behaviours as they occur (i.e. eating). By using direct observation of dietary intake, objective information is obtained from eating events [113], while keeping participant burden low. However, direct observations of eating are often resource-demanding and time-consuming [113], and due to practical reasons often limited to parts of the day [114]. Typically, for school children, observations are conducted by several observers during school meals. Interobserver reliability (IOR) should therefore ideally be assessed [114]. Because direct observation is susceptible to participant reactivity [115], it is strongly recommended to use unobtrusive observations [116]. In the validation study of the WebFR, we used unobtrusive observations. For that reason, we did not interact with the participants during observations. Additionally, the observations were blinded, so the children did not know who was under observation on a given day.

Indirect assessment of energy intake using accelerometer counts

Accelerometers are electronic motion sensors, providing direct objective measurements of physical activity and sedentary behaviours in free-living conditions [117]. Today, sensors are based on microelectromechanical system technology (MEMS) [118]. They are incorporated into small wearable devices and measure the acceleration of the part of the body on which the accelerometer is placed; then this measured acceleration is converted into a signal, which can be processed into activity counts [119]. By quantifying all activity counts in defined time intervals called epochs, both intensity and duration of physical activity and sedentary behaviours are possible to determine, from defined cut-points for different thresholds of intensity [119]. Such physical activity estimates or activity counts from accelerometers may be combined with data on body size, sex and age, or measured or estimated resting energy expenditure (REE) prediction equations, to calculate TEE [120]. In weight-stable individuals, this estimated TEE can be used as an indirect measure of EI, which was done in Paper III in this thesis.

1.4 Measurement error in dietary assessment

Objective, quantitative measurements of long (usual) or short-term dietary exposure for all nutrients and food groups do not exist. Hence, it is impossible to fully avoid dietary assessment methods based on self-reported data. Unfortunately, such self-reports are particularly troubled with measurement errors, which may attenuate or distort observed associations between dietary exposures and outcomes in nutritional epidemiology [9, 95].

Measurement error is the deviation from the true value [9]. The two main sources of measurement error in dietary assessment are random within-person errors and systematic errors (bias) [121, 122]. The random within-person errors originate from day-to-day variation (deviations from the usual mean intake), in addition to any random error in the measurement (e.g. clicking on the wrong portion size image) [95]. Bias is, on the other hand, consistent deviations from the true intake in a particular direction; the most important types are person-specific bias and intake-related bias [121]. Person-specific bias is related to characteristics of individuals like age and weight status [9], typically manifested as a constant underestimation or overestimation of certain foods by certain individuals, due to social desirability. Intake-related bias results from systematic errors that are proportional to the dietary intake [9]; those with high intakes (e.g. of sweets) may for instance typically under-report their intakes more than moderate- or low intake consumers [121].

The random within-person error will lead to loss of power and inflate the variation in a group and may attenuate the relationship between diet and health [9]. In comparison, the consequences of bias are more complex; they can lead to either exaggerating or attenuating diet-health relationships and can distort group mean intakes and distributions [121].

1.4.1 Identifying measurement errors

Evaluation studies are useful when trying to identify measurement errors in dietary assessment. There are two different types: Reproducibility studies and validation studies [9]. In reproducibility studies, the presence of random within-person errors can be identified, by evaluating the consistency of a method administered more than once, at different time points, to the same individuals [95]. Hence, if the reproducibility of a method is high, that means it is precise.

To what extent a method measures what the method is intended to measure, is defined as its validity [123]. A method may be precise, but not valid if bias is present [123]. Moreover, accuracy describes the degree to which a measurement obtained from a method deviate from the true value [9]; e.g., how much the mean fish intake in a population measured by an FFQ deviate from the true mean fish intake. Accordingly, only a valid method can provide high accuracy of the measurements obtained by that method. To identify bias, we need validation studies [95]. The validity of a 24HR can be assessed by evaluating to which degree the estimated intake reflects the true intake, the previous day, by comparing the 24HR estimates (the test method) to a superior reference method [9]. Objective and independent reference methods that measure the true intake without bias are the ideal options in validation studies [124]. Biomarkers of exposure can serve as such objective reference methods in validation studies [84, 86], but not many recovery biomarkers are available [85], as explained previously. Direct observation is another option for obtaining objective reference measures, but observation of individuals' dietary intake for entire days, over extended periods of time is often not feasible [116]. Methods based on self-reports are therefore often used as a reference tool, despite being biased and having correlated errors with the test method [125]. When such a comparison is made, we use the term relative validation, to indicate that the reference method is imperfect and that the test and reference methods are not independent of each other [124]. Consequently, a relative validation study may result in a high agreement between methods, which may not be due to the high accuracy of the methods, but that they measure the same construct in the same direction or way.

1.4.2 Statistical techniques to reduce measurement errors

It is crucial to reduce measurement errors to a minimum during data collection, e.g. by using standardised protocols, valid tools, repeated measures (e.g. multiple 24HRs) and training of researchers [9]. However, it is also possible to handle measurement errors in the phase of analysing data.

Several statistical techniques have emerged to correct or reduce the impact of measurement errors [95]. Among these approaches, we find techniques for energy adjustment [126], removal of within-person errors (day-to-day variation) [127, 128], and regression calibration [125]. The two first approaches were used in the current thesis. The regression calibration approach is useful in studies of diet-health relationships; the risk estimate can be recalculated, using attenuation factors calculated from a subgroup in the main study, in which a superior dietary method is used in addition to the main method [9]. For example, one could use 24HRs in the main study and recovery biomarkers in a sub-sample.

Combining self-report methods in new ways may be a promising development. FFQ data may be used to estimate the probability of consuming different foods, coupled with 24HR data to determine the amounts consumed [25]. Merging data from self-report methods and biomarkers is also suggested as a possible approach to mitigate measurement errors, and thus improve estimates of dietary intake [129].

1.5 Aim and objectives

The aim of this thesis is to assess the validity of two web-based dietary assessment tools developed for children and adolescents, and adults, respectively; for use in both descriptive- and analytical nutritional epidemiology studies, in addition to clinical studies, in Norway.

Objectives

- The following objectives were set out to validate the first web-based dietary assessment tool for children and adolescents in Norway, the WebFR, using three different reference methods:

  i. To assess the accuracy of school lunch entries in the WebFR, using direct unobtrusive observation as the reference method, in the age group 8–9 years.

  ii. To assess the ranking abilities of the WebFR for carotenoid-rich foods, by comparing reported intakes of carotenoid-rich foods to concentrations of carotenoids in plasma, in the age groups 8–9 and 12–14 years.

  iii. To assess the validity of EI estimated from the WebFR, using TEE calculated from accelerometer outputs, combined with data on weight and sex or combined with REE prediction equations, as the reference method in the age groups 8–9 and 12–14 years.

- The following objectives were set out to validate a new web-based FFQ, the WebFFQ, for assessment of habitual dietary intake among Norwegian adults, using two different reference methods:

  i. To assess the absolute validity of the estimated EI from the WebFFQ, using TEE measured by DLW as the reference method.

  ii. To assess the relative validity of the estimated intakes of macronutrients and food groups from the WebFFQ, using repeated non-consecutive 24HRs as the reference method.

2 Methods and materials

The validity of two web-based dietary assessment methods is assessed in this thesis: namely the web-based food record/recall (WebFR) and the web-based food frequency questionnaire (WebFFQ).

2.1.1 The study sample in the WebFR validation study (Paper I-III)

The WebFR was validated in a study carried out in the period from September- December 2013, in the municipality of Bærum, in Norway. The tool was developed for use in a national dietary survey in Norway, UNGKOST 3, among children and adolescents, in the 4th grade (8-9 years) and the 8th grade (12-14 years), respectively [130]. Thus, we invited 414 pupils in these age groups to participate in the validation study (Figure 2).

Convenience sampling was used; the principals of 11 schools in a short travel distance from the University of Oslo were tried reached by phone, out of which nine responded after calls on no more than two different days. Principals of six schools showed an interest in the project and were formally invited by email or mail, and their respective schools were subsequently included in the study. To increase the variability in the sample, we invited schools from a part of the municipality known to have a relatively heterogeneous population, with regards to their socio-economic status and ethnic background.

Information regarding the study was provided to the invited pupils, and to their parents or guardians, in classrooms during school hours, and at plenary school meetings for parents, respectively. Besides, all invited received written material. Pupils who wanted to be included in the study had to have Wi-Fi at home, and their parents/guardians had to provide the researchers with an email address, of which they were responsible.

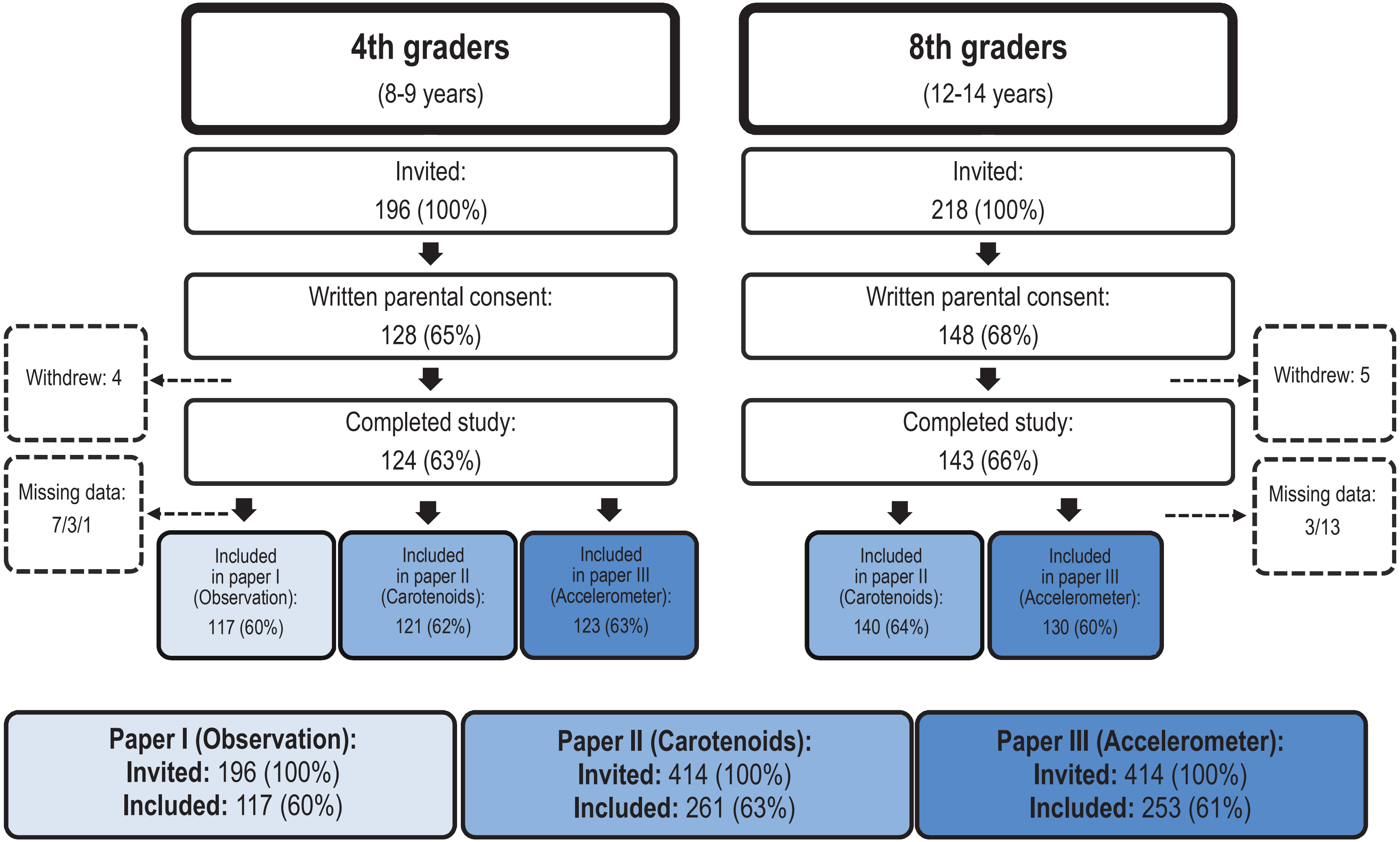


Figure 2. Flow chart showing the recruitment process in the validation study of the WebFR (paper I-III).

2.1.2 Design of the WebFR validation study (Paper I-III)

All participants had to enter types and amounts of all food items and beverages they consumed in the WebFR for four days. Moreover, they had to wear an accelerometer for seven days, during the same week. The youngest participants (8-9 years old) were also observed during school lunch in one out of the four days they recorded in the WebFR. In the following period upon completing the recordings in the WebFR (maximum 11 days), blood drops were collected from all participants after a minor puncturing of one of their fingertips. Their height and weight were also measured.

The WebFR

The WebFR is a hybrid dietary assessment tool, designed as a food record, yet with elements from recall methodology. This is because recordings are not done in real-time, but rather at the end of each recording day. The WebFR is based on a pre-set meal structure, and images are used to estimate portion sizes. Selected screenshots from the WebFR are provided in Appendix I. It is an 'open method', as it is possible to enter information regarding any food or beverage consumed, but not listed in the WebFR. In total, the WebFR contains around 550 items, selected based on data on frequently consumed items in NORKOST 3, a Norwegian national dietary survey from 2011 [131], in addition to unique children's products (e.g. yoghurt), which were selected based on sales statistics in Norway.

The WebFR is based on the Danish Web-based Dietary Assessment Software for Children (WebDASC) [57]. Several aspects of the WebDASC were changed during the process in which the WebFR took its form: all text and audio files, selected aspects of the interface, and types of meals, food lists, selected images and the food composition database. The image series in the WebFR consists of a mix of new image series specially made for the WebFR and image series that originated from the Danish WebDASC. The suitability of the portions sizes shown in the image series was evaluated by experienced dietitians before they were included in the WebFR. Despite these alterations, the underlying construct and basic functions are the same in both the WebDASC and the WebFR.

The WebFR has an interface that intends to be both intuitive and enjoyable for children and adolescents – its target group. An interactive, voice-assisted cartoon character guides the users through each day's eating occasions, chronologically, aided by both audio and text in speech bubbles. Participants enter the foods and beverages they have consumed for each eating event separately. There are three alternative approaches. Participants can either use a search function, or select items from a drop-down list, organised by categories, or use the option for open field entries. The list comprises three levels (e.g. beverages → milk → semi-skimmed milk). To specify the portion size of each food item, participants use the image series, which hold two to four images displayed at once, proving examples of various portion sizes. The participant clicks on the image that is the best fit for the consumed food item or beverage and indicates the number of portions consumed. For some items, the image series for portion size estimations show images of substitute foods. For instance, orange marmalade is illustrated by strawberry jam. Pop-up elements are incorporated to remind the participants to enter in-between snacks, supplements, or other items often omitted from reports, to reduce recall bias.

Entries in the WebFR

All participants were asked to enter everything they consumed in the period of four consecutive days. One out of the four days had to be a weekend day (i.e. Saturday or Sunday). Moreover, they were instructed to conduct the recordings at home, after their last meal, at the end of each recording day. Parents/guardians were instructed to assist the youngest participants (8-9 years).

Direct observation during school lunch - Paper I (8-9-year-olds only)

Direct observations of the 4th graders (8-9 years) were conducted at school, during their regular school lunch break. Participants ate their lunch brought from home, in their classrooms, as they normally did, while being observed. Each participant was observed once, during a weekday, in the same period as they were recording in the WebFR. Careful planning and training were conducted prior to the data collection to ensure that the observations were as unobtrusive as possible. The observations were single blinded: All the children received name tags each observation day and were not informed when they were observed. No contact with the children during the observation was permitted. All school classes were also paid a pre-observational visit, to make the participants familiarised to the observers being in the classroom.

The data collection was preceded by an extensive observer training. An assessment of the IOR was done both prior to and during the data collection. The IOR demonstrates the agreement between the different observers, based on the proportion of observations in agreement for each pair.

Carotenoids in blood - Paper II

Blood was collected from non-fasting participants, by a trained researcher, using the Dried Blood Spot (DBS) method. The school nurse's office or any other appropriate room at school, which could provide a minimum of privacy for the participants, was used during the sampling. A small finger-prick lancet was used, and a few drops of blood from the fingertips of each participant were placed right onto a filter paper, called DBS cards (Protein SaverTM 903R Cards; Whatman, Sanford, ME, USA). Blood sampling was conducted not more than 11 days after the participant had completed their recordings in the WebFR. This was done to analyse the concentrations of carotenoids in blood, and subsequently, to compare the concentrations of the biomarkers with the reported intakes of carotenoid-rich foods. Details of how the DBS samples were handled and later analysed, and how the carotenoid-rich food variables were constructed, are described extensively in Paper II.

Accelerometers - Paper III

Participants were instructed to wear the ActiGraph GT3X+ accelerometer (ActiGraph LLC, Pensacola, FL, USA). They were told to wear the accelerometer for seven consecutive days, including an entire weekend, and only to remove it during water activities (e.g. swimming, showering), and at night.

In Paper III, the mean of three different equations for TEE calculations was used. They were calculated from accelerometer outputs, combined with data on weight and sex or REE prediction equations. All prediction equations are shown in full length in the paper, in addition to a description of how the accelerometer data were handled.

Anthropometric measurements

Height and weight measurements were made, according to standard procedures, at the same time and location as the blood samples were collected. Height was measured to the nearest 1 mm and weight to the nearest 0.1 kg. The digital scale (TANITA TBF-300; Tanita Corporation, Tokyo, Japan) was used for weight measurements. Participants were only allowed to wear light clothing, and no shoes, when they were measured.

Child and parental characteristics

The parents/guardians of all participants were instructed to provide information on the participant's sex and age in the consent form. In addition, they were asked about their education level, ethnicity, and the type of family structure their family had.

2.1.3 The study sample in the WebFFQ validation study (Paper IV)

The study sample in the validation study of the WebFFQ, consisted of 92 participants that were recruited at two different time points from different populations, as shown in Figure 3.

Group 1, was recruited by convenience sampling, during a period of two weeks, in November 2015. An aim was set to recruit 32 women, based on sample size calculations. Thus, only women were recruited, using invites on Facebook, in addition to posters and word of mouth. A total of 58 women responded to the invites, out of whom 42 fulfilled the inclusion criteria, described in detail in Paper IV. Out of these 42 women, ten were excluded, and 32 were included in the study. The included women had the least comparable characteristics, defined by their age, self-reported body weight and height, self-reported physical activity level, and what area they resided. The purpose of this inclusion strategy was to increase the variability in the study sample. One of the included women withdrew from the study before the start of the study. Consequently, she was replaced by one of the10 previously excluded women, who did fulfil the criteria for inclusion. The data collection was conducted in the period from January to June 2016. All 32 women completed all parts of the study.

Group 2 was recruited and data collected continuously in the period from March to December 2016. A random selection was drawn from the Norwegian population between 18-70 years by the Norwegian Tax Administration. Since group 1 consisted of women only, more men than women were invited to group 2. This was done to obtain a more balanced sex ratio in the entire sample. Specifically, a total of 200 individuals comprising of a mix of both men and women, in addition to 100 men, were invited. Thus, a total of 300 individuals were invited. All were sent the invite by mail and then called within one to two weeks. Whenever possible, text messages or voicemail were used if the invited did not respond. If no contact was established, a new phone call was made again after a few days. Then new text messages or voicemail was used if needed. Inclusion criteria are described in detail in Paper IV.

All participants, in both groups, were informed both in writing and orally regarding the study.

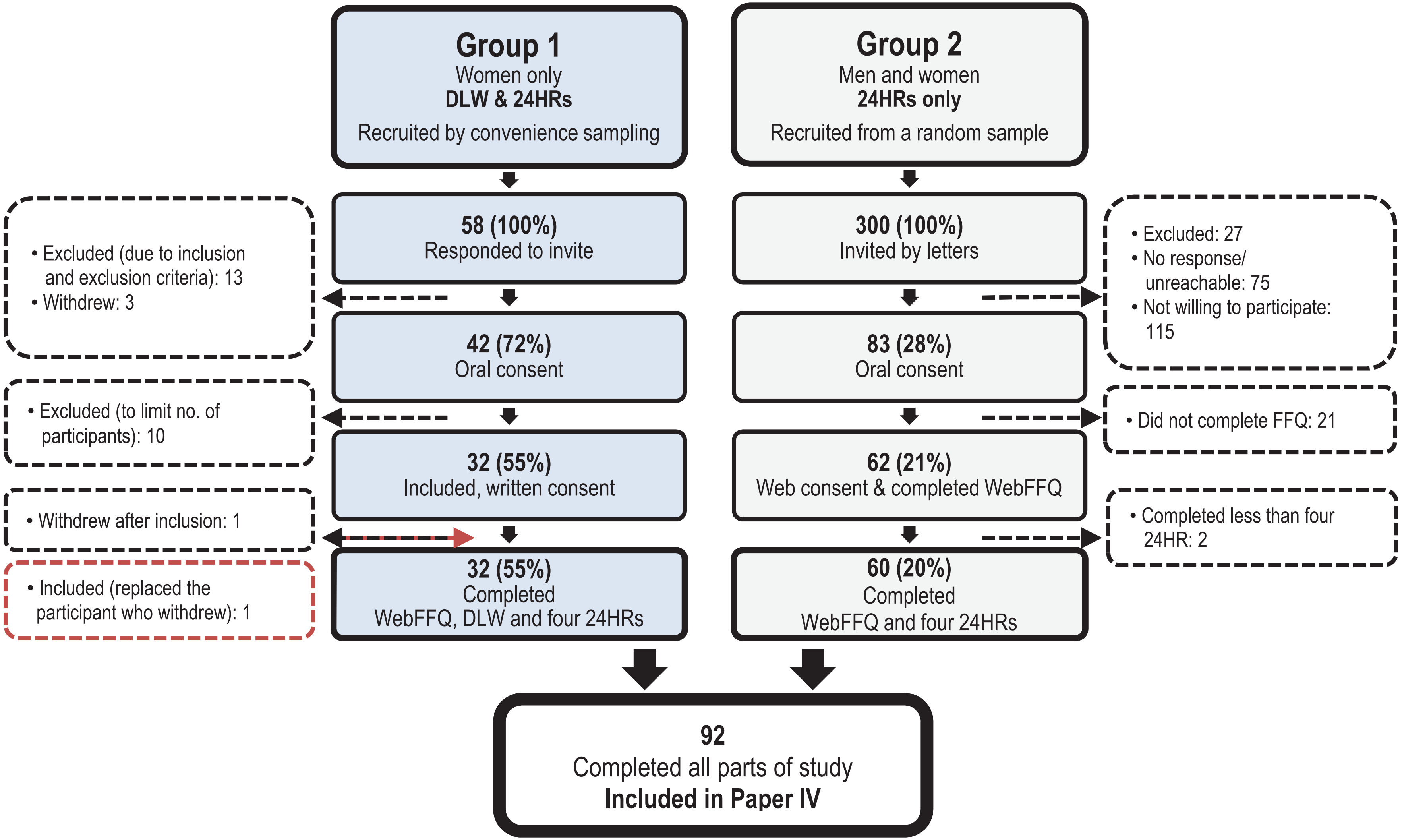


Figure 3. Flow chart showing the recruitment process in the validation study of the WebFFQ (Paper IV).
Adapted from Figure 1 in Paper IV.

