BMJ Open

A web-based approach of life events, habits, and health outcomes in the Nutritionists' Health Study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012081
Article Type:	Protocol
Date Submitted by the Author:	28-Mar-2016
Complete List of Authors:	Folchetti, Luciana; University of Sao Paulo, School of Public Health Silva, Isis; University of Sao Paulo, School of Public Health Almeida-Pititto, Bianca; Universidade Federal de Sao Paulo Escola Paulista de Medicina, Department of Preventive Medicine Ferreira, Sandra Roberta G.; University of Sao Paulo, Department of Epidemiology, School of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Public health, Research methods, Nutrition and metabolism
Keywords:	Nutritionists' Health Study, web-based system, early-life events, gut microbiota, cardiometabolic risk factors

A web-based approach of life events, habits, and health outcomes in the Nutritionists' Health Study

Luciana Gavilan Dias Folchetti[1], Isis Tande da Silva[1], Bianca de Almeida-Pititto[2], Sandra Roberta. G. Ferreira[3]

1 Department of Nutrition, School of Public Health, University of Sao Paulo, São Paulo, Brazil

2 Department of Preventive Medicine, Federal University of Sao Paulo, São Paulo, Brazil

3 Department of Epidemiology, School of Public Health, University of Sao Paulo, São Paulo, Brazil

Correspondence

Prof. Sandra Roberta G. Ferreira

Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo

Av. Dr. Arnaldo, 715 – CEP 01246-904 – São Paulo, SP, Brasil

Fone 55 11 3061-7870 E-mail: sandrafv@usp.br

KEYWORDS

Nutritionists' Health Study, web-based system, early-life events, gut microbiota, cardiometabolic risk factors

Word Count: 2,928

ABSTRACT

INTRODUCTION Non-communicable chronic diseases (NCCDs) represent a burden for public health. Alongside, the established cardiometabolic risk factors – high blood pressure and disorders of glucose and lipid metabolism – living habits and nutritional status at different stages of life are seen as contributors for this scenario. Gut microbiota composition and subclinical inflammation have been pointed out as underlying mechanisms of metabolic disorders and NCCDs. Studies involving health professionals have brought relevant contributions to the knowledge about risk factors. Technological advances facilitate data collection and analysis for big samples. Web-based cohort study, able to identify these risk factors, is highly desirable. The objective of the Brazilian Nutritionists' Health Study – NutriHS is to gather online information on early-life events, daily habits, emergent cardiometabolic risk factors and health outcomes of a specific subset of the Brazilian population. **METHODS AND ANALYSIS** The NutriHS, developed at the School of Public Health-University of Sao Paulo, Brazil, is a research initiative that enrolls undergraduates of nutrition courses from Brazilian universities and graduated volunteers. A web-based self-administered system was designed to collect health-related data. After fulfilling online questionnaires (socioeconomic, early-life events and lifestyle data), participants are invited to a clinical visit for physical examination and lab procedures (blood sampling, feces collection and body composition). At a 3-year interval, they will be invited to repeat similar procedures. **ETHICS AND DISSEMINATION** The NutriHS research protocol was approved by the Institutional Ethics Committee and is providing promising data which contribute to the understanding of pathophysiological links between early life events, body composition, gut microbiota and inflammatory and metabolic risk profile. The combination of a friendly tool with the innovative purposes of NutriHS offers a remarkable resource for testing hypotheses about mechanisms of nutrition-related diseases and further planning of preventive programs in the field of the public health.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- A web-based prospective study for collection of self-reported multidisciplinary data

- Low cost for development and potential of amplification of the web-based system
- Availability of self-reported and measured data which allows validation study
- Innovative data to be associated with early outcomes
- Characteristics of the population sample which has skills to accurate answers to self-reported technical questionnaires and potential high compliance in the long-term study
- Language-related limitation offered in the web-based system (only in Portuguese)
- Non-representative sample of general population

INTRODUCTION

Numerous factors from intrauterine life and birth determine adult health and well-being. Monitoring of early-life events and daily habits has provided tools to protect individual health from major damage and other threats. Health outcomes depend on a combination of genetic and environmental factors.

Delivery conditions, birth weight, and early events in infancy have been involved in the genesis of future diseases [1-4]. Such a relationship, initially raised by Barker [5], has been further confirmed [6-9], although underlying mechanisms are not yet completely understood. More recently, the types of delivery and child nutrition have been shown to influence the gut microbiota composition which, in turn, modulates immune system and the risk of diseases [10-13].