2.1.4 Design of the WebFFQ validation study (Paper IV)

In this validation study, all participants started by completing the WebFFQ. The WebFFQ ask about the habitual diet, that is their average dietary intake over the last 12 months. Then, a total of four telephone-administered non-consecutive 24HRs were collected for all the individuals included in the study. The interviews were conducted by trained nutritionists. Additionally, TEE was measured by the DLW method in all participants in group 1.

The WebFFQ

The WebFFQ is a self-administered, web-based FFQ. It is designed to assess the habitual dietary intake, specified as the usual dietary intake over the last 12 months. Study participants access the WebFFQ by using a direct link sent to their email.

Researchers from the Department of Nutrition in addition to the staff at the University Center for Information Technology, both at the University of Oslo, developed the WebFFQ. It is based on former paper-based FFQs [132-136].

The WebFFQ includes 279 foods or beverages, typically consumed in an adult Norwegian population. Images are used to assist participants when they estimate portion sizes (Appendix II). Moreover, to reduce the burden on participants, skip-algorithms are used. Specifically, that means it is possible to skip entire food categories (i.e. meat-based dishes) when a participant ticks off the box for non-consumers for the particular food category. Due to automatic error detection, the WebFFQ evades the problem with missing data, which is a widespread issue when using paper-based FFQs. That is, one cannot proceed to the next page if there are any questions left unanswered. Questions regarding the characteristics of the participants, for instance, smoking habits or educational level, are included at the end of the WebFFQ.

Doubly labelled water (group 1 only)

The DLW method was used in the participants included in group 1, to measure TEE. The TEE was later compared to the estimated EI from the WebFFQ.

Participants in group 1 were all visited three times each, in their own home, during the study (Figure 4). The first visit was made after they had completed the WebFFQ. At visit one, they were given practical information and were provided with the equipment they needed for sampling and storage of urine samples. During the second visit, a pre-dose (baseline) sample of urine was collected from the participants, and they had their weight and height measured. At the end of visit number two, participants were given a dose of DLW. A multi-sample protocol over a period of two weeks was used, described more in detail in Paper IV. Finally, urine samples were collected, and new weight and height measurements were conducted during visit three.

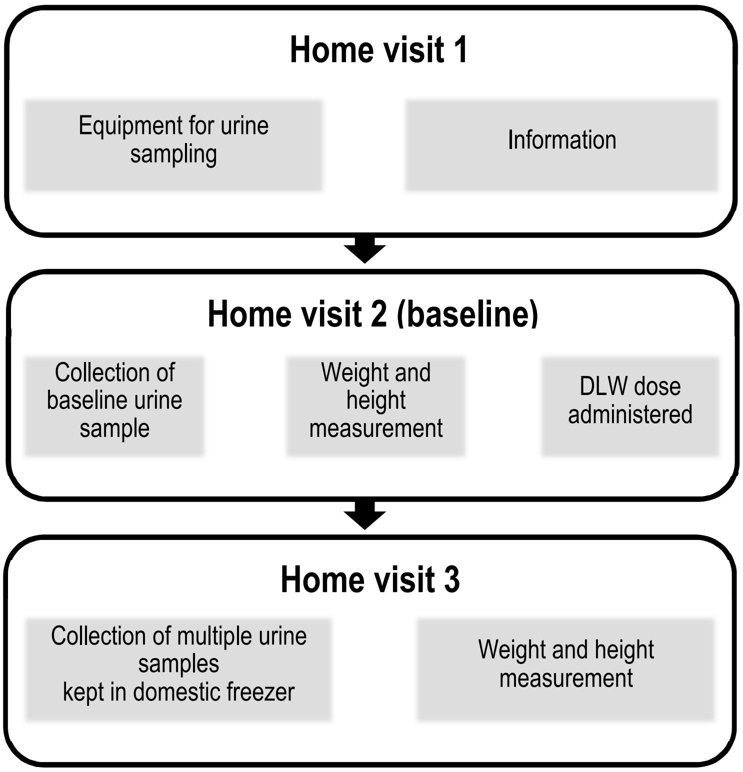


Figure 4. Overview of home visits in group 1 in the validation study of the WebFFQ (paper IV).

Anthropometric measurements (group 1 only)

A portable stadiometer (Seca 213, Seca GmbH & Co. KG., Hamburg, Germany) was used to measure the height of participants, to the nearest mm. A digital scale (TANITA TBF-300, Tanita Corporation, Tokyo, Japan) was used to measure weight, to the nearest 0.1 kg. All measurements were conducted in the morning, after a night's fast, in very light clothing or underwear.

Multiple 24HRs (group 1 and group 2)

Four non-consecutive 24HRs were completed for each participant by telephone, using the 24-hour multiple-pass recall module, integrated into the food and nutrient composition database and calculation system KBS, developed at the Department of Nutrition, University of Oslo, Norway [137].

The 24-hour multiple-pass recall module of KBS is designed to be used in a three-step sequence (Figure 5), resembling the approach of the United States Department of Agriculture's Automated Multiple-Pass Method [27]. In step one, the respondent describes what was consumed the previous day freely; that is, without being interrupted by the interviewer. In step two, the interviewer recaps everything the respondent reported, in chronological order. Moreover, the interviewer inquires about portion sizes, and probes regarding probably omitted items (e.g. sugar or milk, if tea is reported without specifying any details), or omitted meals or snacks. In the final third step, the interviewer prompts for foods, beverages and supplements, frequently omitted from recalls, using a pre-defined fixed list.

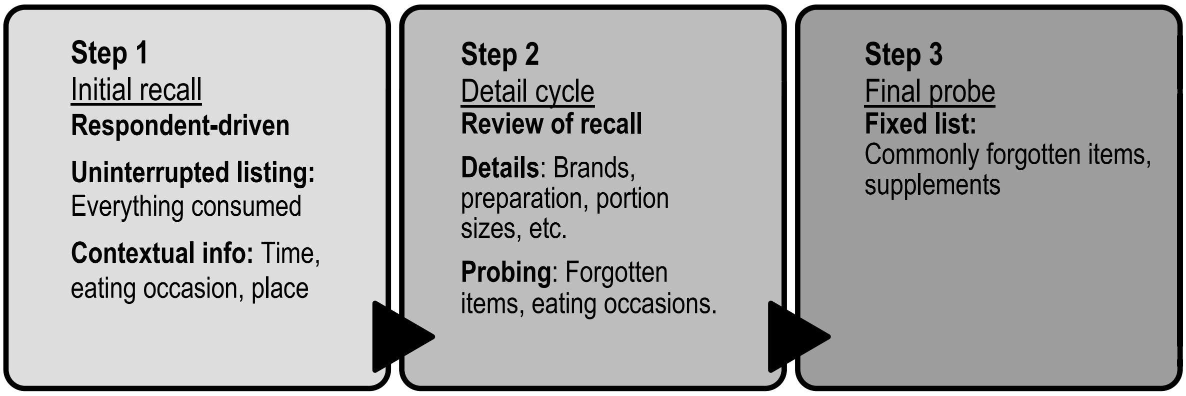


Figure 5. A description of the 'Interview-assisted and computer-based 24-hour multiple-pass recall module' of KBS, from the Department of Nutrition, University of Oslo.

During the interviews all participants had access to image series consisting of four images each, displaying different portions sizes of the same food, to ease the portion size estimations. The image series were available in paper format, or electronically, as a PDF file. Of the four 24HR-days, one was either a Friday, a Saturday or a Sunday. The interviews were, for the most part, not prearranged (93%), to avoid reactivity.

Self-reported weight and height (all)

All participants reported their weight and height in the WebFFQ.

Other subject characteristics (all)

The WebFFQ included questions regarding educational level, smoking habits and birth date. All participants in group 1 provided information about their physical activity level, over the phone, when they were considered for inclusion in the study.

2.2 Ethical statements

Paper I-III

The study was conducted in accordance with the Declaration of Helsinki. The Norwegian Data Protection Official for Research (NSD) approved the study (Project No. 32968). Child assent and written parental consent were obtained from all participants. All participants who completed the study were given two tickets to the cinema, in the form of a personal gift card.

Paper IV

The study was conducted in accordance with the Declaration of Helsinki and all procedures involving human subjects were approved by the Data Protection Official for Research in Norway (NSD), project numbers: 44876 and 45712. Written informed consent was obtained from all participants. No economic compensation or incentives were given to the participants.

3 Summary of papers

Paper I:
Evaluation of a web-based food record for children using direct unobtrusive lunch observations: A validation study

<u>Aim:</u> To assess how accurately children could record their school lunch using the WebFR. The reference method used for comparison was direct, unobtrusive observation during school lunch.

<u>Subjects and setting:</u> Children, 8-9 years old (n=117), from Bærum, Norway. Data was collected between September-December 2013.

<u>Methods:</u> Participants recorded their dietary intake for four consecutive days in the WebFR, assisted by their parents/guardians, and were observed in the same period, while eating their lunch at school. Three observers conducted all observations. IOR was assessed and found satisfactory. Data from observations was compared to the participants' school lunch recordings, and variables for 'matches', 'omissions' and 'intrusions' were constructed. 'Matches' are defined as foods/beverages both observed being consumed and recorded in the WebFR; 'omissions' are defined as foods/beverages observed being consumed, but not recorded in the WebFR; 'intrusions' are defined as foods/beverages not observed being consumed, but recorded in the WebFR. Match rates ((matches/observed eaten foods)*100), omission rates ((omissions/observed eaten foods)*100), and intrusion rates ((intrusions/(recorded eaten foods)*100), were calculated. These rates were calculated to evaluate to what degree the participants were able to register their school lunch in the WebFR correctly. Rates were calculated separately for food categories, and for all foods/beverages combined. Moreover, a logistic regression analysis was conducted to examine whether body mass index (BMI), parental educational level, parental ethnicity or family structure were associated with a 'Low match rate', defined as ≤70%. Excel (version 2010, Microsoft Excel), IBM SPSS (version 21.0, 2012, IBM Corp.) and R (version 3.0.1., 2013, The R Foundation for Statistical Computing) were used to create the variables, and in the analyses.

<u>Results:</u> The average match, omission and intrusion rates varied widely between food categories. Recording accuracy was better for bread products and milk as compared to spreads, fruit, berries, vegetables and salads. For all food groups combined, the mean match, omission

and intrusion rates were 73%, 27% and 19%, respectively. We observed that parental educational level and parental ethnicity were associated with match rate. Specifically, the mean match rate was 52% for children of the lower educated parents, versus 77% for children of the higher educated parents ($p< 0.01$). Moreover, the mean match rate was 57% for children of two non-Norwegian parents, versus, 75% for the others ($p=0.04$). In the logistic regression model, only parental ethnicity remained statistically significant, with an adjusted odds ratio of 6.9, and 95% confidence interval between 1.3- 36.4. Nevertheless, the parental educational level variable was borderline significant with an odds ratio of 3.8.

Conclusions: We have demonstrated that some of the 8-9 year-old children included in the current study were not able to record their dietary intake from school lunch adequately. Lower parental educational levels and having two non-Norwegian parents were linked to more recording errors. However, these findings must be interpreted with caution, due to the low number of participants in the subgroups with these characteristics. The WebFR seems to be in line with other web-based tools for children. By including additional prompts for foods that had high omission rates, we may improve the WebFR. We suggest that participants with language difficulties may benefit from extra support and information in future studies using the WebFR.

Paper II:
Associations between reported intakes of carotenoid-rich foods and concentrations of carotenoids in plasma: a validation study of a web-based food recall for children and adolescents

Aim: To validate the recorded intakes of carotenoid-rich foods in the WebFR. Measured concentrations of carotenoids in blood, converted to plasma values, were used as an objective reference method.

Subjects and setting: Children and adolescents, in age groups 8-9 years and 12-14 years (n=261), from Bærum, Norway. Data was collected between September-December 2013.

Methods: All participants used the WebFR to record their dietary intake for four consecutive days. Within 11 days after completing the recordings, a few drops of blood from the fingertip were collected from all, using the DBS method. Concentrations of carotenoids (β-carotene, α-carotene, β-cryptoxanthin, lycopene, lutein and zeaxanthin) were analysed using standard procedures of high-performance liquid chromatography. The carotenoid-rich food variables that were created, comprised foods with a significant content of carotenoids, that had been consumed in the current study. Cross-classifications and Spearman's rank correlations were used to assess the relationship between concentrations of carotenoids from the DBS and the recorded intake of foods with a high content of carotenoids. Excel version 2010 and KBS (database AE-10, version 7.2 Department of Nutrition, University of Oslo, Norway), and IBM SPSS Statistics Version 21.0 (2012) were used to create the variables and in the analyses.

Results: The median recorded consumption of vegetables, fruits and juice combined was 225 grams/day, and the median intake of all carotenoid-rich foods was 81 grams/day. Data from all participants, on recorded dietary intakes of carotenoid-rich foods and the concentrations of the corresponding carotenoids in plasma, showed Spearman's correlations between 0.30 - 0.44. Moreover, we observed that 72–77% of all participants were classified in the same or adjacent quartile if the results of lutein and zeaxanthin were excluded. The correlation between recorded intakes of vegetables and total carotenoids in plasma were significantly different between 8–9-year-olds (r=0.47) and the 12–14-year-olds (r=0.14).

Conclusion: The ranking abilities of the WebFR were acceptable for foods rich in α-carotene, β-carotene, β-cryptoxanthin and lycopene, in a sample of children and adolescents. The WebFR is a suitable tool to assess the intake of foods rich in carotenoids, especially in the age group 8-9 years.

Paper III:
Validation of energy intake from a web-based food recall for children and adolescents

Aim: To validate estimated EI from the WebFR, by comparing EI to estimated TEE, calculated from accelerometer counts, in combination with data on sex and body weight, or combined with REE equations.

Subjects and setting: Children and adolescents in the age groups 8-9 years and 12-14 years (n=253), from Bærum, Norway. Data was collected between September-December 2013.

Methods: Participants recorded everything they consumed for four consecutive days in the WebFR, and their physical activity was measured using an accelerometer (ActiGraph GT3X+) for seven consecutive days, during the same week. Counts from the accelerometer were used to calculate the individual physical activity level for all participants. REE was calculated, based on age, sex and measured weight and height. Subsequently, three different equations were used to calculate TEE. They were based on either accelerometer counts and sex and weight specific equation, or accelerometer counts, and a sex-specific equation and REE, or REE and physical activity level. EI was estimated from recordings in the WebFR. Pearson's correlation between EI and TEE was calculated. The proportion of acceptable-, under- and over-reporters of energy was calculated using two different approaches. A Bland-Altman plot was created to assess the agreement between EI and TEE. Also, a linear regression analysis was used to see which variables contributed to the misreporting of EI. ActiLife (version 6.0, ActiGraph LLC, Pensacola, FL, USA), MS Excel (version 2010, Microsoft, Redmond, WA, USA) and IBM SPSS (version 22.0, 2013, IBM Corp, Armonk, NY, USA) were used for the calculations and the analyses.

<u>Results:</u> The mean EI for all participants was 6.85 MJ/day, and the mean TEE was 8.67 MJ/day. More than one-third (36-37%) were defined as under-reporters of energy, but only 2-4% were identified as over-reporters of energy. Pearson's correlation was 0.16 for the entire sample, 0.31 for the 8-9-year-olds, and 0.08 for the 12-14-year-olds. The mean EI was under-reported by -1.83 MJ/day, for the complete sample. In a multiple linear regression model, increased energy under-reporting was observed for overweight and obese participants, the oldest age group, boys, those with parents/legal guardians with a low educational level, and those living in a non-traditional family. Among these variables, weight status showed the strongest association with misreporting of energy: Participants who were either overweight or obese underreported their EI by -2.35 MJ/day more as compared to those with a normal body weight.

<u>Conclusion:</u> Estimated EI from the WebFR was significantly underestimated. The degree of underestimation was affected by weight status, sex, age, parental educational level and family structure. EI from the WebFR should be used with caution in children and adolescents.

Paper IV:
The validity of a web-based food frequency questionnaire assessed by doubly labelled water and multiple 24-hour recalls

<u>Aim:</u> To assess the validity of the estimated habitual dietary intake from the WebFFQ, using the DLW and multiple 24HRs as reference methods.

<u>Subjects and setting:</u> A total of 92 adults, born in Scandinavia and living in Norway, were included in the study. Data collection was conducted in the period between January-December 2016.

Methods: All participants completed the WebFFQ in addition to four non-consecutive 24HRs. In a subsample of 32 women, TEE was measured using the DLW method for comparison with their estimated EI. The relative validity of the WebFFQ's estimated intakes of macronutrients and food groups was assessed in the entire sample (n=92), by comparing them to the estimated mean intakes from the 24HRs. We used various techniques in these validity assessments, including calculating absolute differences, ratios, crude and deattenuated correlations, cross-classifications, and a Bland-Altman plot. Besides, we plotted the misreporting of energy (EI-TEE) against the relative misreporting of food groups (WebFFQ-24HRs). KBS (KBS, version 7.3, database AE14, University of Oslo, Oslo, Norway), IBM SPSS (version 22.0, 2013, IBM Corp, Armonk, NY, USA) and MS Excel (version 2010, Microsoft, Redmond, WA, USA) were used for all the calculations and the analyses.

Results: The EI estimated by the WebFFQ was not significantly different from the TEE measured by DLW on group level (0.7 MJ/day). However, ranking abilities were poor (r=-0.18). The relative validation showed an overestimation by the WebFFQ for the majority of the variables using absolute intake; especially the food groups 'vegetables' and 'fish and shellfish' were largely overestimated. We observed an improved agreement between the test and the reference tool for energy-adjusted intakes. Deattenuated correlation coefficients were in the range 0.22-0.89. Moreover, for the majority of the energy-adjusted variables for macronutrients and food groups, we observed low levels of grossly misclassified participants (0-3%).

Conclusion: The WebFFQ is not able to rank individuals correctly according to their reported EI and is therefore not suitable to estimate EI at the individual level. For the energy-adjusted macronutrients and the majority of the energy-adjusted food groups, both the estimated intakes on group level and the ranking abilities seem acceptable. Consequently, the WebFFQ appears to be a suitable tool for both future dietary surveys and nutritional epidemiology studies. Nevertheless, there is a need to confirm the results from the relative validation using objective reference methods.

4 Discussion

In this thesis two web-based dietary assessment tools are validated in different samples, using several different reference methods with varying strengths and weaknesses. Relevant methodological considerations are discussed in section 4.1, and a discussion of the main findings is found in section 4.2. Finally, the advantages of web-based tools versus traditional dietary assessment tools are discussed in 4.3.

4.1 Methodological considerations

4.1.1 Selection bias and external validity

Selection bias arises when the study sample is not representative of the target population about which conclusions are to be drawn [138]. Such bias is a challenge in dietary surveillance studies, with the potential to lead to distorted estimates of a populations' intake, caused by deviations between estimates from the non-representative sample, and the true population's intake. Trying to avoid selection bias is equally important in validation studies, because an unrepresentative sample may report their intake more accurately, or eat differently, than the target population [123]. One may for instance, mistakenly conclude that a dietary assessment tool gives a valid estimate of energy, although this could primarily be due to an overrepresentation of individuals able to accurately estimate their intake, compared to the target population. The validity of a tool will, therefore, vary due to the study sample's characteristics. A randomly selected representative sample from the target population is, therefore, the ideal, but unfortunately not always obtainable in validation studies.

Paper I-III

The primary target populations of the WebFR were 4th graders (8-9 years) and 8th graders (12-13 years) in Norway, as the WebFR was first and foremost developed for use in a national representative dietary surveillance study (UNGKOST 3) in these age groups [130]. For the validation study of the WebFR, convenience sampling was used to select the schools from which children in the target age groups were invited. The municipality, from which the schools were chosen, was selected due to its proximity to the University of Oslo. However, attempts were made to counter selection bias, by selecting schools from an area in the municipality known to have a relatively diverse population in terms of ethnicity and socio-economic status. All children at the selected schools were invited, and participation rates were relatively high (>60%). Compared to the general Norwegian population, the ethnic diversity was high in our sample. In Paper I-III, a total of 19-23 % of all parents/guardians of participants were non-Norwegian, and 9-13% of the participants had two non-Norwegian parents/guardians. In comparison, according to Statistics Norway, in 2013, a total of 14% of the population were either immigrants or Norwegian-born to immigrants [139]. The parental educational level was high in the included sample. A total of 71-74% of all parents/guardians of the recruited participants had higher education (university college or university), whereas the average in the overall Norwegian population, in the age group 25–59 years, was 42% in 2014 [140]. Finally, the proportion of overweight and obese children and adolescents was slightly lower for the sample included in Paper I-III, as compared to children and adolescents from the general population in Norway (13-14% versus 16%) [141]. The population's educational level in Norway is higher in typical urban or semi-urban areas, as compared to rural areas in Norway. Specifically, the proportion of individuals over 16 years who have higher education, and who live in one of the four largest cities in Norway or the surrounding areas of the capital city Oslo, is in the range of 41-52% [142]. In comparison, this proportion is in the range of 13-30% for 324 out of the 337 smaller municipalities with <10.000 inhabitants [142]. The proportion of immigrants is especially high in Oslo and the surrounding municipalities of the capital city (17-30%)[139]. Moreover, a study which included a nationally representative sample of 8-year-olds showed that the mean BMI was higher among children living in rural areas as compared to urban areas in Norway [143]. In conclusion, the sample included in Paper I-III resembles the population in urban- and semi-urban parts of Norway. Combined with the relatively high participation rates obtained in the study, it can be argued that the generalizability to the general population, living in semi-urban or urban parts of Norway, is good, resulting in an acceptable external validity.

Paper IV

The target population, for which the WebFFQ is developed, is the Norwegian adult population between 18-70 years of age. In the validation study presented in Paper IV, the recruitment of participants was done at two different time points, using different approaches. This gave rise to group 1 and group 2, described previously in Figure 3.

Group 1 had to accept a relatively high participant burden, involving a multi-sample DLW protocol. Thus, it was decided to approach and recruit highly motivated participants, with a high likeliness to complete all parts of the study. Convenience sampling was used, recruiting through social media, word of mouth and posters. Moreover, several, but important criteria for inclusion have undoubtedly contributed to a reduced representativeness of the sample. Attempts were made to increase the variability of the sample, by selecting participants that differed most, based on age, self-reported height, weight and physical activity level, in addition to the area they resided, as described previously. The characteristics observed for the 32 participants in group 1 diverged from the general population: they consisted of women only, had a high educational level (84%), and only 6% were current smokers. In comparison, 18% of the overall population of adult women in Norway were current smokers (11% daily and 7% occasional) in 2016 [144]. One may speculate if the low number of smokers indicates that the sample in group 1 may consist of more health conscious participants than in the general population, perhaps more susceptible to report their food intake in a socially desirable way.

For group 2, the participation burden was moderate, as compared to group 1; as a consequence, a representative sample was invited. However, despite that the invited individuals received a personal letter, in addition to personal follow-up phone calls, the participation rate was low (20%). Therefore, as for group 1, we cannot rule out selection bias. Participants in group 2 consisted of 57% men and 60% had higher education. This is a high proportion as compared to the general adult population in Norway in 2016, in which 36% of 20-year-olds, or older individuals, had higher education [142]. Moreover, only 13% in group 2 were current smokers, versus 21% of the general adult population (12% daily and 9% occasional) [144]. Because of that, the sample in group 2 seems to deviate from the target population in the same direction as group 1, but to a lesser extent.

4.1.2 Internal validity: Strengths and limitations of the reference methods

Direct observation

The advantage of using direct observation to assess dietary intake is that it allows the observers to get insight about the true intake, not obtainable by any other method. Observations of packed lunches, used as a reference method in this thesis (Paper I), is however believed to be challenging because of the diversity of items and portion sizes, in addition to the challenge with opaque drinking bottles or food containers, compared to observations of served school meals. Nevertheless, it has been demonstrated that accurate and reliable observations of packed school lunches are attainable when using trained observers with a background in nutrition [116].

If direct observations are to be used as an unbiased reference method, reactivity must be limited as much as possible, to avoid drawing misleading conclusions regarding the performance of the test method. To minimise reactivity, we must conduct observations unobtrusively [113]. One may argue that truly unobtrusive observation, can only be done behind a glass mirror, or similar [115]. In Paper I, avoiding reactivity was attempted both in the planning phase and during the data collection, using the following strategies: Familiarising participants to the observers before the data collection, the use of single blinding, and avoidance of contact with the children during school lunch. Others have demonstrated that children observed during school lunch are not affected by the observations [145, 146]; thus reactivity was most likely limited in Paper I.

The observations were restricted to school lunch, during weekdays only, in Paper I. This is a limitation, as we cannot rule out the possibility that the meals under observation in the current study may not be representative. Specifically, the recording accuracy for other types of meals and eating occasions, at other times of the day, and week, may be different from the recording accuracy for the school lunch. To what degree observers can truly observe what is eaten, may be affected by the number of individuals under observation at the same time, per observer. In Paper I, this was limited to a maximum of three per observer, which is found to be acceptable elsewhere [147]. Moreover, when using several observers, it is important to assess the agreement between them during the study, using IOR [114]. An agreement of 85% is often put forward as the lowest acceptable limit of IOR [114]. In the current study (Paper I), IOR was assessed before and continuously during the data collection. An IOR of 92% was obtained for all food items on average, ranging from 88-96% between different observer pairs. This secured a minimum of standardisation, resulting in an acceptable internal validity.

Concentrations of carotenoids in plasma (Paper II)

Concentrations of carotenoids in plasma were used as objective measures of dietary exposure of foods rich in carotenoids in Paper II. As explained previously, concentration biomarkers cannot be used to quantify the absolute intake, but may be used to rank individuals according to their intake. Additionally, between-person variation in the concentration of these biomarkers can reflect other factors than the variability of the dietary exposure of the biomarker itself. That is, individuals with the same dietary exposure may vary in the way they absorb and metabolise a biomarker [84]. Many factors can affect the concentration of carotenoids in blood. A meal high in fat increases absorption, and between-person variation in postprandial metabolism can influence the concentration [148]. Moreover, overweight and obesity are linked to lower levels of plasma carotenoids [149], and it has also been reported that inflammation may reduce β-carotene concentrations [150]. None of these factors were taken into account in the validation study of the WebFR. Thus it cannot be ruled out that they may have contributed to a possible misclassification of participants in Paper II.

The time of exposure reflected by the biomarker is of high importance [84]. Carotenoids in plasma and serum are observed to have a half-life between 1-11 weeks [101, 151, 152]. Thus they reflect the dietary exposure of carotenoids over the last weeks. In this study (paper II), blood samples were drawn once, within 11 days upon completion of the recording of dietary intake in the WebFR. Thus, the discrepancy in time between the sampling and recordings is a limitation, and ideally, more than one sample should have been drawn, to assess and address any within-person variation.

Accelerometer derived TEE (Paper III)

There is a large assortment of accelerometers available; not all are validated. In a recent review, Jeran et al. identified studies that used accelerometers to estimate activity-related energy expenditure (and thus indirectly TEE) in adults in a free-living setting and compared it to DLW data [117]. The results varied largely, from poor to good estimates; thus the authors concluded that accelerometer derived energy expenditure should be used with caution [117]. ActiGraph is the type of accelerometer used in the current thesis (Paper III). TEE calculated from accelerometer counts from the ActiGraph (former known as CSA/MTI) has, however, shown to correlate reasonably with TEE measured by DLW; specifically, correlation coefficients in the range of 0.68-0.96 have been observed in children and adolescents [153].

The mean of three different approaches to estimate TEE was used in Paper III, to make our estimates more robust. This is based on the principle 'wisdom of select crowds' [154], which suggests that averaging a few carefully chosen estimates based on expertise, will often perform superior to a single estimate taken as the best, as errors randomly distributed between uncorrelated estimates will cancel each other out by averaging. This assumption of uncorrelated estimates has probably been violated to some degree in Paper III. However, because aggregate prediction equations, in other research areas, have outperformed single prediction equations [155], this approach was used in the current thesis.

There are several limitations to consider. First, the appropriateness of the equations and cut-points used may be questioned. The equations used from Ekelund [120] were developed for 9-year-old children, and may be less appropriate for the older age group included in Paper III. Besides, there is a lack of consensus regarding what thresholds should be used for different physical activity intensities [156]. Secondly, it is well known that accelerometers are unable to register certain activities, involving static movement and weight bearing activities, and they cannot discriminate between walking and walking up a hill [156]. This may lead to underestimations of physical activity. The placement of the sensor is also of importance. Placing the accelerometer on the hip, as done in Paper III, captures major body motions, but tend to underestimate activities like upper-arm movements or cycling [157].

Reactivity is a potential challenge, as participants may move more than normal, not providing information on the usual activity pattern and overestimating the usual physical activity level. Participants included in Paper III started wearing the accelerometer 24 hours before the start of the recordings. Thus the most immediate alterations in activity have been omitted. Additionally, to avoid reactivity, it is of importance that the accelerometer is small, robust, to not interfere with the activity of the individual [153]. The ActiGraph used in Paper III satisfies these requirements.

In individuals with a stable weight, there is a good agreement between EI and TEE [158]. Therefore, estimates of reported EI can be compared to estimates of TEE [159]. In the validation study of the WebFR, the height and weight of the young participants were measured only once during data collection. Therefore, the assumption of weight stability may have been violated. However, the period under study was limited to one week, and any substantial weight alterations seem implausible.

The DLW method (Paper IV)

The DLW method, used in a subsample in Paper IV, is regarded as the best technique to obtain optimal measurements of individual TEE in normal free-living conditions [93]. Still, the DLW method can be troubled by random error [160]. Validation studies of the DLW method, in which indirect calorimetry was used as the reference method, have demonstrated a difference between methods of +/- 1-3% on group level [92]. On the individual level, the average difference between methods has been found to be substantially larger (around 10%) [92]. In addition to errors caused during analysis in the laboratory, other factors can contribute to errors in the DLW methods. Any spillage of dose during administration, or mixing up which dose is given to which subject, will introduce errors. Also, the background enrichment of the naturally present isotopes can change rapidly [161] and may distort samples at the end of the sampling period, in which the enrichment of body water is approaching background levels. The type of medium used is also of importance. Blood is regarded as the preferred medium, which reflects the exact enrichment of body water at the time of sampling [162]. In the study included in this thesis, urine was used to sample body water, to reduce the burden of the participants. Urine in the bladder is not in equilibrium with the rest of the body water and is rather reflecting the period since the last void of urine [162]. To complicate it further, if the bladder is not emptied properly during urination, the retained urine will mix with the newly produced urine, and distort the samples, which typically is an issue in older individuals [162]. To keep the errors to a minimum, we used a strict sampling protocol, and a highly specialised lab analysed all samples. Additionally, as a multi-sample protocol reduces the analytical variation and improves the precision of the method [93], this approach was used for the DLW assessment of TEE, in Paper IV. Still, three participants had to be excluded due to invalid TEE estimates; one of them most probably due to a minor spillage of dose.

A premise for the use of TEE as a proxy for EI is weight stability of participants, in the period under study. This is why body weight was measured both in the beginning and after the DLW sampling period, in the WebFFQ validation study. No significant weight change was observed, which indicates that the participants were in energy balance, giving reason to rely on TEE estimates, as a marker of EI. Nevertheless, a limitation is that the DLW data reflects the TEE only in the short period under study, whereas the estimated EI from the WebFFQ, to which the TEE was compared, reflects the habitual intake. To obtain the habitual TEE for an individual, we would need multiple assessments of DLW [95]. This was not feasible, and an attempt to bypass this limitation was therefore made, by selecting individuals with a history of being weight-stable, indicative of fairly stable mean TEE and EI.

Multiple 24HRs (Paper IV)

It is not always feasible to use objective reference methods. Other dietary assessment methods, based on self-reports, are therefore often being used as reference methods in validation studies.

This was also done in Paper IV, despite the limitations of these reference tools. Willet has stated that dietary records are the best alternative for comparison with FFQs, due to the fact that dietary records share the least number of errors with the FFQ, out of all dietary methods [95]. Compared to dietary records, errors in the 24HRs are more likely to be correlated with the FFQ, because both methods rely on memory, and portion size estimation [95]. However, one may argue that dietary records are still not the best option, as they are subject to reactivity [163], and will typically result in under-eating during the days of recording. Such reactivity, in the form of under-eating, may distort the results from the validation studies largely, possibly in the direction of attenuating the agreement. In the current study, quadruple non-consecutive recalls were collected, out of which 93% were not pre-scheduled. The number of non-consecutive recalls obtained per participant, and the use of unannounced recalls, which are not subject to reactivity [164], are strengths of Paper IV. However, the period, which the dietary intake from the WebFFQ and the 24HRs reflects, does not coincide, which is an obvious limitation.

The validity of the estimated EI from the mean of four non-consecutive 24HRs, used as the relative reference tool in Paper IV, was also assessed using TEE measured by DLW. A deattenuated Pearson's correlation coefficient of 0.34 was observed between DLW-derived TEE and EI from the mean of four 24HRs. Moreover, EI from the mean of the four 24HRs was on average underestimated by 17%, compared to DLW-derived TEE. These results show that despite underestimation of EI on group level, the 24HRs could correctly rank individuals according to their true EI far better than the WebFFQ, which strengthens the 24HRs' value as a reference tool.

4.2 Results in context: Discussion of the main findings

4.2.1 Energy intake from dietary self-reports (Paper III, IV)

The validity of the EI estimated from the WebFR in children and adolescents, and the EI estimated from the WebFFQ in adults, was assessed by TEE derived from accelerometer counts and DLW, respectively. Keeping the previous discussion on methodological considerations in mind, the TEE derived from DLW is superior to TEE from accelerometer counts. Results showed that the EI was estimated with substantial errors by both tools. Low abilities of ranking were observed: Correlation coefficients between EI and TEE were 0.16 and -0.18 for the WebFR and the WebFFQ, respectively. Only the WebFFQ provided a fair estimate of EI on group level, given the non-significant underestimation of 0.7 MJ/day (6%). However, the group mean was caused by large over- and under-reporting of EI at the individual level that cancelled each other out. The EI estimated from the WebFR was underestimated by 1.8 MJ/day (23%), on average across all individuals. Thus, the findings in this thesis, demonstrate that both the WebFR and WebFFQ provide poor individual estimates of EI, and the WebFR provides poor group level means as well. Such poor results are in agreement with findings reported by others. In a systematic review by Burrows et al., the validity of estimated EI from different dietary assessment tools in children was compared to DLW data [165]; results showed that underreporting was common, ranging from 19-41% for food records, whereas over-reporting of EI was seen for 24HRs [166]. Thus, the WebFR resembles the foods records, more than recalls in this respect. In the study by Freedman et al., who pooled results from several large validation studies using DLW in adults [167], they demonstrated that assessment of self-reported EI was largely troubled with energy misreporting: The deattenuated correlations for women were between 0.11-0.34.

In the wake of the study by Freedman et al. [167], recommendations have been put forward, advising that self-reported EIs should not be used to derive estimates of absolute EI [163, 168]. Subar et al. argue that the reason for this is that EI is especially prone to measurement errors, as many errors, both small and large, add up because most of what is consumed contain energy [163]. Moreover, Subar et al. further argue that especially FFQs are not suitable for measuring EI, because of their finite lists of items and low level of detail, which also applies to any web-based FFQ, like the WebFFQ, validated in this thesis. With regards to the WebFR, the element of recall, introduced by instructing the participants to record their dietary intake at the end of each recording day, may have increased the number of omitted items in Paper I and contributed to the under-reporting of energy observed in Paper III. This proposition fits well with observations done by Baxter et al., which showed clearly how reporting accuracy decreased as time after the eating event increased [169].

By adapting the WebFR for use on portable platforms, like smartphones or tablets, we would enable real-time recordings, avoiding the problems with memory related to recalls. A validation study of My Meal Mate, which is an example of such digital real-time record for smartphones for adults, showed good agreement on the group level, but not individual level, when compared to 24HRs [66]. Shifting to real-time recordings is not an option for the WebFFQ. Besides, the harsh critique, targeting FFQs in particular, makes it relevant to consider whether turning to objective measures of EI is the better solution. An alternative to self-reported EI exists through the DLW method. But as it is costly to use, and as resources are often a limiting factor, the DLW method will probably prove difficult to implement in large-scale studies. Using the DLW method in sub-samples of a study population is, however, a feasible option. This will also provide the opportunity to correct for measurement errors in self-reported EI, in regression models used in nutritional epidemiology, by applying the statistical approach regression calibration [170]. However, when doing so, repeated measures of the biomarker is necessary to take the variance of the biomarker into account in the model [171]. Moreover, it has been demonstrated that by adjusting other dietary self-reported data using self-reported EIs, the estimates improve. This was seen in Paper IV, in which intakes on group level and ranking abilities of the WebFFQ relative to the 24HRs for the energy providing nutrients and foods groups improved for the energy-adjusted intakes. For vegetables, 8% of individuals were classified in the extreme opposite quartile for crude intakes, whereas this was reduced to 2% for the energy-adjusted intakes. In line with this, others have demonstrated that nutrient densities of protein from an FFQ were stronger correlated with the measured true intakes, as compared to the absolute protein intake [122]. In conclusion, future estimates of EI from the WebFFQ and WebFR are not useless but must be applied appropriately.

4.2.2 Omission and intrusions in school lunch entries by 8-9-year-olds (Paper I)

The observations of school lunch in Paper I provided insights into what was truly eaten, and what was omitted and intruded in children's school lunch entries in the WebFR. Across all individuals for all food groups combined, the omission rate was 27%, and the intrusion rate was 19%, which confirms the observed underreporting of energy in Paper III.

Baxter (former Domel) and co-workers have published results from numerous validation studies of American schoolchildren's records and recalls from 1994 and up till 2017, using observation of school meals as the reference tool [172, 173]. Among the studies from Baxter et al., the ones most suitable for comparison are same-day-recalls, because the time of recordings in these same-day-recalls coincides with the time of recordings of the WebFR. Omission rates in these same-day-recalls are in the range of 27-56%, and intrusion rates are between 8-39% [174-179]. Hence, with a few exceptions of intrusion rates, the results from these studies are in line, or considerably worse, than the results in Paper I. The participants in Paper I were assisted by parents or guardians because 8-9-year-olds are believed to need assistance [180]. However, children this age were not assisted by their parents in the studies by Baxter et al. This difference between studies may partly explain why recording accuracy was better in Paper I. Nevertheless, one can speculate if parental assistance is less useful for school meals. Parents do not necessarily know what their child eats at school. During the observations in Paper I, trading of foods between the children occurred occasionally. Also, some children did not drink their school milk, or ate the food provided by their caretakers; they poured the milk down the sink or threw their packed lunch in the bin.

There are a few validation studies of self-administered web-based 24HRs, in addition to the WebDASC, which the WebFR is based on, that have used school lunch observations as the reference method. These studies are discussed in detail in Paper I and show that the WebFR seems to be at least as good as most other comparable tools. For example, the WebFR had lower omission rates and lower intrusion rates than what was found in the studies of ASA24-KIDS [40] and CAAFE [181]. However, children were not assisted by their parents in these two studies. In contrast, reporting accuracy in the validation study of the WebDASC [182], was considerably better than in Paper I; omission rates were in the range of 0-5%. This may be explained, not only by the different approach used to calculate these rates but by the fundamentally different way the observations were conducted in the WebDASC-study. The observations appear to have been far from unobtrusive, possibly resulting in an improved reporting accuracy.

The observations in Paper I pinpoint the sources of misreporting of what was consumed during school lunch. In Figure 6, an overview of omitted and intruded items is shown.

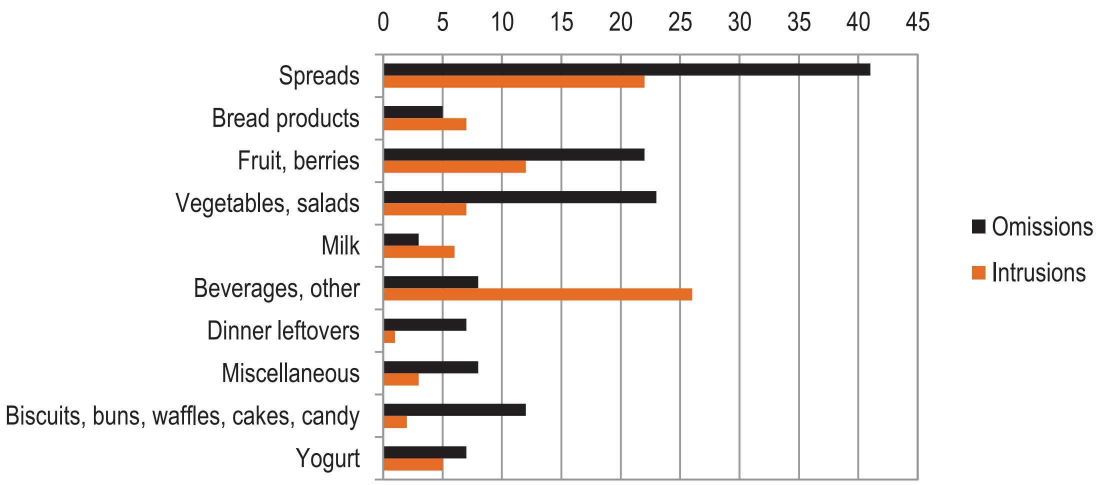


Figure 6. Number of omitted and intruded food items out of 495 observed items and 450 recorded items, in 8-9-year-olds

A typical school lunch for children in Norway includes open-faced sandwiches with spreads, brought from home [183]. It is interesting that spreads, but not bread products were often omitted. Moreover, only some of the omissions of spreads had coinciding intrusions. The remarkably many intrusions of 'beverages, other', are simply water (96%). As a consequence of these discoveries regarding omissions, the WebFR has been slightly altered, to include additional prompts for spreads, to remind children to record these commonly omitted items.