Epidemiological and clinical studies have contributed towards identifying modifiable risk factors and deepening the understanding of the pathophysiology of non-communicable chronic diseases (NCCDs), such as obesity, type 2 diabetes mellitus, and cardiovascular disease. These lifestyle-related morbidities represent a heavy burden for public health systems due to complications, increased mortality, and high costs [14-16].

Knowledge of risk factors for NCCDs, collected prospectively, is essential for enhancing evidence and planning interventions. Epidemiological studies involving health professionals have brought significant contributions to the knowledge about the role of behavioral factors in health outcomes [17,18]. This kind of study often requires large population samples, followed for long periods, which generate a significant amount of data, high costs, and the need for trained professionals [15-18]. More commonly, health surveys have been based on paper-and-pencil questionnaires.

Printed questionnaires-based surveys are complex and time-consuming, and their quality depends on accuracy in data collecting, processing and uploading to create a consistent database before analysis. Despite extending study duration and representing a financial burden, these steps are crucial for adequate data analysis to achieve reliable conclusions [19,20].

Technological advances can minimize errors and duplicated information, reduce costs, as well as provide the almost immediate construction of a database using

bioinformatic tools [20,21]. Multiple and diverse software for data analysis tailored to specific questions are available, but very few are designed specifically to investigate the role of early-life events and daily habits.

The Internet is a promising tool for overcoming some of the logistic and financial feasibility constraints in research, and is gaining considerable popularity [21]. In contrast to other fields – such as in psychological and marketing research – in epidemiology, less than 1% of the reported studies is internet-assisted [14]. Successful examples of large-scale studies in which enrolment and follow-up occur over the Internet are the Millennium Cohort Study [22], the Nurses and Midwives e-Cohort Study [23], and the Danish Web-based Pregnancy Planning Study [24]. In nutritional epidemiology, there is the NutriNet-Santé Study, a web-based prospective study launched in France in 2009 [25].

In middle-income countries, similar initiatives using the Internet as a tool for data collection are scarce, and when considering the inclusion of health professionals they are inexistent. The development of web-based surveys able to identify environmental risk factors in developing countries with wide geographic area and limited resources would be highly desirable. As far as we know, the use of self-administered online questionnaires for this kind of population has not yet been reported.

The purpose of the current study was to describe the methodology of the Nutritionists' Health Study (NutriHS), a prospective Web-based study developed at the School of Public Health, University of São Paulo (FSP/USP), Brazil, to gather information on early-life events, daily habits, emergent cardiometabolic risk factors, and health outcomes of a specific subset of the Brazilian population – undergraduates and graduates in Nutrition. The outcome of major interest is excessive body weight and obesity-related morbidities.

METHODS AND ANALYSIS

Implementation of the NutriHS study included the following steps: 1) development of the web-based system; 2) collection of data related to early-life events, lifestyle habits, anthropometry, and body composition; 3) assessment of the cardiovascular profile based on traditional and emerging risk factors; 4) analysis of the association of

early-life events and lifestyle habits with the gut microbiota composition, and cardiovascular risk factors.

- **The Web-based system**

With the support of the FSP/USP Informatics Center, a dedicated Web-based online self-administered system (http://www.fsp.usp.br/nutrihs/index.html) was developed, named *e*-NutriHS, to collect health-related data based on previously validated printed questionnaires. Details on its development have been described elsewhere [26]. In brief, the Web-based system was built using free programming languages, such as HyperText Markup Language (HTML), Cascading Style Sheets (CSS), JavaScript, Hypertext P reprocessor (PHP), Structured Query Language (SQL), and the MySQL for database management system.

The colours chosen for NutriHS website, along with logotype and slogan designed, call attention to the need for increased knowledge on the role of nutrition for health.

The development also included skip patterns to hide questions (depending on the person's answer to a prior question), randomly organized questions, and validation checks providing additional information to simplify and help responding, pointing out questions that needed attention. The *e*-NutriHS data can be transferred to Excel format, allowing further analyses. Since the database is located at the FSP/USP server, its security was guaranteed.

Prior to releasing the *e*-NutriHS to participants, a prototype was created (26). Twenty volunteers, graduates at the FSP/USP who have been working in the health area, aged 25 to 60 years, tested the system and checked its ease-of-use, data entry, and website design. Based on their opinions, a user-friendly version of *e*-NutriHS was released.