4.2.3 Fruit and vegetable intakes from young individuals' self-reports (Paper I, II)

The WebFR's ability to assess carotenoid-rich foods was assessed in Paper II. The carotenoid-rich foods included in that paper, are fruits and vegetables, or constitutes of foods and beverages derived from fruits and vegetables (e.g. ketchup). Not many validation studies of dietary assessment tools for the younger age groups have used carotenoids in plasma as an objective reference method to assess the accuracy of reported intakes of fruits and vegetables. In Table 3, results from the WebFR are compared to a few relevant studies, covering the same age groups as in Paper II.

The comparison in Table 3 shows that the WebFR is in line, and even seems to outperform some, but not all of these other studies. Both non-fasting and fasting samples, plasma and serum are used in the studies included in Table 3. This is not a hindrance for comparison, because carotenoids in plasma and serum will, for the most part, reflect the habitual intake of carotenoid-rich foods, as the half-life of carotenoids is weeks, not days [101, 151, 152]. Moreover, serum and plasma values of carotenoids may be used interchangeably [184]. However, five out of the seven other assessment tools included in this comparison with the WebFR are FFQs. Moreover, both the paper of Burrow et al. [149] and Byers et al.[185] are based on parental reports, and the study populations are dissimilar. This makes the comparison difficult. For example, correlations between plasma or serum β-carotene and reported intakes of vegetables are reported in both Paper II, Biltoft-Jensen et al.'s study [182], and Slater et al.'s study [186], and are in the range of 0.07-0.38. That is, Slater et al.'s 24HR had the lowest correlation, followed by their FFQ, the WebFR, and finally the WebDASC. Age of the study populations differs in these studies, and could perhaps explain part of the differences in results. But these study populations also differ in other essential ways. Slater et al.'s study is from Brazil, and the studies of the WebDASC and WebFR are both from Scandinavia. Dissimilar populations are not expected to eat the same types of vegetables. Thus, we are comparing associations between the intake of different kinds of vegetables to β-carotene in plasma or serum, without taking into account that the different vegetables may contain highly different levels of β-carotene [102]. Nevertheless, considering the results in Paper II independently, the WebFR has acceptable ranking abilities for foods rich in carotenoids (r >0.30), except for foods rich in lutein and zeaxanthin (r= 0.20).

Table 3. A comparison of validation studies of dietary assessment tools for school-age children and adolescents using concentrations of carotenoids in plasma or serum as a marker of intake of carotenoids, or foods rich in carotenoids

Paper (ref)	Year	Age, years	Test method	Reference method	Correlations biomarker *v.* dietary intake	Correlation type	r
Paper II	2016	8-9 and 12-14	WebFR	DBS - plasma carotenoids (NF)	Single carotenoids v. single carotenoid-rich foods	Spearman's	0.20-0.44
					Total carotenoids v. total carotenoid-rich foods	Spearman's	0.31
					β-carotene v. total vegetables	Spearman's	0.23
		8-9			Total carotenoids v. total vegetables	Spearman's	0.47
		12-14			Total carotenoids v. total vegetables	Spearman's	0.14
Nguyen et al. [187]	2015	9-12	FFQ	Plasma carotenoids (F)	Total carotenoids v. total carotenoids	Pearson's crude	0.39
					Total carotenoids v. total vegetable intake (not	Pearson's crude	0.26
Biltoft-Jensen et al. [182]	2013	8-11	WebDASC	Plasma carotenoids (F)	Total carotenoids v. total fruits, juice and vegetables	Spearman's	0.58
					Total carotenoids v. total carotenoids	Spearman's	0.43
					β-carotene v. total vegetables	Spearman's	0.38
					Total carotenoids v. total vegetables	Spearman's	0.33
Slater et al. [186]	2010	13	24HR x 2	Serum β-carotene (F)	β-carotene v. total vegetables	Pearson's adjusted	0.07[a]
			FFQ	Serum β-carotene (F)	β-carotene v. total vegetables	Pearson's adjusted	0.15[a]
Burrows et al. [149]	2009	5-12	FFQ parental report	Plasma carotenoids (F)	Single carotenoids v. single carotenoids	Pearson's crude/adjusted	-0.09-0.25/ 0.16-0.56[b]
					β-carotene v. β-carotene	Pearson's crude/adjusted	0.17/ 0.56[b]
Neu-houser et al. [188]	2001	12-17	FFQ	Plasma carotenoids (NF)	Single carotenoids v. single carotenoids	Pearson's adjusted	0.08-0.38[c]
					β-carotene v. β-carotene	Pearson's adjusted	0.15[c]
Byers et al. [185]	1993	6-10	FFQ parental report	Serum carotenoids (NF)	Total carotenoids[d] v. 35 fruits and vegetables	Spearman's	0.30

v., versus; NF, non-fasting sample; F, fasting sample.
[a] Adjusted for BMI, total fat, total cholesterol and fibre
[b] Adjusted for BMI
[c] Adjusted for age, sex, race, BMI, total serum cholesterol
[d] Including β-Carotene, α-Carotene and cryptoxanthin

In contrast to this, observations of school lunch entries (Paper I), discussed in section 4.2.2, showed that the food categories 'fruit, berries' and 'vegetables, salads' were among the categories in which omission rates were especially high. Specifically, a total of 42 (36%) and 33 (28%) children ate foods within the categories 'fruit, berries' and 'vegetables, salads' during their school lunch, respectively. These children had a mean omission rate of 39% and 45%. This is contrary to what one would expect based on the results in Paper II, in which the WebFR's ranking abilities for such foods were found satisfactory. A possible explanation is that the intake of 'fruit, berries' and 'vegetables, salads' during school lunch may only have contributed to a small fraction of the total intake of these food groups. Only about 1/3 of the children ate something in these categories during school lunch, and the omissions were predominantly of small portion sizes. This is further supported by data from a dietary survey in Norwegian adults, which demonstrated that the lunch meal only contributed with 13% of the intake of fruits, and 16% of vegetables [189]. To sum up, the WebFR's ranking abilities for fruits and vegetables were satisfactory for the mean intake, across all meals, despite the fact that a noteworthy proportion of small portions of these foods were omitted from recordings of school lunch.

4.2.4 Person-specific bias in children and adolescents (Paper I-III)

To identify possible person-specific bias, assessments of misreporting linked to participants' characteristics, were conducted in Paper I-III.

Overweight and obesity

In a recent review by Sharman et al., studies using either observations, the DLW method or the double portion technique, in 6-12-year-olds, were included: Factors associated with the accuracy of dietary recalls conducted without parental assistance were investigated [190]. Results from this review showed specifically that higher omission rates, but lower intrusion rates were associated with overweight [190]. Previous studies have demonstrated that under-reporting of EI is associated with a higher BMI in children and adolescents. Fisher et al. reported that 4-11-year-olds overestimated their EI by 14% in parental assisted 24HRs, as compared to the DLW method; however, children with a higher relative weight to height, were more likely to underestimate their EI, as compared to those with lower relative weight [191]. Lioret et al. identified 26% of 11-17-year-old children as under-reporters of energy, and 60% among those who were overweight, when using the Goldberg cut-off method for 7-days food records [192]. Moreover, Murakami et al. identified under-reporters of energy from a diet history questionnaire in 6-15-year-olds, using the Goldberg cut-off approach [193]; 32% were under-reporters, and the odds of being an under-reporter were three and six times higher for overweight and obese individuals, respectively. Results from Paper I-II in this thesis show no observed association between overweight/obesity and misreporting, but, this could be due to lack of power. For example, in Paper II, a correlation between the concentration of total carotenoids in plasma and recorded intake of vegetables was 0.30 for normal weight children, and 0.13 for overweight or obese children. The non-significant difference between these correlation coefficients is probably due to the limited number of overweight and obese children in the sample. However, in Paper III, a statistically significant increased underestimation of 2.4 MJ/day was observed in overweight and obese participants, as compared to normal weight children (reference), after adjusting for sex, age, parental educational level and family structure in a multiple regression analysis. This significant increase in underreporting in overweight and obese individuals corroborates the previously mentioned studies.

Age

The association between age and reporting accuracy was studied in Paper II and III, but not in Paper I, in which the study sample consisted of the 8-9-year-olds only. In Paper II, significantly higher correlations between carotenoids in plasma and recorded intakes of vegetables were observed for the 8-9-year-olds, as compared to the 12-14-year-olds, showing $r=0.47$ v. $r= 0.14$, respectively. Moreover, in Paper III, there was a significant increase in underestimation of energy for the 12-14-year-olds (0.69 MJ/day), as compared to the 8-9-year-olds (reference), adjusted for sex, BMI-category, parental educational level and family structure. This means that the 8-9-year-olds' dietary reports were the most accurate. In the review of Sharman et al., described in the previous section, increasing age was related to improved reporting accuracy of recalls in children from 6 to 12 years, not assisted by their parents, in 10 out of 13 studies [190]. This can be explained by the fact that older children are much more likely to have fully developed cognitive abilities, and have more extensive knowledge regarding their food intake as compared to younger children. Consequently, one could expect that the oldest participants (12-14 years) included in Paper II and III, would outperform the younger children (8-9 years). Nevertheless, the opposite was observed. There may be several factors that can explain this finding. Firstly, only the parents or guardians of 8-9-year-olds were instructed to assist their children during recording. Lack of assistance from parents or caregivers was associated with misreporting in 6-15-year-olds in Japan who completed a diet history questionnaire [193].

Although this study had a highly different study population, culture-wise, as compared to the validation studies of the WebFR, this may indicate that parental support may be valuable also in older children and adolescents, although perhaps not feasible. Secondly, social desirability emerges in older children or adolescents and can affect the recordings negatively [194]. Thirdly, the irregular eating structure characterising adolescence may furthermore impair the recording accuracy [195].

Parental educational level and ethnicity

Misreporting was associated with a low parental educational level in Paper I and Paper III. This is also reported elsewhere [192, 193, 196]. In Paper I, a lower match rate, representing poorer reporting accuracy, was also associated with having both parents/guardians of non-Norwegian ethnicity. However, a limited number of participants in the study had this characteristic. The review by Sharman et al. did not find associations between reporting accuracy in recalls (not parental assisted) and race or ethnicity, in 10 out of the 11 studies included in the review [190]. In the only study that showed an association between ethnic groups and accuracy of the recalls, the study sample consisted of recent immigrant and refugee children [190]. Comparing the findings of Sharman et al.'s review to the findings in this thesis may be problematic. This is because the studies in the review are primarily from the US, which means that these study samples may differ substantially from the study sample in this thesis, regarding ethnic origin, culture and other person-specific characteristics. In Paper II and Paper III, no association between ethnicity and recording accuracy was observed. This result is consistent with the previous-mentioned review of Sharman et al. It is noteworthy that an association between parental ethnicity and misreporting was observed in Paper I, in which the study sample consisted of 8-9-year-old children only, who depend on parental assistance when recording. Using the WebFR requires an understanding of the Norwegian language. These findings might reflect that non-Norwegian parents/guardians may have been unable to assist their children to a satisfactory degree, due to language barriers. Besides, a limited number of ethnic foods are included in the food list of the WebFR, which may have caused further challenges, although any food could be entered in the WebFR using an open field function.

Sex

No association between sex and recording accuracy was found for the WebFR, except for in Paper III, in which under-reporting of EI was assessed using accelerometer counts. A larger underestimation of EI was observed for boys as compared to girls. There are no obvious explanations for this observation, but it is consistent with observations in an earlier validation study of a paper-based pre-coded food record in Norwegian 9-year-olds [197]. However, other findings in the literature regarding the association between sex and misreporting, in children and adolescents, are inconsistent. Murakami et al. reported an association between female sex and under-reporting of EI in 6-15-year-olds [193], whereas Lioret et al. [192] and Sharman et al. [190] found no association between under-reporting and sex in 3-17-year-olds and 6-12-year-olds, respectively.

4.2.5 Social desirability bias in adults (Paper IV)

Dietary assessment may be affected by social desirability bias, which reflects individuals' desire to conform to social norms [198]. Specifically, this could manifest itself as a tendency to alter the reported intake in the direction that meets the dietary recommendations put forward by health authorities. An example of this is a recent study of Di Noia et al. which showed that social desirability was significantly associated with vegetable intake, assessed by a questionnaire in a group of American women with a mean age of 29 years [199]. However, it is not possible to conclude from that study whether women with high social desirability overestimates their vegetable intake, or if they simply eat more vegetables.

The validity of the WebFFQ relative to repeated 24HRs showed that the estimated intakes on group level and ranking abilities for macronutrients and most food groups were acceptable. The exceptions were the intakes of 'vegetables' and 'fish and shellfish'. Although the ranking abilities were acceptable for the energy-adjusted intake of these foods, both the absolute intakes and the estimated intakes for energy-adjusted intakes of these foods were overestimated significantly and to a large degree. Specifically, the WebFFQ estimate of vegetable intake and fish and shellfish intakes was 205% and 169% of the 24HRs, respectively. Dietary guidelines promoted by the health authorities in Norway recommend increasing the intake of these foods [200]. Additionally, FFQs and short food-lists assessing habitual intakes may be more prone to social desirability bias, because the diet is assessed as a characteristic of the individual, as opposed to the 24HRs that assess the acute intake [201]. One may, therefore, speculate whether this observed overestimation of 'vegetables' and 'fish and shellfish' by the WebFFQ may be due to social desirability. In support of this, 'cakes' were also underestimated significantly by the WebFFQ relative to the 24HRs, but 'alcohol' and 'sweets, desserts, sugars' were not, despite that all these items could be expected to be under-reported by individuals prone to social desirability.

4.3 Web-based versus traditional tools

4.3.1 Web-based FFQs versus paper-based FFQs

Several advantages promote the use of web-based FFQs over the paper-based ones. They provide complete data, through error checks, and may improve portion size estimations through the use of images. Using multiple images displayed simultaneously as an aid of portion size estimation has shown to increase reporting accuracy in a self-administered online 24HRs [202], which may suggest that this is also true for web-based FFQs. The web-based FFQs reduce the burden of data handling, as there is no need for manual checks, or to transfer data from paper to electronic formats. Additionally, they may also improve user-friendliness and compliance, due to easy access, the possibility of flexible completion, and the use of reminder messages and skip-algorithms. Lastly, as for the WebFFQ validated in this thesis, the possibility of electronic consent may increase user-friendliness further.

User-friendliness

Increased user-friendliness, if resulting in increased participation rates, may reduce selection bias, through obtaining a more representative selection of the target population. However, there are inconsistent findings in the literature regarding whether the web-format of FFQs actually improves the user-friendliness. For instance, the web-format of an FFQ was preferred over the paper version among 59%, in a relatively large Canadian study among adults, yet response rates were higher for the paper-based version than for the web-version [53]. In another study, response rates of web-based and paper-based surveys (not specifically FFQs) were compared [203]. The highest response rates were obtained when administrating both the web- and paper-based versions at the same time, letting the participants decide on which one to use. The second best rates were obtained using the web-based version, whereas the lowest response rates were obtained when participants were given the paper-based version only. These findings indicate that the web-format is preferred over paper-based methods for some, but not all participants. Depending on the sample, using a web-based FFQ may be perceived as a technical challenge that may perhaps be a hindrance, and partly explain these findings. Also, a paper-based FFQ is portable and does not require any Internet access or computer literacy, which may be an advantage for some.

Performance

The built-in error checks of the web-based FFQs can reduce random error. It could, therefore, be expected that web-based FFQs, would outperform the paper-based versions. The WebFFQ validated in Paper IV builds upon and resembles a paper-based version, also validated using DLW [134]. Results showed that the WebFFQ tool is neither superior nor worse in estimating EI than the paper-based FFQ: EI was under-reported on group level by 0.70 MJ/day and 0.96 MJ/day by the WebFFQ and paper-based FFQ, respectively, and both tools showed poor ranking abilities for EI. Nevertheless, the study populations were dissimilar; women in the validation study of the paper-based FFQ were young university students with a mean age of 24 years, and a mean BMI of 22 kg/m^2, whereas the women in Paper IV were much more diverse in age and BMI. Moreover, 180 items were included in the paper-based FFQ, as compared to 279 items, including many images for portion size estimations, in the WebFFQ. Another Norwegian paper-based FFQ holding 270 items, also based on the previous described paper-based FFQ, overestimated the intake of vegetables by 51% relative to a 7-days weighed food record in a study sample of adult men and women [133]. In comparison, in Paper IV we observed that the WebFFQ overestimated the intake of vegetables significantly and to a large extent (121%), relative to the 24HRs, across all participants for absolute intakes. In view of this, the direction and the magnitude of the misreporting is, to some extent, comparable between the WebFFQ and the two paper-based FFQs. Nevertheless, due to the differences between study populations, one cannot conclude that the WebFFQ is superior to these two paper-based FFQs, or vice versa. One may speculate whether the lack of convincing and clear improvements in the accuracy of the web-based FFQ format is due to the fact that improvements of the web-based FFQs, e.g. through increased completeness, are insignificant compared to overall measurement errors still associated with FFQs. Web-FFQs are still cognitive complex tools, with the same underlying structure as the paper-based ones.

A few studies have assessed the inter-version reliability of similar paper-based and web-based FFQs, by administering them both to the same study population. Forster et al. compared the Food4Me, a web-based FFQ with 157 items, to a similar paper-based version of that FFQ in a study of 113 adults, with a mean age of 30 years [204]. They showed crude, unadjusted correlations between methods in the range 0.41-0.90, including nutrients and foods groups, and a 2.8 MJ/day higher EI for the web-based Food4Me as compared to the paper-based version. In a study of 31 middle aged women, adequate reliability was observed for the web-based FFQ as compared to the paper-based version. [55]. That is, correlation coefficients between methods for energy, macronutrients and food groups were between 0.68-0.86, and the estimates of EI were not statistically different between the web- and paper-based version. This study of the middle-aged women is in line with a recent Canadian study of 347 adults comparing a web-based FFQ with a comparable paper-based version, in which the estimates of nutrients were for the most part similar in the different versions [53]. Although, the web-based version of the FFQ in this Canadian study showed a small, significant higher estimate of EI for women only (0.2 MJ/day), as compared to the paper-based version. Then again, in a pilot study using a crossover design in a group of Spanish university students, significantly lower intakes, e.g., 1.9 MJ/day for energy, were observed for a web-based FFQ compared to a paper-based version of the same FFQ [205]. These findings indicate that the web-based format, for the most part, seems to be similar to the paper-based versions of the same FFQs; thus the mode of administration appears to be of little importance. However, the studies mentioned above are few, and several have a small sample size. Thus it is not possible to draw definite conclusions. Nevertheless, if we were to conclude that results are similar regardless of the administration mode, the reduced burden of data handling associated with the web-based FFQs, is alone enough to strongly advocate the use of this format, over the paper-based one.

4.3.2 Web-based versus other dietary assessment tools for young individuals

Given the perspective of researchers, there is undoubtedly a preference for self-administered web-based 24HRs and web-based food records, like the WebFR. This is above all due to the significant reduction in manual data handling compared to the traditional paper-based food records or interviewer administered 24HRs. Nevertheless, web- and image-based portable food records, typically developed for smartphones, do not necessarily reduce the burden of data handling for researchers. The Remote Foods Photography Method (RFPM) is an example of this, used in both adults and children [61]. The RFPM currently requires a human operator to manually look through all captured images, because the food and portion size identification process is still not fully automatic. However, progress has been made in developing technology that automatically identifies all foods and portion sizes from images. This can be exemplified by the TADA [60], a web- and image-based food record for smartphones, in which all images are automatically processed, which reduces the burden for both the researchers and the users. The

TADA is believed to be specifically suitable for adolescents, as they are used to adapt to new technology [60].

There are two major aspects to consider when concluding whether the web-based methods for children and adolescents are superior to the traditional ones. The first comprises the question regarding if the web-based methods are more user-friendly or not. If they are, we may obtain more representative samples, and reduce selection bias, as argued in section 4.2.1, for the WebFFQ. The second important aspect is whether the magnitude of errors is reduced for the web-based methods, as compared to the traditional paper- or interview based techniques.

User-friendliness

Technology-based dietary assessment tools are assumed to be especially accepted and preferred over traditional methods, among children and adolescents, as they are familiar with the use of technology in many aspects of their life. A few studies have evaluated the mode of preference for young individuals. Dutch children (10-12 years) preferred a web-based questionnaire over an identical paper-based questionnaire [206], and Dutch adolescents (13-17 years) also favoured a web-based health questionnaire over a paper-based one [207]. In the study validating the WebFR (Paper I-III), the preference of the web-mode versus a paper-based mode was not assessed. However, we had few dropouts and a relatively high participation rate, which may indicate that the WebFR was well accepted among 8-9-year-olds and 12-14-year-olds. Moreover, the participation rates, in the national dietary survey in Norway (UNGKOST 3) from 2015 [130], in which the WebFR was used in the same age groups as in the current thesis, were also acceptable. Participation rates were 55% and 53%, for 8-9-year-olds and 12-14-year-olds, respectively. This further indicates that the WebFR is user-friendly. Boushey and co-workers assessed 11-15-year-olds' preferences for six different modes of dietary assessment and found that technology-based tools were preferred over paper-based dietary records [208]. Specifically, capturing dietary intake using either a disposable camera or PDA with a camera, were the most popular methods, followed by a PDA with search functions. This shows that there might be an even stronger preference for image based real-time assessment, than web-based tools like the WebFR, at least in older children and adolescents. One may speculate if this is due to the fact that capturing images requires minimal work for the participant, or because adolescents may already be used to capturing images of what they eat with their smartphones, and even sharing them on social media platforms like Instagram. In conclusion, it seems as if the web-format is well accepted and preferred by the Internet generation, over traditional methods. Using web-based dietary assessment tools, like the WebFR or image-based dietary assessment, may therefore possibly reduce selection bias as compared to the use of traditional methods among young individuals.