The website provides details on the study purposes, the team involved, and contact information, and also represents an opportunity for learning with recent research reports from international literature. In order to increase awareness of the NutriHS, a social network page on the Internet was created (https://www.facebook.com/nutrihs) with a direct link to the *e*-NutriHS homepage. This website plays an important role in stimulating dialogue among researchers and participants.

- **Study design, recruitment, and inclusion**

Given the growing recognition of the importance of the life course approach for understanding the determinants of NCCDs, the NutriHS was designed to be a cohort study in which retrospective and prospective data are being collected. The study was designed to include all undergraduates and graduates of Nutrition courses, aged ≥18 years. Undergraduates of Nutrition courses from Brazilian universities and graduate volunteers were selected based on their potential skill to navigate the Internet and on the expected high quality of data provided. Exclusion criteria are pregnancy and no access to the Internet.

Volunteer recruitment was done using multiple strategies. Invitations were made to the university staffs, lectures for undergraduates or graduates were offered, news was published on the FSP/USP website and on the NutriHS social network page, and banners were placed in strategic places.

After signing the electronic informed consent, users are required to fill out personal data and to create a unique login and password. Then, their access to questionnaires is released. Once questionnaires completed, participants are instructed to schedule, on the *e*-NutriHS, a face-to-face visit at the Health Center for clinical examination and biological sampling. After a 3-year interval, they will be invited to complete a similar set of questionnaires.

- **Questionnaires and variables of interests**

The baseline set of questionnaires deals with socioeconomic, early-life events, and lifestyle data, including physical activity and eating habits. The order for filling out these questionnaires is chosen by the participant.

1. Demographic, socioeconomic lifestyle, and clinical questionnaires

The structured questionnaires obtained data regarding skin color, educational level, marital status, person(s) they live with, parents or guardians, family income, number of people in the house, work, work hours, use of medications, medical history, familial medical history, a restrictive diet, alcohol consumption, tobacco consumption, height, and weight.

2. Early-life events questionnaire

The contents of this questionnaire allow one to obtain retrospective information about the mothers of the participants, their pregnancies, and data related to birth and childhood, as well as maternal health and educational level during the pregnancy period.

Data collected are divided into two parts regarding mothers and participants. Mother's age at the participant's birth, educational level, occupation, height, weight before pregnancy, use of medications, pregnancy conditions, such as weight gain and health problems, are reported. Also, participants are asked to answer as to prematurity, twins, birth weight, breastfeeding, time of breastfeeding, time of introduction of formula, medications used, weight gain during childhood and adolescence, fruit and vegetable consumption, and physical activity during childhood.

3. Physical activity questionnaire

The short version of the International Physical Activity Questionnaire (IPAQ), validated for the Brazilian population [27], is being employed. Intensity of physical activity (walking, moderate and intense), frequency (per week), and duration are obtained. These data enable estimates of total physical activity as quantitative or categorical variables (low, moderate, and high). A sedentary pattern has been assessed by time spent sitting.

4. Dietary assessment

Diet is assessed using a validated 102-item food frequency questionnaire [28] regarding eating habits over the last year. Data inserted into the *e*-NutriHS are immediately transformed into the equivalent food amounts for one day. The system is able to provide reports of macro- and micronutrients, based on the USDA National Nutrient Database for Standard Reference.

5. Clinical assessment

Completion of questionnaires is required in order to access the link for scheduling the face-to-face visit to the FSP/USP Public Health Center. A couple of days before the visit, an e-mail with instructions for biological sampling and body composition determination is sent to each participant. This visit represents an opportunity to double-check previously collected selected data.

- ***Anthropometry and blood pressure***

Weight is obtained on a digital scale accurate to 100g and height to the nearest 0.5 cm, and they are used to determine the Body Mass Index. Waist circumference is taken at the midpoint between the lowest rib and the iliac crest. Body composition is being determined by dual-emission x-ray absorptiometry (DXA Lunar GE) that allows accurate estimates of body compartments, including visceral fat. Sitting blood pressure levels are measured using an automatic device.