Performance

The heterogeneity in technology-based assessment tools available for children and adolescents, and in the evaluation studies published, makes it challenging to confidently conclude whether or not the web-based methods are superior to the traditional methods. Specifically, for children, there is a lack of high-quality validation studies using objective reference methods and with a large sample size.

As discussed previously, the WebFR was not able to capture EI with accuracy. EI was underestimated by 1.8 MJ/day on group level across all individuals in both age groups, 36-37% of all participants were identified as under-reporters and 2-4% as over-reporters (Paper III). These findings are almost in complete agreement with results from a paper-based food record validated in 9-year-olds in Norway by Lillegaard et al. [197]. In that study, the mean EI was underestimated by 1.8 MJ/day across all individuals, and, 39% and 4% were classified as under- and over-reporters, respectively. The youngest participants in Paper III – the 8-9-years-olds – underestimated their EI by 1.4 MJ/day on group level, in comparison. These observations indicate that the WebFR is in line with, or perhaps slightly better than this traditional comparable paper-based tool, in respect of estimating the EI.

Results from the validation studies of the WebFR and results from the validation studies of the Danish WebDASC [182, 209], which the WebFR is based on, have been compared and discussed extensively in Paper I-III. The results show that the WebDASC appears to be superior to the WebFR in all aspects. However, as the tools are relatively similar, it is more likely that the different results obtained are caused by other factors. The lack of parental assistance in the oldest age group in Paper II and III could perhaps be such a contributing factor. However, the WebDASC also performed superior to the WebFR in Paper I in which just the youngest participants, who received assistance from their parents/guardians, were included. Other explanatory factors may include reactivity, which might have been an issue in the WebDASC study, or dissimilarities in the age of the study participants. The fact that the study population in the validation study of the WebFR was fairly diverse in respect of ethnicity, and had a higher proportion of overweight and obese individuals, as compared to the validation studies of the WebDASC may also explain the dissimilar results between the validation studies.

In respect of improving the WebFR, it would be of great interest to know whether the web- and image-based portable food records or tools, in which participants may capture images of eating events as a memory aid, perform better than the WebFR. Svensson et al. showed that overweight and obese 8-12-year-olds underestimated their dietary intake by 2.8 MJ/day (24%) on group level, using a traditional food record combined with a digital camera to assess diet, and an accelerometer-based reference method assessing TEE [210]. In comparison, EI was underestimated by 4.1 MJ/day (43%) across all the overweight or obese individuals, including both the 8-9-year-olds and the 12-14-year-olds, in Paper III. Another Swedish study, also by Svensson et al., using objective reference estimates of TEE, assessed 14-16-year-olds' ability to report their diet using a mobile phone app [67]. The EI was underestimated by 2.8 MJ/day (29%) using median values on group level, as compared to a mean underestimation of 2.3 MJ/day (25%) in our sample of 12-14-year-olds. In the latter study of Svensson et al., a small subsample (n=15) used a web-based, non-portable tool in addition to the app; no significant differences between the two methods were observed [67]. Despite the fact that these studies to a certain degree may be comparable in respect of reference methods used, it is difficult to isolate whether dissimilarities in study populations have had an impact on the results. It is hard to conclude based on these studies, but one can speculate if the superior accuracy in the first out of these two studies by Svensson et al. is partly due to the use of the camera.

4.3.3 Cost-effectiveness of web-based dietary assessment

Repeatedly it has been stated that shifting to technology-based dietary assessment methods, including web-based methods are cost-effective [30, 42, 211]. The cost-effectiveness of a web-based 24-hour dietary record, as compared to a dietitian-conducted 24HR, was found to be in favour of the web-based-tool, especially because the cost of interviewers diminished [212]. Nevertheless, there is a lack of studies that have compared the cost-effectiveness of paper-based FFQs and records to web-based ones. Intuitively, one may argue that despite the fact that the methods may be expensive to develop, the web-based methods reduce the cost as there is no need for postal services in surveys or large studies, and because of lower cost due to reduced need for manual data handling. However, the cost of keeping software up-to-date or other needs of maintenance may be significant and call for specialised computer expertise. Moreover, to obtain acceptably high participation rates, face-to-face recruitment may be necessary [213]; thus travel expenses is not automatically avoided. When taking all these aspects into consideration, it remains unclear at this point whether the total cost is reduced.

5 Conclusion

The validity of the WebFR is assessed in this thesis in a sample of children and adolescents. The reference methods used were direct observations, the concentration biomarkers carotenoids and accelerometer-derived TEE as a marker of EI.

- ✓ Observations of 8-9-year-olds' school lunches showed that both omissions and intrusions were common, with a mean omission rate of 29% and a mean intrusion rate of 19%, across all food groups and all children.

- ✓ Spreads, fruits and vegetables were among the specific foods that were an important source of misreporting in the school lunch entries, but the omitted proportions of fruits and vegetables were mostly of small portion sizes. Nevertheless, the WebFR has acceptable ranking abilities for carotenoid-rich foods, demonstrated by using concentrations of carotenoids in plasma as a biomarker of intake.

- ✓ The estimated EI from the WebFR should be used with caution, given the mean underreporting of EI by 1.83 MJ/day and a correlation of 0.16 across the entire sample, which was demonstrated using accelerometer counts to derive TEE as a marker of EI.

- ✓ Overweight or obesity, older age, and a low parental educational level were the most important factors associated with misreporting in the WebFR. Parental ethnicity seems to be important for young children who need parental assistance during recordings. Hence, younger children with non-Norwegian parents would possibly benefit from receiving extra assistance during recordings.

- ✓ As other dietary assessment tools based on self-reported data, the WebFR suffers from misreporting, but are in line with other comparable web-based tools in these age groups.

- ✓ The WebFR would potentially benefit from being adapted for use on portable platforms, like smartphones or tablets, enabling real-time recordings that perhaps can reduce omissions and intrusions, and improve the user-friendliness. Alternatively, users in future studies could probably benefit from taking images with their camera phone during the day, as a memory aid to improve the recordings.

The validity of the WebFFQ is assessed in this thesis in an adult study sample. The reference methods used were the DLW method and repeated 24HRs.

- ✓ The group mean EI from the WebFFQ was not significantly different (-6%) from TEE measured by the DLW method in a sample of adult women.
- ✓ The WebFFQ's ability to range women according to their EI was poor, given the correlation coefficient of -0.18 between DLW derived TEE and EI from the WebFFQ.
- ✓ Consequently, the WebFFQ seems able to estimate EI on group level, but not on an individual level. Estimated absolute individual EI from the WebFFQ should, therefore, not be used. Estimates of energy may still be valuable for energy adjustments.
- ✓ The relative comparison between the WebFFQ and 24HRs, in a sample of men and women, showed that the estimated absolute intakes of macronutrients and most food groups from the WebFFQ were acceptable on group level, except for 'vegetables' and 'fish and shellfish', which were significantly and largely overestimated by the WebFFQ.
- ✓ The WebFFQ was able to rank individuals correctly according to their reported intake of macronutrients and most food groups, especially when using the energy-adjusted intakes.
- ✓ The WebFFQ bears similarities to paper-based FFQs, and it appears to be neither better nor worse than other comparable paper-based FFQs.
- ✓ The WebFFQ would be a suitable tool in future dietary surveys and epidemiological studies and will reduce the burden on researchers. Nevertheless, it is important to bear in mind the limitations of FFQ-data.

6 Final remarks and future perspectives

Over 20 years ago, Beaton stated that: *"Dietary intake cannot now be estimated without error; it never will be!"* [214]. Since then, substantial development in dietary assessment methodologies, discoveries of new biomarkers and development of new statistical approaches have materialised; and Beaton's statement still holds. It seems inevitable that self-reports will always be troubled with misreporting, due to the nature of humans. Memory, skills and knowledge, and social desirability are among the factors compromising self-reports. The results from the validation studies of the WebFR and WebFFQ, presented in this thesis, corroborate this.

The reason why we use and even develop new imperfect methods based on self-reports is that despite all the errors, they provide insights and rich data on the complex behaviour of eating that has not been feasible to obtain in any other way. Observational studies in nutritional epidemiology using data from FFQs and 24HRs have provided valuable insights on dietary intake and health [215]. Examples include the association between folate intake and neural tube defects, or the effect of trans-fatty acids on LDL cholesterol [216]. On the other hand, we are probably also failing to show several existing associations, due to the errors in dietary assessment.

As discussed previously, especially the FFQ has been criticised substantially, and in particular, the energy estimates derived from it. Some even argue that the FFQ should be abandoned in future studies, and replaced by multiple short-term instruments [9]. Doing so may prove feasible, due to the use of web-based 24HRs and records which require minimal data handling. However, for episodically consumed foods, the FFQ has some distinct advantages over the short-term instruments. Moreover, no other method can replace the FFQ or web-based FFQ in large case-control studies in the future, or in other situations in which information about dietary exposure in the past is needed.

Combining tools is also a promising approach. One may get information on the probability of consumption of specific foods from an FFQ and details regarding portion sizes from 24HRs or food records. A group in Germany has used such approach: Multiple short 24HR-food lists and FFQ data have been combined with data on portion sizes from food records, using novel statistical methods [25]. Hence, combining data from the WebFFQ and multiple days of dietary records or recalls may be a way forward. This does require extensive collaborations with statisticians, which will probably become even more important in future nutritional studies. Adapting the WebFR to adults, or developing a new self-administered web-based food record/recall tool for adults, would probably also prove beneficial to reduce data handling, as stand-alone tools, or in combination with the WebFFQ, as described above.

Technology is rapidly altering society, and will probably change dietary assessment far beyond what is seen today. At the moment, the ambulatory assessment, which comprises real-time assessment using portable tools, like smartphones, seems promising because of its memory independent nature. The WebFR validated in this thesis, can perhaps benefit from being adapted to smartphone technology. However, we need additional validation studies to investigate this further.

Nevertheless, to fully avoid errors associated with self-reports, it seems as if a shift to objective methods is the only solution. Automatic recording in real time is one alternative, but getting there will be challenging. The eButton [76], or its like, in which video and audio are captured, may be compromising the protection of personal information. Sensor technology and big data, derived from data on e.g. grocery shopping, will probably also prove helpful. Biomarkers of exposure is another promising objective alternative. Both new biomarkers derived from metabolomics, giving info on short-term intakes, and the promising stable isotopes, providing info on the long-term intake, will most likely supplement self-reports or other methods to assess dietary intake. However, biomarkers will probably not, at least in the immediate future, be a real alternative to assess the entire dietary intake, and will never be able to provide contextual data for the eating events.

To summarise, there are several promising and interesting emerging approaches to dietary assessment that are likely to improve our data. Still, in the immediate future, tools based on self-report, like the WebFFQ and the WebFR, will have a dominant position in dietary assessment.

7 Reference list

1. Forouzanfar MH, Alexander L, Anderson HR *et al.* Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* **386**, 2287-2323.
2. World Health Organization (2017) Epidemiology. http://www.who.int/topics/epidemiology/en/ (accessed April 2017).
3. Langseth L (1996) Nutritional Epidemiology: Possibilities and Limitations. In *ILSI Europe Concise Monograph*. [Kok FJ and Jardine NJ, editors]. Brussels, Belgium: The International Life Sciences Institute (ILSI).
4. Carpenter KJ (2003) A short history of nutritional science: part 1 (1785-1885). *J Nutr* **133**, 638-645.
5. Carpenter KJ (2003) A short history of nutritional science: part 3 (1912-1944). *J Nutr* **133**, 3023-3032.
6. Medlin C, Skinner JD (1988) Individual dietary intake methodology: a 50-year review of progress. *J Am Diet Assoc* **88**, 1250-1257.
7. Widdowson EM (1936) A Study of English Diets by the Individual Method: Part I. Men. *J Hyg (Lond)* **36**, 269-i264.
8. Burke BS (1947) The dietary history as a tool in research. *J Am Diet Assoc* **23**, 1041-1046.
9. Slimani NF, H. Illner, A-K. Huybrechts, I. (2015) Methods to Determine Dietary Intake. In *Nutrition Research Methodologies* [Lovegrove JAH, Sharma L, Lanham-New SA, editors]. West Sussex, UK: John Wiley & Sons, Ltd.
10. Stefanik PA, Trulson MF (1962) Determining the frequency intakes of foods in large group studies. *Am J Clin Nutr* **11**, 335-343.
11. Willett WC, Sampson L, Stampfer MJ *et al.* (1985) Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* **122**, 51-65.
12. Thompson FE, Subar AF (2013) Dietary assessment methodology. In *Nutrition in the Prevention and Treatment of Disease,* 3rd ed. [Coulston AM, Boushey CJ, editors].San Diego: Academic Press.
13. Burke BS, Stuart HC (1938) A method of diet analysis. *The Journal of Pediatrics* **12**, 493-503.
14. Isaksson B (1980) Urinary nitrogen output as a validity test in dietary surveys. *Am J Clin Nutr* **33**, 4-5.
15. Schoeller DA, van Santen E (1982) Measurement of energy expenditure in humans by doubly labeled water method. *J Appl Physiol Respir Environ Exerc Physiol* **53**, 955-959.
16. Plakke T, Berkel J, Beynen AC *et al.* (1983) Relationship between the fatty acid composition of the diet and that of the subcutaneous adipose tissue in individual human subjects. *Hum Nutr Appl Nutr* **37**, 365-372.
17. Ziegler RG, Subar AF, Craft NE *et al.* (1992) Does beta-carotene explain why reduced cancer risk is associated with vegetable and fruit intake? *Cancer Res* **52**, 2060s-2066s.
18. Schwab K (2016) The Fourth Industrial Revolution: what it means, how to respond. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/ (accessed February 2017).
19. Illner AK, Freisling H, Boeing H *et al.* (2012) Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. *Int J Epidemiol* **41**, 1187-1203.
20. Boushey CJ, Spoden M, Zhu FM *et al.* (2016) New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. *Proc Nutr Soc*, 1-12.
21. Gemming L, Utter J, Ni Mhurchu C (2015) Image-assisted dietary assessment: a systematic review of the evidence. *J Acad Nutr Diet* **115**, 64-77.
22. Storey KE (2015) A changing landscape: web-based methods for dietary assessment in adolescents. *Curr Opin Clin Nutr Metab Care* **18**, 437-445.
23. Forster H, Walsh MC, Gibney MJ *et al.* (2016) Personalised nutrition: the role of new dietary assessment methods. *Proc Nutr Soc* **75**, 96-105.
24. Timon CM, van den Barg R, Blain RJ *et al.* (2016) A review of the design and validation of web- and computer-based 24-h dietary recall tools. *Nutr Res Rev* **29**, 268-280.
25. Conrad J, Nothlings U (2017) Innovative approaches to estimate individual usual dietary intake in large-scale epidemiological studies. *Proc Nutr Soc*, 1-7.

26. Kristal AR, Kolar AS, Fisher JL *et al.* (2014) Evaluation of web-based, self-administered, graphical food frequency questionnaire. *J Acad Nutr Diet* **114**, 613-621.
27. Raper N, Perloff B, Ingwersen L *et al.* (2004) An overview of USDA's Dietary Intake Data System. *Journal of Food Composition and Analysis* **17**, 545-555.
28. Subar AF, Kirkpatrick SI, Mittl B *et al.* (2012) The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. *J Acad Nutr Diet* **112**, 1134-1137.
29. Subar AF, Thompson FE, Potischman N *et al.* (2007) Formative research of a quick list for an automated self-administered 24-hour dietary recall. *J Am Diet Assoc* **107**, 1002-1007.
30. Foster E, Hawkins A, Delve J *et al.* (2014) Reducing the cost of dietary assessment: self-completed recall and analysis of nutrition for use with children (SCRAN24). *J Hum Nutr Diet* **27 Suppl 1**, 26-35.
31. Slimani N, Casagrande C, Nicolas G *et al.* (2011) The standardized computerized 24-h dietary recall method EPIC-Soft adapted for pan-European dietary monitoring. *Eur J Clin Nutr* **65 Suppl 1**, S5-15.
32. Baranowski T, Islam N, Baranowski J *et al.* (2002) The food intake recording software system is valid among fourth-grade children. *J Am Diet Assoc* **102**, 380-385.
33. Vereecken CA, Covents M, Matthys C *et al.* (2005) Young adolescents' nutrition assessment on computer (YANA-C). *Eur J Clin Nutr* **59**, 658-667.
34. Zoellner J, Anderson J, Gould SM (2005) Comparative validation of a bilingual interactive multimedia dietary assessment tool. *J Am Diet Assoc* **105**, 1206-1214.
35. Smith BA, Morgan SL, Vaughn WH *et al.* (1999) Comparison of a computer-based food frequency questionnaire for calcium intake with 2 other assessment tools. *J Am Diet Assoc* **99**, 1579-1581.
36. Wong SS, Boushey CJ, Novotny R *et al.* (2008) Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, Hispanic, and non-Hispanic white youth. *J Am Diet Assoc* **108**, 539-543.
37. Slattery ML, Murtaugh MA, Schumacher MC *et al.* (2008) Development, implementation, and evaluation of a computerized self-administered diet history questionnaire for use in studies of American Indian and Alaskan native people. *J Am Diet Assoc* **108**, 101-109.
38. Edwards SL, Slattery ML, Murtaugh MA *et al.* (2007) Development and Use of Touch-Screen Audio Computer-assisted Self-Interviewing in a Study of American Indians. *American Journal of Epidemiology* **165**, 1336-1342.
39. Matthys C, Pynaert I, Roe M *et al.* (2004) Validity and reproducibility of a computerised tool for assessing the iron, calcium and vitamin C intake of Belgian women. *Eur J Clin Nutr* **58**, 1297-1305.
40. Diep CS, Hingle M, Chen TA *et al.* (2015) The Automated Self-Administered 24-Hour Dietary Recall for Children, 2012 Version, for Youth Aged 9 to 11 Years: A Validation Study. *J Acad Nutr Diet*.
41. Vereecken C, Covents M, Maes L *et al.* (2014) Formative evaluation of the dietary assessment component of Children's and Adolescents' Nutrition Assessment and Advice on the Web (CANAA-W). *J Hum Nutr Diet* **27 Suppl 1**, 54-65.
42. Moore HJ, Hillier FC, Batterham AM *et al.* (2014) Technology-based dietary assessment: development of the Synchronised Nutrition and Activity Program (SNAP). *J Hum Nutr Diet* **27 Suppl 1**, 36-42.
43. Baranowski T, Islam N, Douglass D *et al.* (2014) Food Intake Recording Software System, version 4 (FIRSSt4): a self-completed 24-h dietary recall for children. *J Hum Nutr Diet* **27 Suppl 1**, 66-71.
44. Carvalho MA, Baranowski T, Foster E *et al.* (2014) Validation of the Portuguese self-administered computerised 24-hour dietary recall among second-, third- and fourth-grade children. *J Hum Nutr Diet*.
45. Bradley J, Simpson E, Poliakov I *et al.* (2016) Comparison of INTAKE24 (an Online 24-h Dietary Recall Tool) with Interviewer-Led 24-h Recall in 11-24 Year-Old. *Nutrients* **8**.
46. Carter MC, Albar SA, Morris MA *et al.* (2015) Development of a UK Online 24-h Dietary Assessment Tool: myfood24. *Nutrients* **7**, 4016-4032.
47. Wardenaar FC, Steennis J, Ceelen IJ *et al.* (2015) Validation of web-based, multiple 24-h recalls combined with nutritional supplement intake questionnaires against nitrogen excretions to determine protein intake in Dutch elite athletes. *Br J Nutr* **114**, 2083-2092.
48. Christensen SE, Moller E, Bonn SE *et al.* (2013) Two new meal- and web-based interactive food frequency questionnaires: validation of energy and macronutrient intake. *J Med Internet Res* **15**, e109.
49. Fallaize R, Forster H, Macready AL *et al.* (2014) Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. *J Med Internet Res* **16**, e190.