- ***Biochemical data***

Blood samples are collected for plasma glucose and lipid profile, determined using the hexokinase method and enzymatic colorimetric assay, respectively. Aliquots are frozen at -80^{o}C for further determinations of apolipoproteins, hormones, and inflammatory markers. Insulin is determined by enzyme-linked immunoenzymatic assay (ELISA) and high-sensitivity C-reactive protein by immunochemistry. ELISA kits are also used for the measurements of cytokine concentrations. Insulin resistance will be estimated by the HOMA-IR index [29]. Extra aliquots are stored for future analyses of interest to the NutriHS.

- ***Metagenomic analysis of gut microbiota***

Participants are instructed to collect a fecal sample and refrigerate it before the visit. Aliquots are frozen until the metagenomic analysis. The pyrosequencing of 16S subunit ribosomal amplicons will be performed using next-generation sequencing technology. After obtaining the sequences, they are subjected to phylogenetic analysis to characterize the fecal microbiota. Bioinformatic procedures, comparing the sequences extracted from the fecal samples with standardized databases, allow determination of the abundance of bacteria belonging to phyla and genera.

- **Data and statistical analysis plan**

The NutriHS planned to collect prospectively data on emergent lifestyle-related risk factors of a specific subset of the healthy Brazilian population. For its cross-sectional phase, data analyses included the description of the population sample to be followed. Additionally, a validation sub-study of the web-collected self-reported data is being conducted. A number of statistical analyses has been planned to test associations of knowledge in nutrition and changes in lifestyle, as well as in body composition and biochemical profile.

Analyses are being performed using IBM SPSS Statistics V. 20.0. Collected measures and derived variables will be summarized using means and standard deviation (SD), medians and interquartile range (IQR), and/or frequencies. Student *t* test or one-way analysis of variance when comparing more than two groups will be used for continuous variables with normal distribution. Pearson correlation coefficient will be employed to test correlations. Correspondent non-parametric tests will use for non-normal distributed variables. For validation studies, agreement has been analyzed by Kappa statistics and Bland-Altman plot method. For multivariate analyses, generalized linear regression models will be used or logistic regression for dichotomous variables. For all tests, a p-value < 0.05 will be considered significant.

ETHICS AND DISSEMINATION

Ethics

The NutriHS study protocol and study documentation, including all questionnaires and the informed consent, was approved by the Institutional Ethics Committee, Committee for Ethics in Research, School of Public Health, University of Sao Paulo. The substantiated opinions were registered by the numbers: Of. COEP 991.542/15 and Of. COEP 257.513/13.

When accessing the *e*-NutriHS, participants must sign an electronic informed consent before the registration process. Individual privacy is respected and data will be stored in *e*-NutriHS database following the regulations applied to security. All publications will respect confidentiality.

Discussion and Dissemination

Analyses involving the issue cardiovascular risk factors & public health, anchored in population-based studies or specific population subsets have been conducted in developed countries. Such studies are of great relevance for proposals and formulation of public policies tailored to local situations. Cohorts of health professionals, like the Nurses' Health Study I and II (http://www.channing.harvard.edu/nhs/) and the Health Professionals Follow-up Study (https://www.hsph.harvard.edu/hpfs/), which included thousands of participants in North America, brought remarkable scientific contributions for identifying and

intervening in major risk factors for NCCDs [17-18]. One strong point of studies involving professionals is the high quality of the specific data collected.

The current Brazilian epidemiological scenario of increasing longevity and nutritional transition highlights the importance of obesity-related chronic diseases for mortality. Along with the traditional cardiometabolic risk factors – hypertension and impaired metabolism of glucose and lipids – lifestyle and nutritional status in different stages of life, including the intrauterine period, have been pointed out as relevant determinants of this situation [3,8,30]. Scientific evidence has associated low birthweight, cesarean delivery, and formula milk feeding with excessive body adiposity and type 2 diabetes mellitus in adult life [31-37]. Recently, Brazil exhibited relatively high rates of undernutrition; therefore, mothers of the NutriHS participants might have come from adverse conditions. Since it has been reported that the second generation is the one that suffers the biggest health impact (38,39), our study represents an opportunity for investigating risk factors for chronic diseases or even their early outcomes.

The atherosclerotic process and insulin resistance markers, both directly associated with the immune system, and subclinical inflammation, have been considered emergent risk factors [40-42]. The NutriHS is providing promising data which may contribute towards deepening the understanding of pathophysiological links of early life events, body composition, and inflammatory risk profile. The gut microbiome is meant to have a central role in mediating immune system alterations that predispose to a proinflammatory status and metabolic disorders [43-45]. Studies about gut microbiota have fomented the hypothesis of dysbiosis early in the cycle of life, triggering mechanisms that favor body fat accumulation and insulin resistance in adult life [46]. Associations of events during intrauterine life or at delivery with gut microbiota composition in childhood have been demonstrated [47-49]. Despite the recognition that the microbiota profile results from an interaction between environmental and genetic factors [50], to our knowledge, investigations about its association with early life events and emergent cardiometabolic risk factors in young adults have not yet been reported. The ability of eating habits to change the gut microbiota composition in the NutriHS could help to elucidate how dietary factors influence cardiometabolic risk.

Young adulthood, when preventive strategies could be more effective, is a unique moment for identifying early metabolic disturbances. The NutriHS cohort, by following undergraduates and graduates from Nutrition courses prospectively, has a great potential to enhance knowledge regarding the role of behavioral factors, especially related to diet and lifestyle, and underlying mechanisms of NCCDs. High-quality data are expected to be obtained from this stratus of the young Brazilian population. The age range of the NutriHS participants may confer less risk of recall bias, due to proximity to infancy and to more reliable information provided by parents and guardians present in their lives.

The importance of the initiative of building a website dedicated for this study, tailored to our local situation, should be emphasized. In addition to producing reliable data, the *e*-NutriHS proved to be easy to use, less costly and less time-consuming than paper-and-pencil collections. Furthermore, extra questionnaires can be included in the Web system, once the researchers decide to investigate additional concerns.

Since NutriHS implementation, users' satisfaction seems to be positive. Participants have found the system easy and practical to use. From 1,102 users (983 undergraduates and 119 graduates) the predominance of the female sex (94%) was expected and mean age of the sample is 25.1 (SD 7.4) years (unpublished data). Data on diet, physical activity, and others are ready for preliminary analysis. This scenario is in contrast with many epidemiological studies, which require a long period to get results. Another advantage of the *e*-NutriHS is related to the fact that participants receive a quick feedback about their health conditions.

Taking into consideration all these advantages, we believe that the NutriHS will contribute to elucidate the temporal sequence regarding some causal determinants and pathological effects, scarcely investigated in other cohort studies. Preliminary results (unpublished data) have fulfilled our expectation of establishing associations among events in the early stages of life, diet, gut microbiota, and circulating biomarkers, reinforcing the hypothesis of intrauterine programing interacting with environmental exposures to influence the risk profile for certain NCCDs, particularly for cardiometabolic diseases.

In summary, we conclude that the description of the NutriHS methodology, developed at FSP/USP, Brazil, is a relevant initiative to facilitate research in subsets

of populations. In addition, the *e*-NutriHS proved to be a powerful and user-friendly tool for producing reliable information, which could also be used in the field of the public health. The combination of a friendly tool with the innovative purposes of NutriHS offers a remarkable resource for testing hypotheses, clarifying mechanisms of nutrition-related diseases and further planning of preventive programs.

LIST OF ABBREVIATIONS

NutriHS – Nutritionists' Health Study

NCCDs – non-communicable chronic diseases

ELISA – enzyme-linked immunoenzymatic assay

FSP/USP – School of Public Health, University of São Paulo

IPAQ – International Physical Activity Questionnaire

DXA - dual-emission x-ray absorptiometry

FUNDING

The study has not received external funding.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

LGDF, ITS, BAP and SRGF designed the study; LGDF, ITS, BAP and SRGF participated in the elaboration of the protocol. LGDF programmed the web-based system. LGDF and ITS have been participating in the acquisition of data. LGDF, ITS, BAP and SRGF drafted the manuscript.

All authors have read and have approved the final manuscript.

ACKNOWLEDGEMENTS

This study is being supported by the School of Public Health, University of São Paulo.

The authors thanks Renata Vidonsky Luthold, Angélica M. M. Valente, Briana Duarte Teixeira, Bianca Sansão Waltrick, Priscila Koritar, Angélica Almeida Obara and Marle dos Santos Alvarenga for their contribution to this project.

REFERENCES

1. Rich-Edwards JW, Stampfer MJ, Manson JE, et al. Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. BMJ 1997; 315(7105):396-400.