50. Trijsburg L, de Vries JH, Boshuizen HC *et al.* (2015) Comparison of duplicate portion and 24 h recall as reference methods for validating a FFQ using urinary markers as the estimate of true intake. *Br J Nutr* **114**, 1304-1312.
51. Franco RZ, Alawadhi B, Fallaize R *et al.* (2017) A Web-Based Graphical Food Frequency Assessment System: Design, Development and Usability Metrics. *JMIR Hum Factors* **4**, e13.
52. Verger EO, Armstrong P, Nielsen T *et al.* (2017) Dietary Assessment in the MetaCardis Study: Development and Relative Validity of an Online Food Frequency Questionnaire. *J Acad Nutr Diet* **117**, 878-888.
53. Lo Siou G, Csizmadi I, Boucher BA *et al.* (2017) The Comparative Reliability and Feasibility of the Past-Year Canadian Diet History Questionnaire II: Comparison of the Paper and Web Versions. *Nutrients* **9**.
54. Beasley JM, Davis A, Riley WT (2009) Evaluation of a web-based, pictorial diet history questionnaire. *Public Health Nutr* **12**, 651-659.
55. Boeckner LS, Pullen CH, Walker SN *et al.* (2002) Use and reliability of the World Wide Web version of the Block Health Habits and History Questionnaire with older rural women. *J Nutr Educ Behav* **34 Suppl 1**, S20-24.
56. Medin AC, Astrup H, Kasin BM *et al.* (2015) Evaluation of a Web-Based Food Record for Children Using Direct Unobtrusive Lunch Observations: A Validation Study. *J Med Internet Res* **17**, e273.
57. Biltoft-Jensen A, Trolle E, Christensen T *et al.* (2014) WebDASC: a web-based dietary assessment software for 8-11-year-old Danish children. *J Hum Nutr Diet* **27 Suppl 1**, 43-53.
58. Freese J, Feller S, Harttig U *et al.* (2014) Development and evaluation of a short 24-h food list as part of a blended dietary assessment strategy in large-scale cohort studies. *Eur J Clin Nutr* **68**, 324-329.
59. Timon CM, Blain RJ, McNulty B *et al.* (2017) The Development, Validation, and User Evaluation of Foodbook24: A Web-Based Dietary Assessment Tool Developed for the Irish Adult Population. *J Med Internet Res* **19**, e158.
60. Schap TE, Zhu F, Delp EJ *et al.* (2014) Merging dietary assessment with the adolescent lifestyle. *J Hum Nutr Diet* **27 Suppl 1**, 82-88.
61. Martin CK, Nicklas T, Gunturk B *et al.* (2014) Measuring food intake with digital photography. *J Hum Nutr Diet* **27 Suppl 1**, 72-81.
62. Casperson SL, Sieling J, Moon J *et al.* (2015) A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: usability study. *JMIR Mhealth Uhealth* **3**, e30.
63. Kikunaga S, Tin T, Ishibashi G *et al.* (2007) The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey. *J Nutr Sci Vitaminol (Tokyo)* **53**, 109-116.
64. Rollo ME, Ash S, Lyons-Wall P *et al.* (2011) Trial of a mobile phone method for recording dietary intake in adults with type 2 diabetes: evaluation and implications for future applications. *J Telemed Telecare* **17**, 318-323.
65. Fukuo W, Yoshiuchi K, Ohashi K *et al.* (2009) Development of a hand-held personal digital assistant-based food diary with food photographs for Japanese subjects. *J Am Diet Assoc* **109**, 1232-1236.
66. Carter MC, Burley VJ, Nykjaer C *et al.* (2013) 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. *Br J Nutr* **109**, 539-546.
67. Svensson A, Larsson C (2015) A Mobile Phone App for Dietary Intake Assessment in Adolescents: An Evaluation Study. *JMIR Mhealth Uhealth* **3**, e93.
68. Bucher Della Torre S, Carrard I, Farina E *et al.* (2017) Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. *Nutrients* **9**.
69. Lacson R, Long W (2006) Natural language processing of spoken diet records (SDRs). *AMIA Annu Symp Proc*, 454-458.
70. Pettitt C, Liu J, Kwasnicki RM *et al.* (2016) A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. *Br J Nutr* **115**, 160-167.
71. Eyles H, Jiang Y, Ni Mhurchu C (2010) Use of household supermarket sales data to estimate nutrient intakes: a comparison with repeat 24-hour dietary recalls. *J Am Diet Assoc* **110**, 106-110.
72. Welch JL, Siek KA, Connelly KH *et al.* (2010) Merging health literacy with computer technology: self-managing diet and fluid intake among adult hemodialysis patients. *Patient Educ Couns* **79**, 192-198.
73. Bi Y, Lv M, Song C *et al.* (2016) AutoDietary: A Wearable Acoustic Sensor System for Food Intake Recognition in Daily Life. *IEEE Sensors Journal* **16**, 806-816.

74. Fontana JM, Farooq M, Sazonov E (2014) Automatic ingestion monitor: a novel wearable device for monitoring of ingestive behavior. *IEEE Trans Biomed Eng* **61**, 1772-1779.
75. Hodges S, Williams L, Berry E *et al.* (2006) SenseCam: A Retrospective Memory Aid. In *UbiComp 2006: Ubiquitous Computing: 8th International Conference, UbiComp 2006 Orange County, CA, USA, September 17-21, 2006 Proceedings*, pp. 177-193 [P Dourish and A Friday, editors]. Berlin, Heidelberg: Springer Berlin Heidelberg.
76. Sun M, Fernstrom JD, Jia W *et al.* (2010) A wearable electronic system for objective dietary assessment. *J Am Diet Assoc* **110**, 45-47.
77. Gemming L, Doherty A, Kelly P *et al.* (2013) Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *Eur J Clin Nutr* **67**, 1095-1099.
78. O'Loughlin G, Cullen SJ, McGoldrick A *et al.* (2013) Using a wearable camera to increase the accuracy of dietary analysis. *Am J Prev Med* **44**, 297-301.
79. Trull TJ, Ebner-Priemer U (2013) Ambulatory assessment. *Annu Rev Clin Psychol* **9**, 151-176.
80. Arab L, Estrin D, Kim DH *et al.* (2011) Feasibility testing of an automated image-capture method to aid dietary recall. *Eur J Clin Nutr* **65**, 1156-1162.
81. Strimbu K, Tavel JA (2010) What are biomarkers? *Curr Opin HIV AIDS* **5**, 463-466.
82. Biomarkers Definitions Working G (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89-95.
83. Kuhnle GG (2012) Nutritional biomarkers for objective dietary assessment. *J Sci Food Agric* **92**, 1145-1149.
84. Arab L, Akbar J (2002) Biomarkers and the measurement of fatty acids. *Public Health Nutr* **5**, 865-871.
85. Corella D, Ordovas JM (2015) Biomarkers: background, classification and guidelines for applications in nutritional epidemiology. *Nutr Hosp* **31 Suppl 3**, 177-188.
86. Jenab M, Slimani N, Bictash M *et al.* (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum Genet* **125**, 507-525.
87. Hedrick VE, Dietrich AM, Estabrooks PA *et al.* (2012) Dietary biomarkers: advances, limitations and future directions. *Nutr J* **11**, 109.
88. Scalbert A, Brennan L, Manach C *et al.* (2014) The food metabolome: a window over dietary exposure. *Am J Clin Nutr* **99**, 1286-1308.
89. Bingham SA (2003) Urine nitrogen as a biomarker for the validation of dietary protein intake. *J Nutr* **133 Suppl 3**, 921S-924S.
90. Mickelsen O, Makdani D, Gill JL *et al.* (1977) Sodium and potassium intakes and excretions of normal men consuming sodium chloride or a 1:1 mixture of sodium and potassium chlorides. *Am J Clin Nutr* **30**, 2033-2040.
91. Tasevska N, Runswick SA, Bingham SA (2006) Urinary potassium is as reliable as urinary nitrogen for use as a recovery biomarker in dietary studies of free living individuals. *J Nutr* **136**, 1334-1340.
92. Speakman JR (1998) The history and theory of the doubly labeled water technique. *Am J Clin Nutr* **68**, 932S-938S.
93. International Atomic Energy Agency (2009) *Assessment of Body Composition and Total Energy Expenditure in Humans Using Stable Isotope Techniques*. IAEA Human Health Series No. 3. Vienna, Austria: IAEA.
94. Tasevska N, Runswick SA, McTaggart A *et al.* (2005) Urinary sucrose and fructose as biomarkers for sugar consumption. *Cancer Epidemiol Biomarkers Prev* **14**, 1287-1294.
95. Willett W (2013) *Nutritional Epidemiology*. 3rd ed. New York, United States of America: Oxford University Press.
96. Maiani G, Caston MJ, Catasta G *et al.* (2009) Carotenoids: actual knowledge on food sources, intakes, stability and bioavailability and their protective role in humans. *Mol Nutr Food Res* **53 Suppl 2**, S194-218.
97. Nagao A (2009) Absorption and function of dietary carotenoids. *Forum Nutr* **61**, 55-63.
98. Khachik F, Spangler CJ, Smith JC, Jr. *et al.* (1997) Identification, quantification, and relative concentrations of carotenoids and their metabolites in human milk and serum. *Anal Chem* **69**, 1873-1881.
99. Brevik A, Andersen LF, Karlsen A *et al.* (2004) Six carotenoids in plasma used to assess recommended intake of fruits and vegetables in a controlled feeding study. *Eur J Clin Nutr* **58**, 1166-1173.
100. Broekmans WM, Klopping-Ketelaars IA, Schuurman CR *et al.* (2000) Fruits and vegetables increase plasma carotenoids and vitamins and decrease homocysteine in humans. *J Nutr* **130**, 1578-1583.
101. Rock CL, Swendseid ME, Jacob RA *et al.* (1992) Plasma carotenoid levels in human subjects fed a low carotenoid diet. *J Nutr* **122**, 96-100.

102. United States Department of Agriculture (2014) USDA National Nutrient Database http://ndb.nal.usda.gov/ (accessed February-March 2014).
103. Wishart DS (2008) Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology* **19**, 482-493.
104. O'Gorman A, Gibbons H, Brennan L (2013) Metabolomics in the identification of biomarkers of dietary intake. *Computational and Structural Biotechnology Journal* **4**, e201301004.
105. O'Gorman A, Brennan L (2017) The role of metabolomics in determination of new dietary biomarkers. *Proceedings of the Nutrition Society*, 1-8.
106. Schoenheimer R, Rittenberg D (1935) Deuterium as an Indicator in the Study of Intermediary Metabolism. *Science* **82**, 156-157.
107. Umpleby M, Fielding BA (2015) Stable Isotopes in Nutrition Research. In *Nutrition Research Methodologies* [Lovegrove JAH, Sharma L, Lanham-New SA, editors]. West Sussex, UK: John Wiley & Sons, Ltd.
108. O'Brien DM (2015) Stable Isotope Ratios as Biomarkers of Diet for Health Research. *Annual review of nutrition* **35**, 565-594.
109. Lee-Thorp JA (2008) ON ISOTOPES AND OLD BONES. *Archaeometry* **50**, 925-950.
110. Gannes LZ, Martinez del Rio C, Koch P (1998) Natural abundance variations in stable isotopes and their potential uses in animal physiological ecology. *Comp Biochem Physiol A Mol Integr Physiol* **119**, 725-737.
111. Patel PS, Cooper AJ, O'Connell TC *et al.* (2014) Serum carbon and nitrogen stable isotopes as potential biomarkers of dietary intake and their relation with incident type 2 diabetes: the EPIC-Norfolk study. *Am J Clin Nutr* **100**, 708-718.
112. Taylor-Powell E, Steele S (1996) Collecting Evaluation Data: Direct Observation. http://learningstore.uwex.edu/pdf/G3658-5.PDF (accessed March 2017).
113. Simons-Morton BG, Baranowski T (1991) Observation in assessment of children's dietary practices. *J Sch Health* **61**, 204-207.
114. Baglio ML, Baxter SD, Guinn CH *et al.* (2004) Assessment of interobserver reliability in nutrition studies that use direct observation of school meals. *J Am Diet Assoc* **104**, 1385-1392.
115. Kazdin AE (1979) Unobtrusive measures in behavioral assessment. *J Appl Behav Anal* **12**, 713-724.
116. Richter SL, Vandervet LM, Macaskill LA *et al.* (2012) Accuracy and reliability of direct observations of home-packed lunches in elementary schools by trained nutrition students. *J Acad Nutr Diet* **112**, 1603-1607.
117. Jeran S, Steinbrecher A, Pischon T (2016) Prediction of activity-related energy expenditure using accelerometer-derived physical activity under free-living conditions: a systematic review. *Int J Obes (Lond)* **40**, 1187-1197.
118. Andrejašic M (2008) *MEMS Accelerometers*. Ljubljana, Slovenia: Faculty for mathematics and physics, University of Ljubljana.
119. Migueles JH, Cadenas-Sanchez C, Ekelund U *et al.* (2017) Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med*.
120. Ekelund U, Sjostrom M, Yngve A *et al.* (2001) Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc* **33**, 275-281.
121. National Cancer Institute (2017) Dietary Assessment Primer - Measurement Error. https://dietassessmentprimer.cancer.gov/concepts/error/error-types.html (accessed April 2017).
122. Kipnis V, Subar AF, Midthune D *et al.* (2003) Structure of Dietary Measurement Error: Results of the OPEN Biomarker Study. *American Journal of Epidemiology* **158**, 14-21.
123. Gibson RS (2005) *Principles of Nutritional Assessment*, 2nd ed. New York, United States of America: Oxford University Press.
124. National Cancer Institute (2017) Dietary Assesment Primer - Validation. https://dietassessmentprimer.cancer.gov/concepts/validation/ (accessed April 2017).
125. Freedman LS, Schatzkin A, Midthune D *et al.* (2011) Dealing With Dietary Measurement Error in Nutritional Cohort Studies. *JNCI: Journal of the National Cancer Institute* **103**, 1086-1092.
126. Willett WC, Howe GR, Kushi LH (1997) Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* **65**, 1220S-1228S; discussion 1229S-1231S.
127. Laureano G, Torman V, Crispim S *et al.* (2016) Comparison of the ISU, NCI, MSM, and SPADE Methods for Estimating Usual Intake: A Simulation Study of Nutrients Consumed Daily. *Nutrients* **8**, 166.

128. Beaton GH, Milner J, Corey P *et al.* (1979) Sources of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. *Am J Clin Nutr* **32**, 2546-2559.
129. Freedman LS, Tasevska N, Kipnis V *et al.* (2010) Gains in Statistical Power From Using a Dietary Biomarker in Combination With Self-reported Intake to Strengthen the Analysis of a Diet-Disease Association: An Example From CAREDS. *American Journal of Epidemiology* **172**, 836-842.
130. Hansen L, Myhre J, Johansen A *et al.* (2016) *UNGKOST 3 Landsomfattende kostholdsundersøkelse blant elever i 4. -og 8. klasse i Norge, 2015 (UNGKOST 3 Nationally representative dietary survey in Norwegian 4th and 8th graders)*. Oslo, Norway: Norwegian Institute of Public Health.
131. Totland TH, Melnæs BK, Lundberg-Hallén N *et al* (2012) *Norkost 3 En landsomfattende kostholdsundersøkelse blant menn og kvinner i Norge i alderen 18-70 år, 2010-11 (Norkost 3 A nationally representative dietary survey in men and women in Norway aged 18-70 years, 2010-11).* Oslo, Norway: The Norwegian Directorate of Health.
132. Nes M, Frost Andersen L, Solvoll K *et al.* (1992) Accuracy of a quantitative food frequency questionnaire applied in elderly Norwegian women. *Eur J Clin Nutr* **46**, 809-821.
133. Carlsen MH, Karlsen A, Lillegaard IT *et al.* (2011) Relative validity of fruit and vegetable intake estimated from an FFQ, using carotenoid and flavonoid biomarkers and the method of triads. *Br J Nutr* **105**, 1530-1538.
134. Andersen LF, Tomten H, Haggarty P *et al.* (2003) Validation of energy intake estimated from a food frequency questionnaire: a doubly labelled water study. *Eur J Clin Nutr* **57**, 279-284.
135. Carlsen MH, Lillegaard IT, Karlsen A *et al.* (2010) Evaluation of energy and dietary intake estimates from a food frequency questionnaire using independent energy expenditure measurement and weighed food records. *Nutr J* **9**, 37.
136. Andersen LF, Solvoll K, Johansson LR *et al.* (1999) Evaluation of a food frequency questionnaire with weighed records, fatty acids, and alpha-tocopherol in adipose tissue and serum. *Am J Epidemiol* **150**, 75-87.
137. Myhre JB, Loken EB, Wandel M *et al.* (2014) Eating location is associated with the nutritional quality of the diet in Norwegian adults. *Public Health Nutr* **17**, 915-923.
138. Hammer GP, du Prel J-B, Blettner M (2009) Avoiding Bias in Observational Studies: Part 8 in a Series of Articles on Evaluation of Scientific Publications. *Deutsches Ärzteblatt International* **106**, 664-668.
139. Statistics Norway (2014) *Innvandrere og norskfødte med innvandrerforeldre i 13 kommuner (Immigrants and Norwegian-born to immigrant parents in 13 municipalities),* [Høydahl E, editor]. Oslo, Norway: Statistics Norway.
140. Statistics Norway (2015) Population´s level of education, 1 October 2014. http://www.ssb.no/en/utdanning/statistikker/utniv/ (accessed August 2015).
141. Norwegian Institute of Public Health (2015) Norhealth. http://www.norgeshelsa.no/norgeshelsa/?language=en (accessed May 2017).
142. Statistics Norway (2017) Educational attainment of the population, 1 October 2016. https://www.ssb.no/en/utdanning/statistikker/utniv/aar/2016-06-20?fane=tabell&sort=nummer&tabell=270244 (accessed June 2017).
143. Biehl A, Hovengen R, Groholt EK *et al.* (2013) Adiposity among children in Norway by urbanity and maternal education: a nationally representative study. *BMC Public Health* **13**, 842.
144. Statistics Norway (2017) Smoking habits, 2016. https://www.ssb.no/en/royk (accessed June 2017).
145. Smith AF, Baxter SD, Hardin JW *et al.* (2007) Validation-study conclusions from dietary reports by fourth-grade children observed eating school meals are generalisable to dietary reports by comparable children not observed. *Public Health Nutr* **10**, 1057-1066.
146. Baxter SD, Hardin JW, Smith AF *et al.* (2009) Twenty-four hour dietary recalls by fourth-grade children were not influenced by observations of school meals. *J Clin Epidemiol* **62**, 878-885.
147. Baxter SD, Smith AF, Litaker MS *et al.* (2006) Body mass index, sex, interview protocol, and children's accuracy for reporting kilocalories observed eaten at school meals. *J Am Diet Assoc* **106**, 1656-1662.
148. Faulks RM, Southon S (2005) Challenges to understanding and measuring carotenoid bioavailability. *Biochim Biophys Acta* **1740**, 95-100.
149. Burrows TL, Warren JM, Colyvas K *et al.* (2009) Validation of overweight children's fruit and vegetable intake using plasma carotenoids. *Obesity (Silver Spring)* **17**, 162-168.
150. Erlinger TP, Guallar E, Miller IE *et al.* (2001) Relationship between systemic markers of inflammation and serum β-carotene levels. *Archives of Internal Medicine* **161**, 1903-1908.

151. Micozzi MS, Brown ED, Edwards BK *et al.* (1992) Plasma carotenoid response to chronic intake of selected foods and beta-carotene supplements in men. *Am J Clin Nutr* **55**, 1120-1125.
152. Burri BJ, Neidlinger TR, Clifford AJ (2001) Serum carotenoid depletion follows first-order kinetics in healthy adult women fed naturally low carotenoid diets. *J Nutr* **131**, 2096-2100.
153. Plasqui G, Westerterp KR (2007) Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity (Silver Spring)* **15**, 2371-2379.
154. Mannes AE, Soll JB, Larrick RP (2014) The wisdom of select crowds. *J Pers Soc Psychol* **107**, 276-299.
155. Wells JC, Williams JE, Haroun D *et al.* (2009) Aggregate predictions improve accuracy when calculating metabolic variables used to guide treatment. *Am J Clin Nutr* **89**, 491-499.
156. Hills AP, Mokhtar N, Byrne NM (2014) Assessment of physical activity and energy expenditure: an overview of objective measures. *Front Nutr* **1**, 5.
157. Mannini A, Intille SS, Rosenberger M *et al.* (2013) Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise* **45**, 2193-2203.
158. Rosenbaum M, Ravussin E, Matthews DE *et al.* (1996) A comparative study of different means of assessing long-term energy expenditure in humans. *Am J Physiol* **270**, R496-504.
159. Black AE (2000) The sensitivity and specificity of the Goldberg cut-off for EI:BMR for identifying diet reports of poor validity. *Eur J Clin Nutr* **54**, 395-404.
160. Schoeller DA, Ravussin E, Schutz Y *et al.* (1986) Energy expenditure by doubly labeled water: validation in humans and proposed calculation. *Am J Physiol* **250**, R823-830.
161. Berman ES, Melanson EL, Swibas T *et al.* (2015) Inter- and intraindividual correlations of background abundances of (2)H, (18)O and (17)O in human urine and implications for DLW measurements. *Eur J Clin Nutr* **69**, 1091-1098.
162. Speakman JR, Hambly C (2016) Using doubly-labelled water to measure free-living energy expenditure: Some old things to remember and some new things to consider. *Comp Biochem Physiol A Mol Integr Physiol* **202**, 3-9.
163. Subar AF, Freedman LS, Tooze JA *et al.* (2015) Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr* **145**, 2639-2645.
164. National Cancer Institute (2017) Dietary Assessment Primer - 24-hour Dietary Recall https://dietassessmentprimer.cancer.gov/profiles/recall/ (accessed July 2017).
165. Forrestal SG (2011) Energy intake misreporting among children and adolescents: a literature review. *Matern Child Nutr* **7**, 112-127.
166. Burrows TL, Martin RJ, Collins CE (2010) A systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water. *J Am Diet Assoc* **110**, 1501-1510.
167. Freedman LS, Commins JM, Moler JE *et al.* (2014) Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol* **180**, 172-188.
168. Dhurandhar NV, Schoeller D, Brown AW *et al.* (2015) Energy balance measurement: when something is not better than nothing. *Int J Obes (Lond)* **39**, 1109-1113.
169. Baxter SD, Thompson WO, Davis HC *et al.* (1997) Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. *J Am Diet Assoc* **97**, 1293-1298.
170. Spiegelman D, McDermott A, Rosner B (1997) Regression calibration method for correcting measurement-error bias in nutritional epidemiology. *Am J Clin Nutr* **65**, 1179s-1186s.
171. Neuhouser ML, Tinker L, Shaw PA *et al.* (2008) Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *Am J Epidemiol* **167**, 1247-1259.
172. Domel SB, Baranowski T, Leonard SB *et al.* (1994) Accuracy of fourth- and fifth-grade students' food records compared with school-lunch observations. *Am J Clin Nutr* **59**, 218s-220s.
173. Baxter SD, Hitchcock DB, Royer JA *et al.* (2017) Fourth-grade children's dietary reporting accuracy by meal component: Results from a validation study that manipulated retention interval and prompts. *Appetite* **113**, 106-115.
174. Baxter SD, Thompson WO, Litaker MS *et al.* (2002) Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. *J Am Diet Assoc* **102**, 386-395.
175. Baxter SD, Thompson WO, Litaker MS *et al.* (2003) Accuracy of fourth-graders' dietary recalls of school breakfast and school lunch validated with observations: in-person versus telephone interviews. *J Nutr Educ Behav* **35**, 124-134.