2. Rich-Edwards JW, Colditz GA, Stampfer MJ, et al. Birthweight and the risk for type 2 diabetes mellitus in adult women. Ann Intern Med 1999; 130: 278–284.
3. Lucas A, Fewtrell MS, Cole TJ. Fetal origins of adult disease - the hypothesis revisited. BMJ. 1999; 319(7204):245-9.
4. Newsome CA, Shiell AW, Fall CH, et al. Is birth weight related to later glucose and insulin metabolism? – A systematic review. Diabet Med 2003; 20:339–348.
5. Barker DJP. Fetal and infant origins of adult disease. London: BMJ Publishing; 1992.
6. Paneth N, Susser M. Early origin of coronary heart disease (the ‘‘Barker hypothesis’’). BMJ 1995; 310: 411–412.
7. Kramer MS, Joseph KS. Enigma of fetal/infant origins hypothesis. Lancet 1996; 348: 1254–1255.
8. Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. Science 2004; 305(5691):1733-6.
9. Godfrey K. The developmental origins hypothesis: Epidemiology. In: Gluckman PD, Hanson MA, editors. Developmental Origins of Health and Disease. Cambridge: Cambridge University Press; 2006. p. 6–32.
10. Bäckhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci USA 2004; 101(44):15718-23.
11. Ley RE, Turnbaugh PJ, Klein S, et al. Microbial ecology: human gut microbes associated with obesity. Nature 2006; 444(7122):1022-3.
12. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature 2011; 473: 174-180.
13. Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011; 334(6052):105-8.
14. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010; 172(11):1292-8.
15. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. Nat Rev Cancer 2005; 5(5):388-96.
16. Wilson KM, Giovannucci E, Stampfer MJ, et al. Dietary acrylamide and risk of prostate cancer. Int J Cancer 2012; 131(2):479-87.

17. Satija A, Yu E, Willett WC, et al. Understanding nutritional epidemiology and its role in policy. Adv Nutr 2015; 6(1):5-18.
18. Malik VS, Willett WC, Hu FB. Global obesity: trends, risk factors and policy implications. Nat Rev Endocrinol 2013; 9(1):13-27.
19. Callegaro M, Manfreda KL, Vehovar V. Web Survey Methodology. London: Sage; 2015.
20. Touvier M, Méjean C, Kesse-Guyot E, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. Eur J Epidemiol 2010; 25(5):287-96.
21. Couper MP. Web surveys – A review of issues and approaches. Public Opin Q 2000; 64(4):464-94.
22. Schmidt WC. World-Wide Web survey research: benefits, potential problems, and solutions. Behav Res Meth Instrum Comput 1997; 29. p.274–9.
23. Turner C, Bain C, Schluter PJ, et al. Cohort profile: the Nurses and Midwives e-Cohort Study—a novel electronic longitudinal study. Int J Epidemiol 2009; 38(1):53-60.
24. Mikkelsen EM, Hatch EE, Wise LA, et al. Cohort profile: the Danish Web-based Pregnancy Planning Study-'Snart-Gravid'. Int J Epidemiol 2009; 38(4): 938–943.
25. Hercberg S, Castetbon K, Czernichow S, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. BMC Public Health 2010; 10:242.
26. Folchetti LD, Silva IT, Almeida-Pititto B, et al. Development of Web-based System for a Cohort Study - The Brazilian e-NutriHS. In: Proceedings of the International Conference on Health Informatics; 2015. p. 574-579.
27. Craig CL, Marshall AL, Sjöström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003; 35(8):1381-95.
28. Fisberg RM, Colucci ACA, Morimoto JM, et al. Food frequency questionnaire for adults from a population- based study. Rev Saúde Pública 2008; 42(3):550-554.

29. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985; 28:412–419.
30. Silveira PP, Portella AK, Goldani MZ, et al. Developmental origins of health and disease (DOHaD). J Pediatr (Rio J) 2007; 83(6):494-504.
31. Curhan GC, Willett WC, Rimm EB, et al. Birth weight and adult hypertension, diabetes mellitus, and obesity in US men. Circulation 1996; 94:3246–3250.
32. Barghava SK, Sachdev HS, Fall CH, et al. Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood. N Engl J Med 2004; 350:865-75.
33. Zhao M, Shu XO, Jin F, et al. Birthweight, childhood growth and hypertension in adulthood. Int J Epidemiol 2002; 31:1043-51.
34. Li HT, Zhou YB, Liu JM. The impact of cesarean section on offspring overweight and obesity: A systematic review and meta-analysis. Int J Obes (Lond) 2013; 37(7):893-9.
35. Darmasseelane K, Hyde MJ, Santhakumaran S, et al. Mode of Delivery and offspring Body Mass Index, Overweight and Obesity in Adult Life: A Systematic Review and Meta-Analysis. PLoS One 2014; 9(2):e87896.
36. Prior E, Santhakumaran S, Gale C, et al. Breastfeeding after cesarean delivery: A systematic review and meta-analysis of world literature. Am J Clin Nutr 2012; 95(5):1113-35.
37. Victora CG, Aquino EML, Leal M do C, et al. Maternal and child health in Brazil: Progress and challenges. Lancet 2011; 377(9780): 1863–1876.
38. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. N Engl J Med 1976; 295(7):349-53.
39. Neel JV. Diabetes Mellitus: A "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 1962; 14(4): 353–362.
40. Libby P. Inflammation and cardiovascular disease mechanisms. Am J Clin Nutr 2006; 83(suppl):456S– 60S.
41. Gilstrap LG, Wang TJ. Biomarkers and cardiovascular risk assessment for primary prevention: an update. Clin Chem 2012; 58(1):72-82.
42. Magnussen CG, Niinikoski H, Juonala M, et al. When and how to start prevention of atherosclerosis? Lessons from the Cardiovascular Risk in the

Young Finns Study and the Special Turku Coronary Risk Factor Intervention Project. Pediatr Nephrol 2012; 27(9):1441-52.

43. Wallace TC, Guarner F, Madsen K, et al. Human gut microbiota and its relationship to health and disease. Nutr Rev 2011; 69(7):392-403.
44. Bäckhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci USA 2004; 101(44):15718-23.
45. Moraes ACF, Silva IT, Almeida-Pititto B, et al. Intestinal microbiota and cardiometabolic risk: mechanisms and diet modulation. Arq Bras Endocrinol Metab 2014; 58(4):317-27.
46. Maynard CL, Elson CO, Hatton RD, et al. Reciprocal interactions of the intestinal microbiota and immune system. Nature 2012; 489:231-41.
47. Cabrera-Rubio R, Collado MC, Laitinen K, et al. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 2012; 96:544-51.
48. Azad MB, Konya T, Maughan H, et al. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ 2013; 185:385-94.
49. Jakobsson HE, Abrahamsson TR, Jenmalm MC, et al. Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by Caesarean section. Gut 2014; 63(4):559–566.
50. Ussar S, Griffin NW, Bezy O, et al. Interactions between gut microbiota, host genetics and diet modulate the predisposition to obesity and metabolic syndrome. Cell Metab 2015; 22(3):516-30.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4 and 5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6 and 7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7 to 10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7 to 10
Bias	9	Describe any efforts to address potential sources of bias	10 to 12
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	9 and 10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	9 and 10
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, explain how loss to follow-up was addressed	N/A
		(*e*) Describe any sensitivity analyses	9 and 10
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	N/A
		(b) Give reasons for non-participation at each stage	N/A

		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	12
		(b) Indicate number of participants with missing data for each variable of interest	N/A
		(c) Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	Report numbers of outcome events or summary measures over time	N/A
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A
		(*b*) Report category boundaries when continuous variables were categorized	N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	10 to 12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12
Generalisability	21	Discuss the generalisability (external validity) of the study results	12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for exposed and unexposed groups.

BMJ Open

The Nutritionists' Health Study cohort: a web-based approach of life events, habits, and health outcomes

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012081.R1
Article Type:	Protocol
Date Submitted by the Author:	29-Jun-2016
Complete List of Authors:	Folchetti, Luciana; University of Sao Paulo, School of Public Health Silva, Isis; University of Sao Paulo, School of Public Health Almeida-Pititto, Bianca; Universidade Federal de Sao Paulo Escola Paulista de Medicina, Department of Preventive Medicine Ferreira, Sandra Roberta G.; University of Sao Paulo, Department of Epidemiology, School of Public Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Public health, Research methods, Nutrition and metabolism
Keywords:	Nutritionists' Health Study, web-based system, early-life events, gut microbiota, cardiometabolic risk factors

The Nutritionists' Health Study cohort: a web-based approach of life events, habits, and health outcomes

Luciana Gavilan Dias Folchetti[1], Isis Tande da Silva[1], Bianca de Almeida-Pititto[2], Sandra Roberta. G. Ferreira[3]