176. Baxter SD, Smith AF, Guinn CH *et al.* (2003) Interview format influences the accuracy of children's dietary recalls validated with observations. *Nutr Res* **23**, 1537-1546.
177. Baxter SD, Smith AF, Litaker MS *et al.* (2004) Recency affects reporting accuracy of children's dietary recalls. *Ann Epidemiol* **14**, 385-390.
178. Baxter SD, Guinn CH, Royer JA *et al.* (2009) Accuracy of children's school-breakfast reports and school-lunch reports (in 24-h dietary recalls) differs by retention interval. *Eur J Clin Nutr* **63**, 1394-1403.
179. Baxter SD, Smith AF, Hitchcock DB *et al.* (2015) Effectiveness of Prompts on Fourth-Grade Children's Dietary Recall Accuracy Depends on Retention Interval and Varies by Gender. *J Nutr* **145**, 2185-2192.
180. Livingstone MB, Robson PJ, Wallace JM (2004) Issues in dietary intake assessment of children and adolescents. *Br J Nutr* **92 Suppl 2**, S213-222.
181. Davies VF, Kupek E, de Assis MA *et al.* (2015) Validation of a web-based questionnaire to assess the dietary intake of Brazilian children aged 7-10 years. *J Hum Nutr Diet* **28 Suppl 1**, 93-102.
182. Biltoft-Jensen A, Bysted A, Trolle E *et al.* (2013) Evaluation of Web-based Dietary Assessment Software for Children: comparing reported fruit, juice and vegetable intakes with plasma carotenoid concentration and school lunch observations. *Br J Nutr* **110**, 186-195.
183. The Research Council of Norway (2012) *Hva spiser elevene på skolen? (What are the pupils eating at school?)*. Report from the Research Campaign 2011. Oslo, Norway: The Research Council of Norway.
184. Olmedilla-Alonso B, Granado-Lorencio F, Blanco-Navarro I (2005) Carotenoids, retinol and tocopherols in blood: comparability between serum and plasma (Li-heparin) values. *Clin Biochem* **38**, 444-449.
185. Byers T, Trieber F, Gunter E *et al.* (1993) The accuracy of parental reports of their children's intake of fruits and vegetables: validation of a food frequency questionnaire with serum levels of carotenoids and vitamins C, A, and E. *Epidemiology* **4**, 350-355.
186. Slater B, Enes CC, Lopez RV *et al.* (2010) Validation of a food frequency questionnaire to assess the consumption of carotenoids, fruits and vegetables among adolescents: the method of triads. *Cad Saude Publica* **26**, 2090-2100.
187. Nguyen LM, Scherr RE, Linnell JD *et al.* (2015) Evaluating the relationship between plasma and skin carotenoids and reported dietary intake in elementary school children to assess fruit and vegetable intake. *Arch Biochem Biophys* **572**, 73-80.
188. Neuhouser ML, Rock CL, Eldridge AL *et al.* (2001) Serum concentrations of retinol, alpha-tocopherol and the carotenoids are influenced by diet, race and obesity in a sample of healthy adolescents. *J Nutr* **131**, 2184-2191.
189. Myhre JB, Loken EB, Wandel M *et al.* (2015) Meal types as sources for intakes of fruits, vegetables, fish and whole grains among Norwegian adults. *Public Health Nutr* **18**, 2011-2021.
190. Sharman SJ, Skouteris H, Powell MB *et al.* (2016) Factors Related to the Accuracy of Self-Reported Dietary Intake of Children Aged 6 to 12 Years Elicited with Interviews: A Systematic Review. *Journal of the Academy of Nutrition and Dietetics* **116**, 76-114.
191. Fisher JO, Johnson RK, Lindquist C *et al.* (2000) Influence of body composition on the accuracy of reported energy intake in children. *Obes Res* **8**, 597-603.
192. Lioret S, Touvier M, Balin M *et al.* (2011) Characteristics of energy under-reporting in children and adolescents. *Br J Nutr* **105**, 1671-1680.
193. Murakami K, Miyake Y, Sasaki S *et al.* (2012) Characteristics of under- and over-reporters of energy intake among Japanese children and adolescents: The Ryukyus Child Health Study. *Nutrition* **28**, 532-538.
194. Livingstone MB, Robson PJ (2000) Measurement of dietary intake in children. *Proc Nutr Soc* **59**, 279-293.
195. Tabacchi G, Amodio E, Di Pasquale M *et al.* (2014) Validation and reproducibility of dietary assessment methods in adolescents: a systematic literature review. *Public Health Nutr* **17**, 2700-2714.
196. Rangan AM, Flood VM, Gill TP (2011) Misreporting of energy intake in the 2007 Australian Children's Survey: identification, characteristics and impact of misreporters. *Nutrients* **3**, 186-199.
197. Lillegaard IT, Andersen LF (2005) Validation of a pre-coded food diary with energy expenditure, comparison of under-reporters v. acceptable reporters. *Br J Nutr* **94**, 998-1003.
198. Hebert JR, Clemow L, Pbert L *et al.* (1995) Social desirability bias in dietary self-report may compromise the validity of dietary intake measures. *Int J Epidemiol* **24**, 389-398.
199. Di Noia J, Cullen KW, Monica D (2016) Social Desirability Trait Is Associated with Self-Reported Vegetable Intake among Women Enrolled in the Special Supplemental Nutrition Program for Women, Infants, and Children. *Journal of the Academy of Nutrition and Dietetics* **116**, 1942-1950.

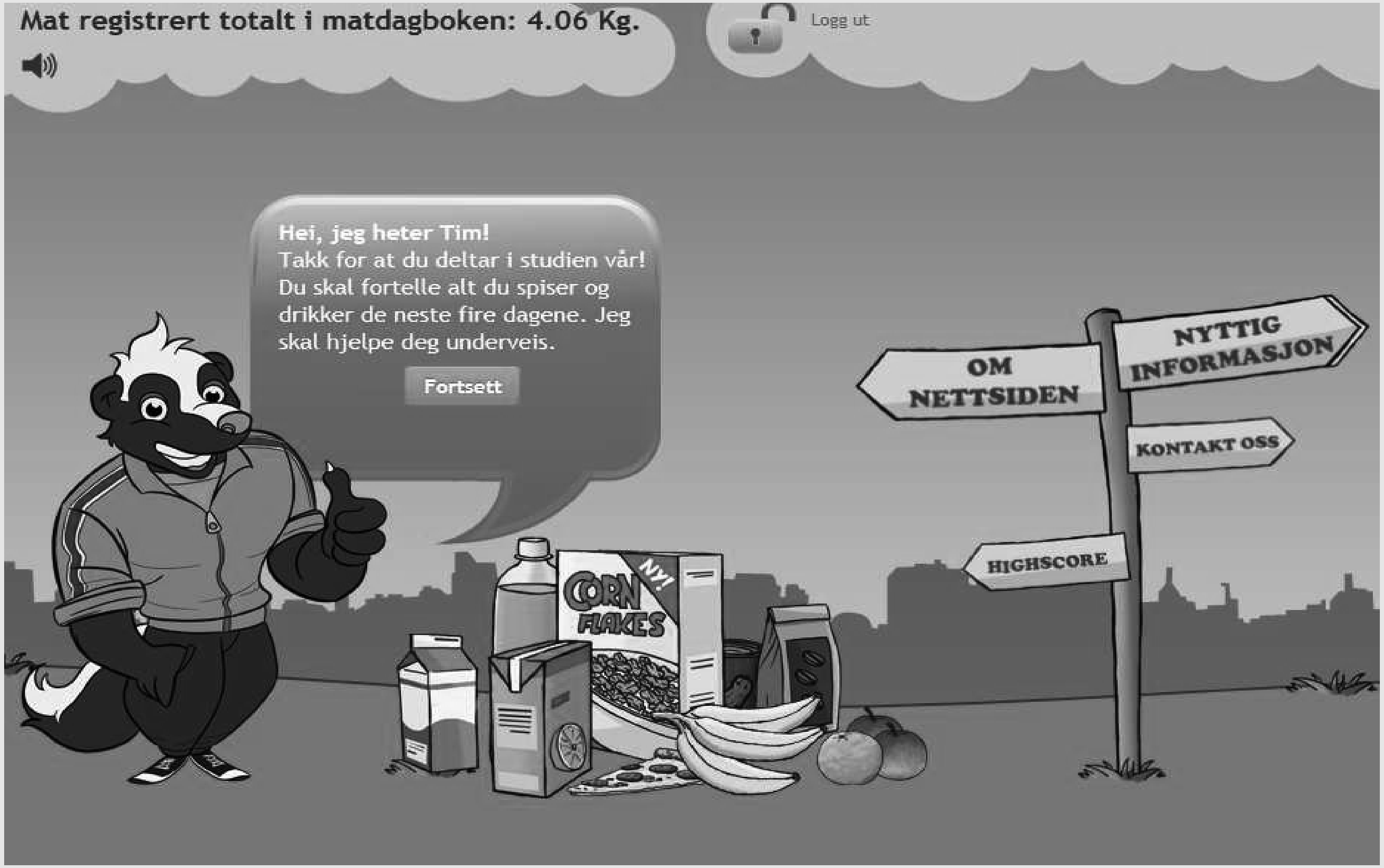
200. The Norwegian Directorate of Health (2014) *Anbefalinger om kosthold, ernæring og fysisk aktivitet (Recommendations on diet, nutrition and physical activity)*. Report IS-2170. Oslo, Norway: The Norwegian Directorate of Health.
201. Hébert JR (2016) Social Desirability Trait: Biaser or Driver of Self-Reported Dietary Intake? *Journal of the Academy of Nutrition and Dietetics* **116**, 1895-1898.
202. Subar AF, Crafts J, Zimmerman TP *et al.* (2010) Assessment of the accuracy of portion size reports using computer-based food photographs aids in the development of an automated self-administered 24-hour recall. *J Am Diet Assoc* **110**, 55-64.
203. Greenlaw C, Brown-Welty S (2009) A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. *Eval Rev* **33**, 464-480.
204. Forster H, Fallaize R, Gallagher C *et al.* (2014) Online dietary intake estimation: the Food4Me food frequency questionnaire. *J Med Internet Res* **16**, e150.
205. Gonzalez Carrascosa R, Garcia Segovia P, Martinez Monzo J (2011) Paper and pencil vs online self-administered food frequency questionnaire (FFQ) applied to university population: a pilot study. *Nutr Hosp* **26**, 1378-1384.
206. Mangunkusumo RT, Duisterhout JS, de Graaff N *et al.* (2006) Internet versus paper mode of health and health behavior questionnaires in elementary schools: asthma and fruit as examples. *J Sch Health* **76**, 80-86.
207. Mangunkusumo RT, Moorman PW, Van Den Berg-de Ruiter AE *et al.* (2005) Internet-administered adolescent health questionnaires compared with a paper version in a randomized study. *J Adolesc Health* **36**, 70 e71-76.
208. Boushey CJ, Kerr DA, Wright J *et al.* (2009) Use of technology in children's dietary assessment. *Eur J Clin Nutr* **63**, S50-S57.
209. Biltoft-Jensen A, Hjorth MF, Trolle E *et al.* (2013) Comparison of estimated energy intake using Web-based Dietary Assessment Software with accelerometer-determined energy expenditure in children. *Food Nutr Res* **57**.
210. Svensson A, Waling M, Backlund C *et al.* (2012) Overweight and Obese Children's Ability to Report Energy Intake Using Digital Camera Food Records during a 2-Year Study. *J Nutr Metab* **2012**, 247389.
211. Liu B, Young H, Crowe FL *et al.* (2011) Development and evaluation of the Oxford WebQ, a low-cost, web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies. *Public Health Nutr* **14**, 1998-2005.
212. Touvier M, Kesse-Guyot E, Mejean C *et al.* (2011) Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* **105**, 1055-1064.
213. Galea S, Tracy M (2007) Participation rates in epidemiologic studies. *Ann Epidemiol* **17**, 643-653.
214. Beaton GH (1994) Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am J Clin Nutr* **59**, 253s-261s.
215. Hébert JR, Hurley TG, Steck SE *et al.* (2014) Considering the Value of Dietary Assessment Data in Informing Nutrition-Related Health Policy. *Advances in Nutrition: An International Review Journal* **5**, 447-455.
216. Satija A, Yu E, Willett WC *et al.* (2015) Understanding nutritional epidemiology and its role in policy. *Adv Nutr* **6**, 5-18.

Appendix I

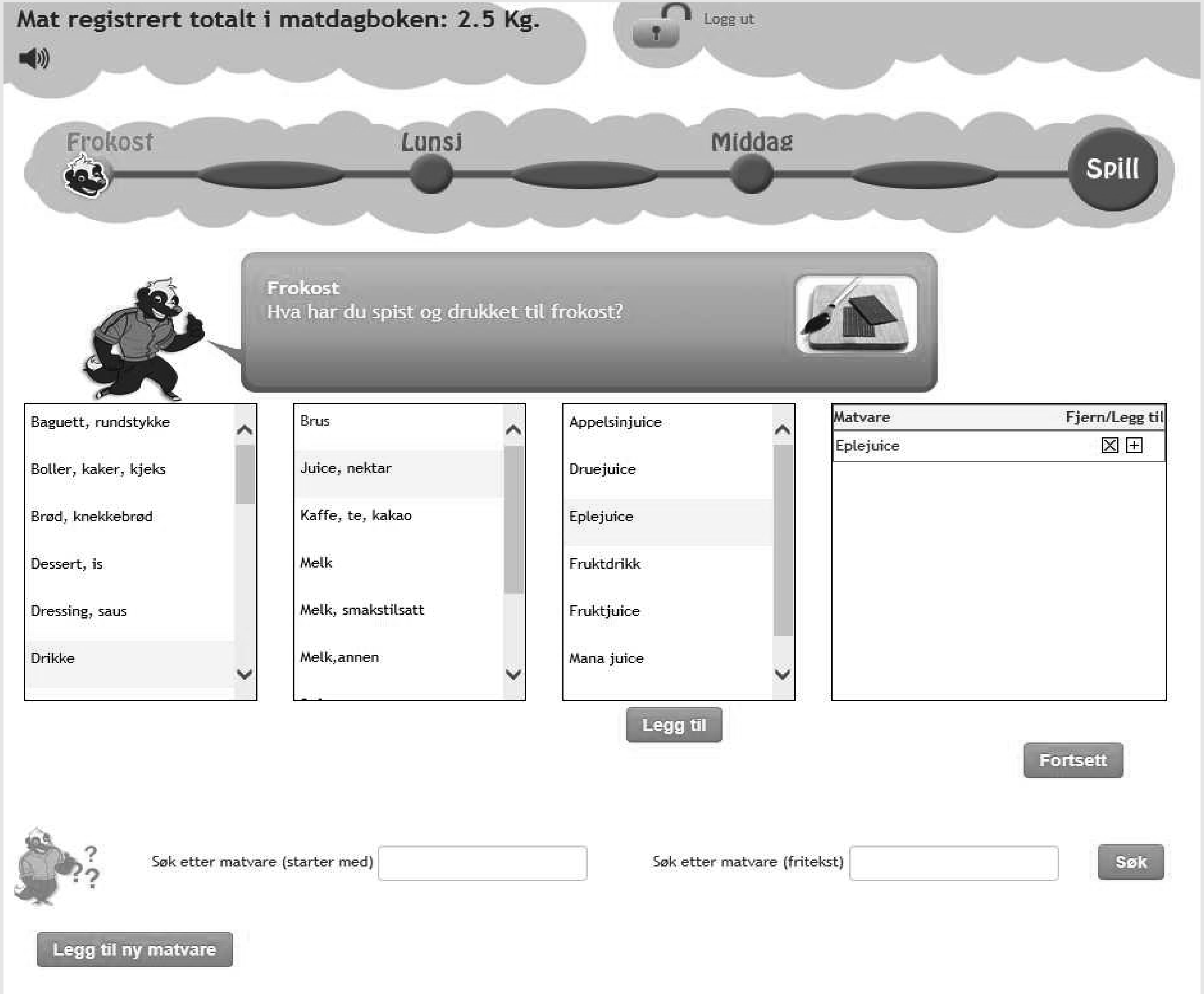
Selected screenshots from the WebFR

The WebFR login page

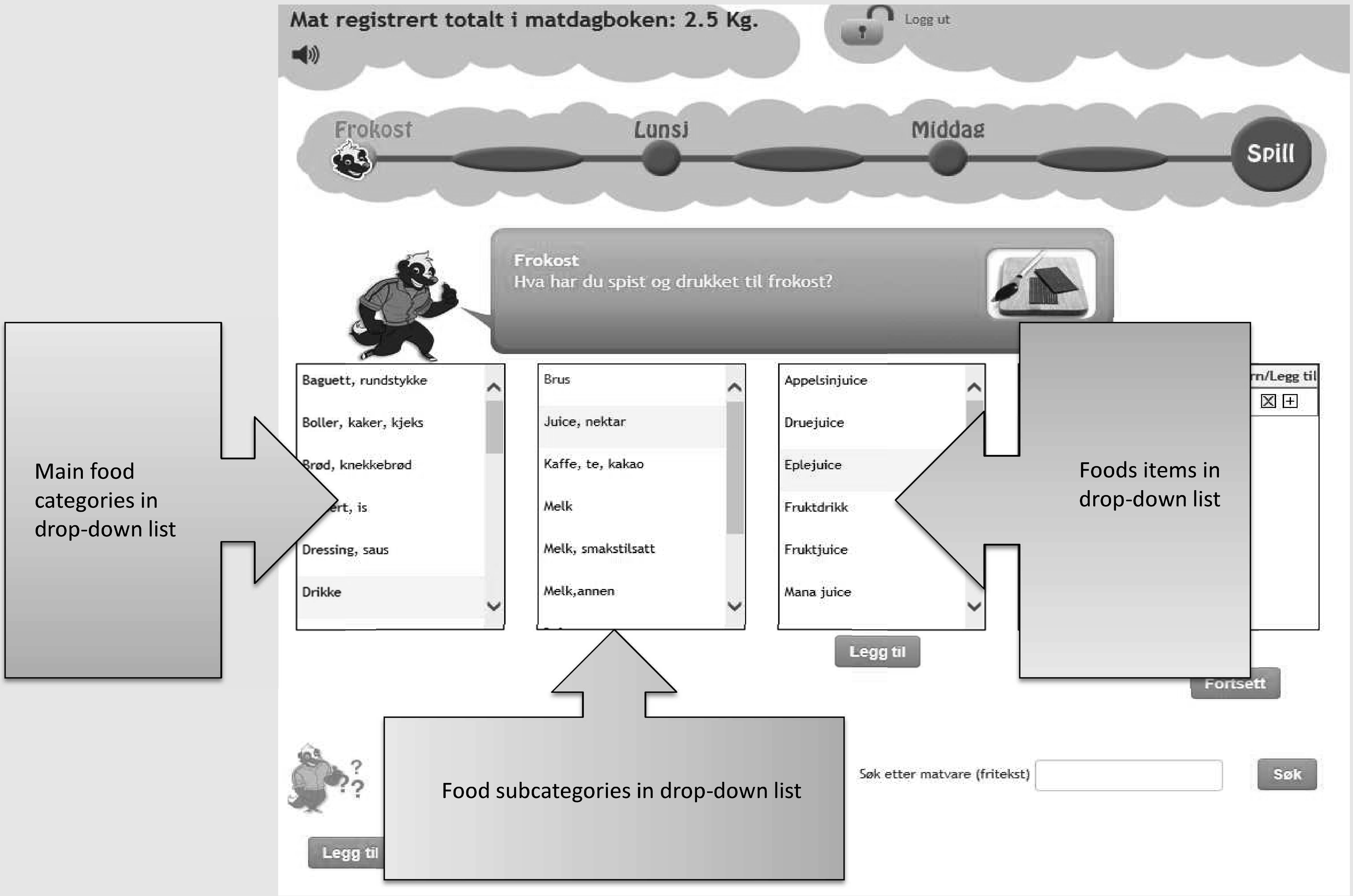
A voice-assisted cartoon character guides the participant through the recordings



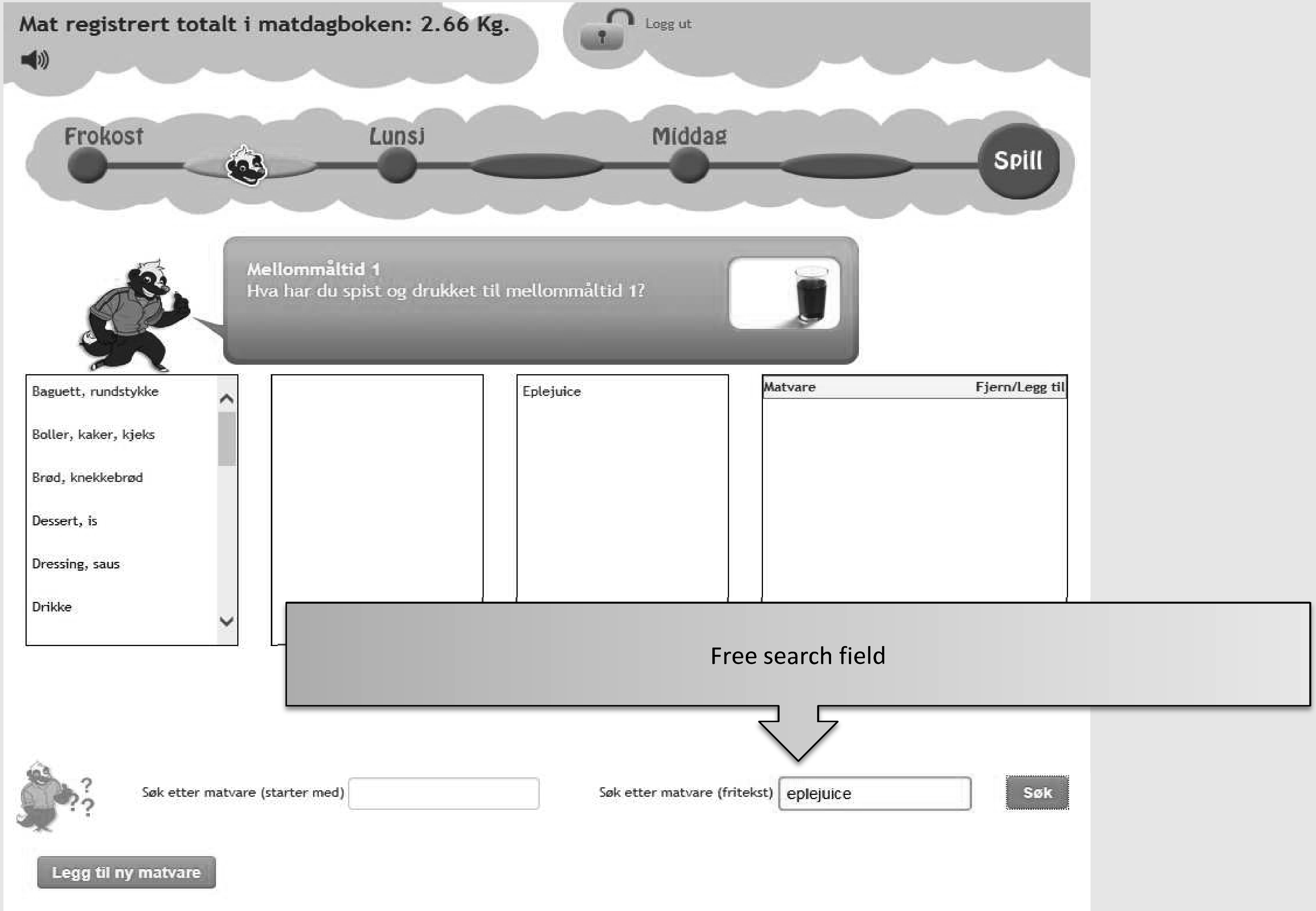
Drop-down lists and a free text search field



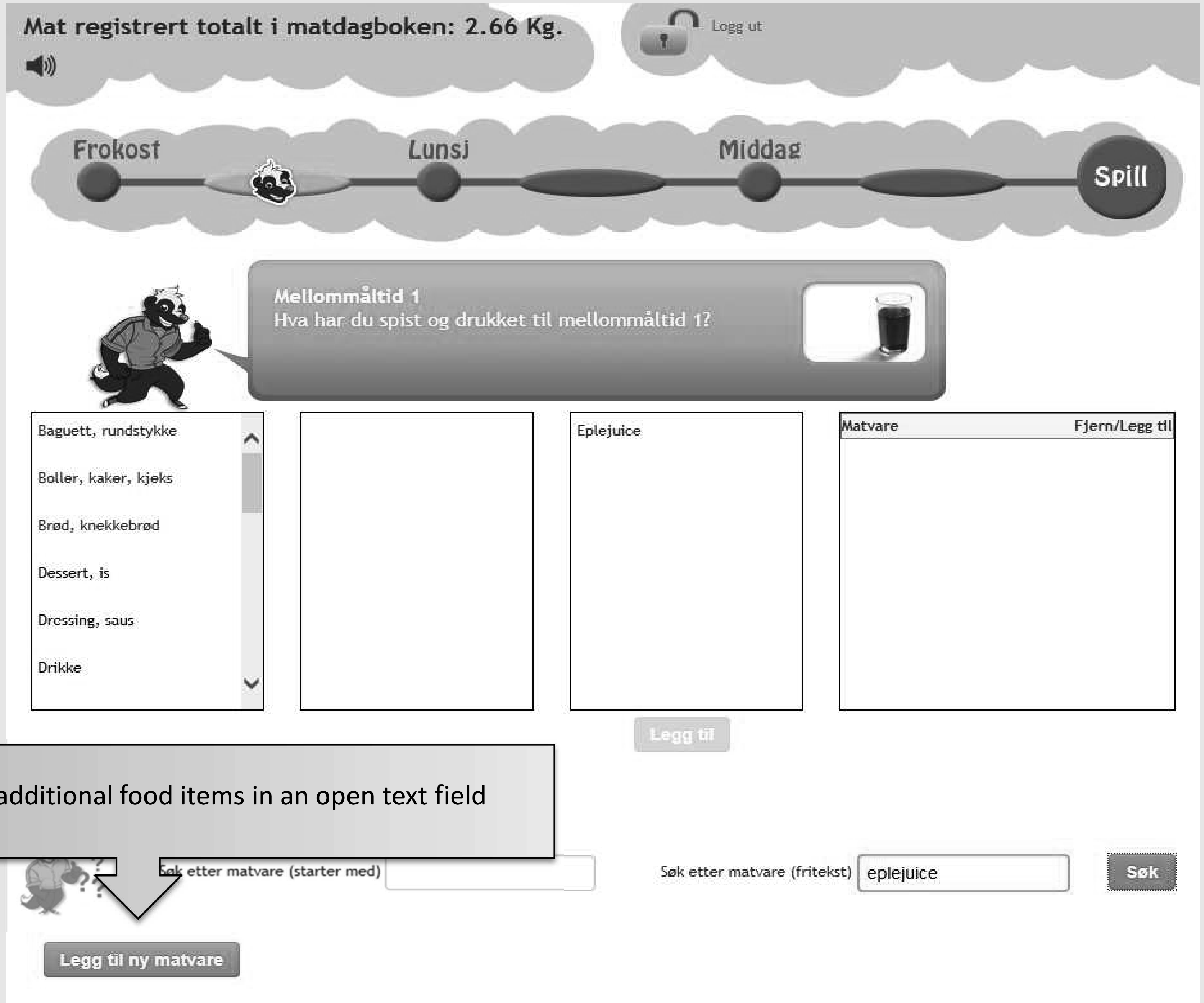
Drop-down lists



A free text search field for food items

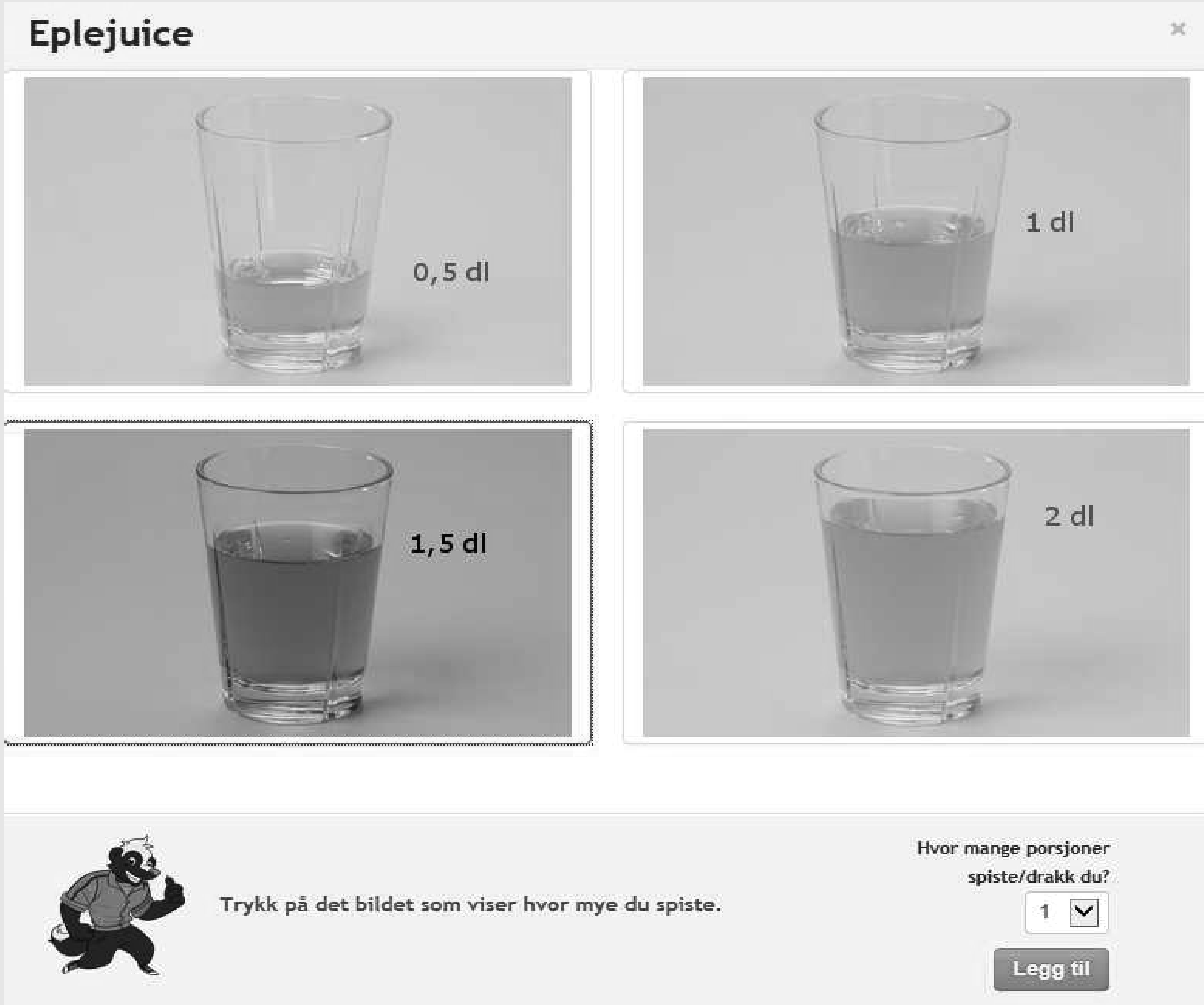


Adding new foods in an open text field

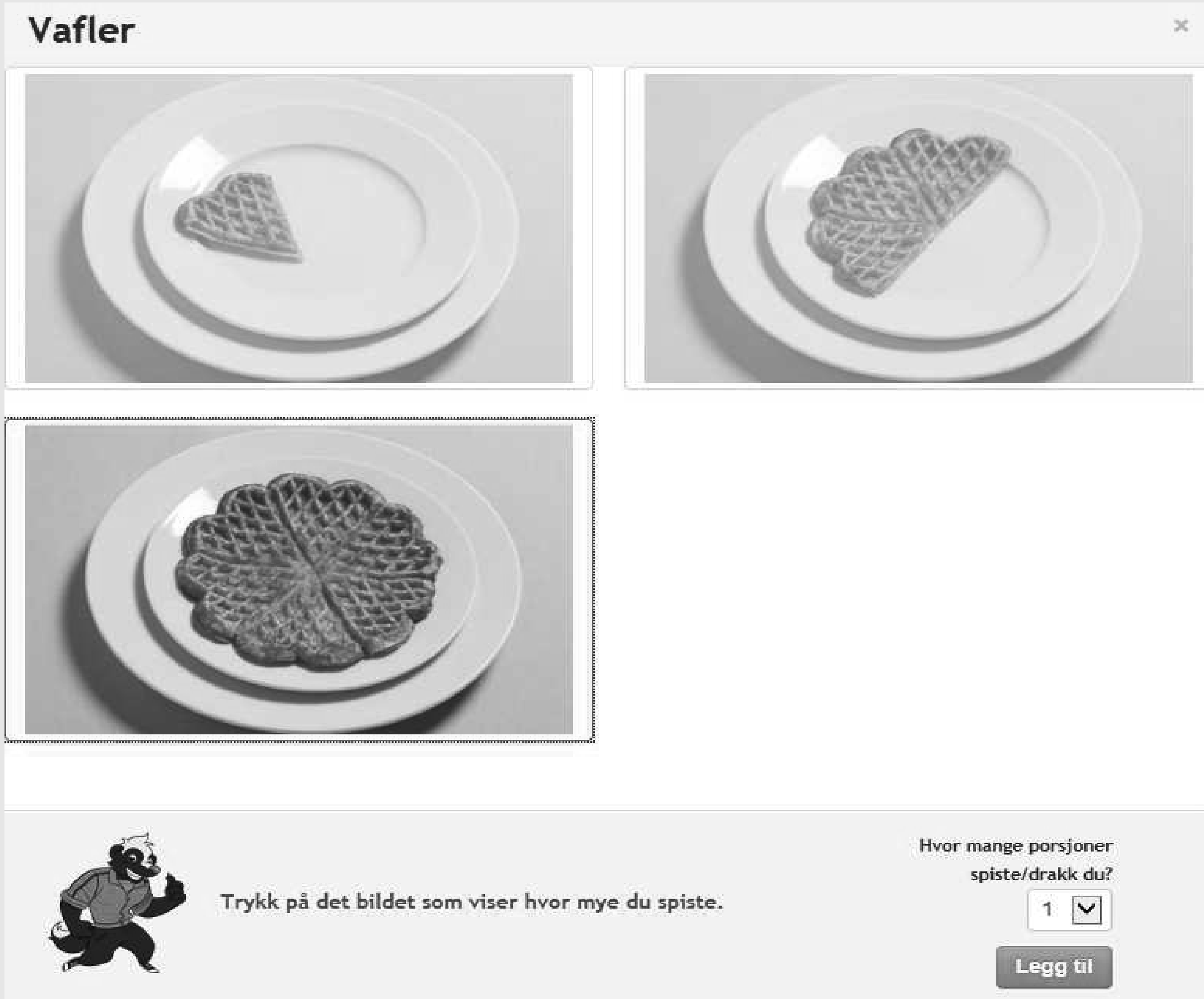


Possibility to add additional food items in an open text field

Selection of portion size from photo series



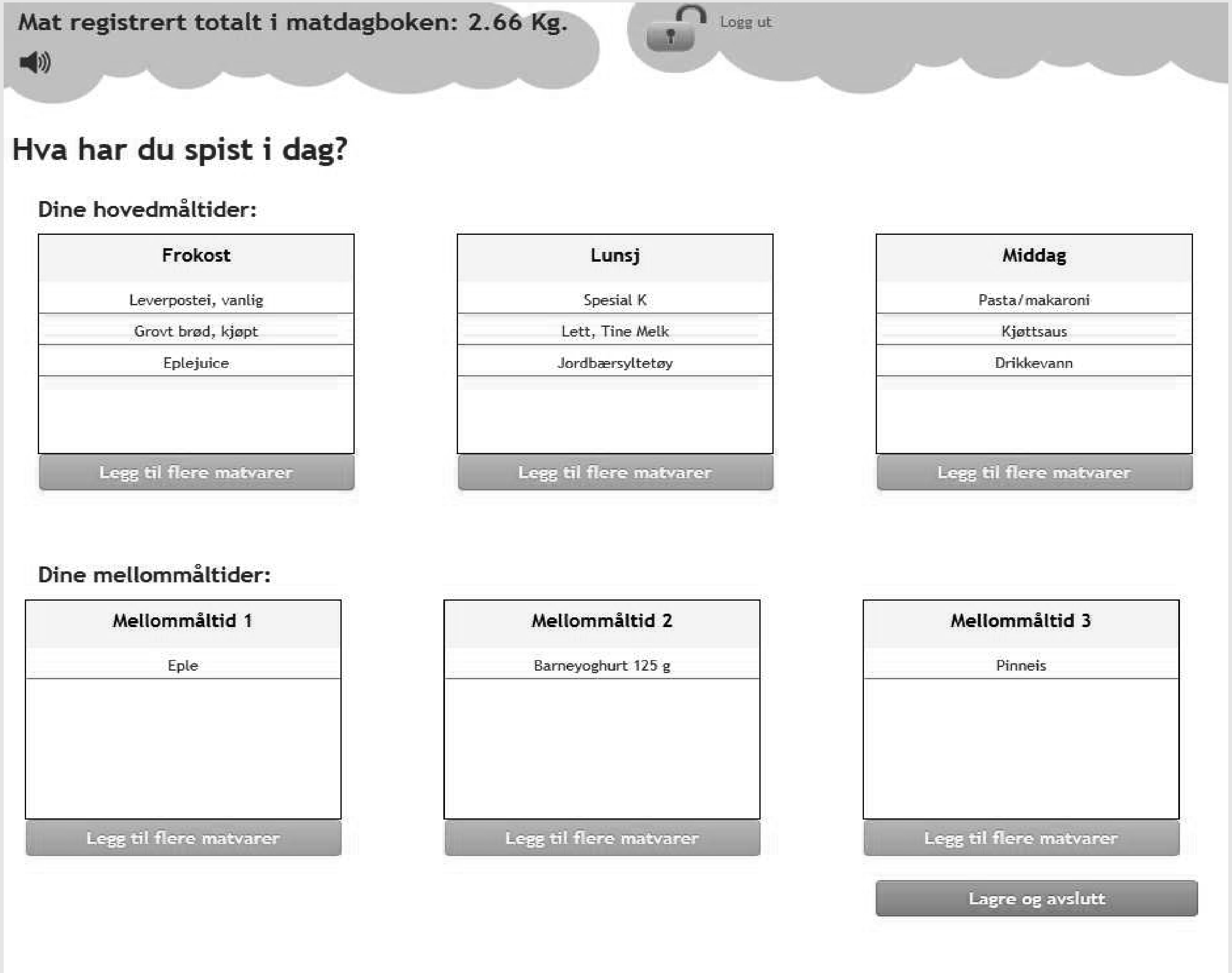
Selection of portion size from photo series



Selection of portion size from photo series

Prompting using pop-ups

Summary at the end of each recording day



Appendix II

Selected screenshots from the WebFFQ

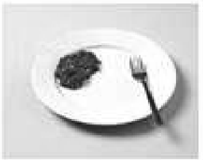
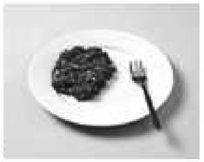
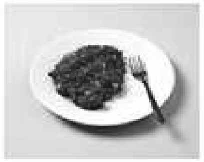
Images are used as an aid for portion size estimations

Hvor ofte spiser du kjøttsaus, gryterett med kjøttdeig eller karbonadedeig av storfe/svin? *

○ Aldri ○ Sjelden ◉ 1/mnd ○ 2/mnd ○ 3/mnd ○ 4/mnd ○ 5-6/mnd ○ 7-8/mnd

○ ≥ 9/mnd

Hvor stor porsjon spiser du? *

○ A  ○ B  ○ C  ○ D 

Hvor mange porsjoner spiser du til et måltid? *

Velg ...

Hvor ofte spiser du kyllinggryte? *

○ Aldri ○ Sjelden ◉ 1/mnd ○ 2/mnd ○ 3/mnd ○ 4/mnd ○ 5-6/mnd ○ 7-8/mnd ○ ≥ 9/mnd

Hvor stor porsjon spiser du? *

○ A 
○ B 
○ C 
○ D 

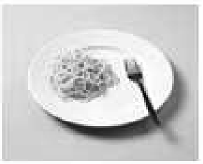
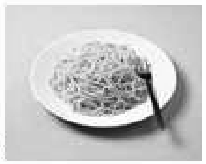
Hvor mange porsjoner spiser du til et måltid? *

Velg ...

Hvor ofte spiser du spaghetti, makaroni, pasta? *

○ Aldri ○ Sjelden ○ 1-3/mnd ◉ 1/uke ○ 2-3/uke ○ 4-5/uke ○ 6-7/uke ○ ≥ 8/uke

Hvor stor porsjon spiser du? *

○ A 
○ B 
◉ C 
○ D 

Hvor mange porsjoner spiser du til et måltid? *

Velg ...

Hvor ofte spiser du brokkoli? *

○ Aldri ○ Sjelden ○ 1-3/mnd ◉ 1/uke ○ 2-3/uke ○ 4-5/uke ○ 6-7/uke ○ ≥ 8/uke

Hvor stor porsjon spiser du? *

○ A  ○ B  ○ C  ○ D 

Hvor mange porsjoner spiser du til et måltid? *

Velg ...

Hvor ofte spiser du blandet salat? *

○ Aldri ○ Sjelden ○ 1-3/mnd ○ 1/uke ◉ 2-3/uke ○ 4-5/uke ○ 6-7/uke ○ ≥ 8/uke

Hvor stor porsjon spiser du? *

○ A  ○ B  ○ C  ○ D 

Hvor mange porsjoner spiser du til et måltid? *

Velg ...