1 Department of Nutrition, School of Public Health, University of Sao Paulo, São Paulo, Brazil

2 Department of Preventive Medicine, Federal University of Sao Paulo, São Paulo, Brazil

3 Department of Epidemiology, School of Public Health, University of Sao Paulo, São Paulo, Brazil

Correspondence

Prof. Sandra Roberta G. Ferreira

Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo

Av. Dr. Arnaldo, 715 – CEP 01246-904 – São Paulo, SP, Brasil

Fone 55 11 3061-7870 E-mail: sandrafv@usp.br

KEYWORDS

Nutritionists' Health Study, web-based system, early-life events, gut microbiota, cardiometabolic risk factors

Word Count: 3,568

ABSTRACT

INTRODUCTION Non-communicable chronic diseases (NCCDs) represent a burden for public health. Alongside, the established cardiometabolic risk factors – high blood pressure and disorders of glucose and lipid metabolism – living habits and nutritional status at different stages of life are seen as contributors for this scenario. Gut microbiota composition and subclinical inflammation have been pointed out as underlying mechanisms of NCCDs. Studies involving health professionals have brought relevant contributions to the knowledge about risk factors. Technological advances facilitate data collection and analysis for big samples. Web-based survey addressed to collect data from a cohort study, able to identify NCCDs risk factors, is highly desirable. The objective of the Brazilian Nutritionists' Health Study – NutriHS is to gather online information on early-life events, daily habits, emergent cardiometabolic risk factors and health outcomes of a specific subset of the Brazilian population. **METHODS AND ANALYSIS** The NutriHS, developed at the School of Public Health – University of Sao Paulo, Brazil, is a research initiative that enrolls undergraduates of nutrition courses from Brazilian universities and graduated volunteers. A web-based self-administered system was designed to collect health-related data. After fulfilling online questionnaires (socioeconomic, early-life events and lifestyle data), participants are invited to a clinical visit for physical examination and lab procedures (blood sampling, feces collection and body composition). At a 3-year interval, they will be invited to repeat similar procedures. **ETHICS AND DISSEMINATION** The NutriHS research protocol was approved by the Institutional Ethics Committee and is providing promising data which contribute to the understanding of pathophysiological links between early life events, body composition, gut microbiota and inflammatory and metabolic risk profile. The combination of a friendly tool with the innovative purposes of NutriHS offers a remarkable resource for testing hypotheses about mechanisms of nutrition-related diseases and further planning of preventive programs in public health.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- A web-based prospective study for collection of self-reported multidisciplinary data

- Low cost for development and potential of amplification of the web-based system
- Availability of self-reported and measured data which allows validation study
- Innovative data to be associated with early outcomes
- Characteristics of the population sample which has skills to accurate answers to self-reported technical questionnaires and potential high compliance in the long-term study
- Language-related limitation offered in the web-based system (only in Portuguese)
- Non-representative sample of general population

INTRODUCTION

Numerous factors from intrauterine life and birth determine adult health and well-being. Monitoring of early-life events and daily habits has provided tools to protect individual health from major damage and other threats. Health outcomes depend on a combination of genetic and environmental factors.

Delivery conditions, birth weight, and early events in infancy have been involved in the genesis of future diseases [1-4]. Such a relationship, initially raised by Barker [5], has been further confirmed [6-9], although underlying mechanisms are not yet completely understood. More recently, the types of delivery and child nutrition have been shown to influence the gut microbiota composition which, in turn, modulates immune system and the risk of diseases [10-13].

Epidemiological and clinical studies have contributed towards identifying modifiable risk factors and deepening the understanding of the pathophysiology of non-communicable chronic diseases (NCCDs), such as obesity, type 2 diabetes mellitus, and cardiovascular disease. These lifestyle-related morbidities represent a heavy burden for public health systems due to complications, increased mortality, and high costs [14-16].

Knowledge of risk factors for NCCDs, collected prospectively, is essential for enhancing evidence and planning interventions. Epidemiological studies involving health professionals have brought significant contributions to the knowledge about the role of behavioral factors in health outcomes [17,18]. This kind of study often requires large population samples, followed for long periods, which generate a significant amount of data, high costs, and the need for trained professionals [15-18]. More commonly, health surveys have been based on paper-and-pencil questionnaires.

Printed questionnaires-based surveys are complex and time-consuming, and their quality depends on accuracy in data collecting, processing and uploading to create a consistent database before analysis. Despite extending study duration and representing a financial burden, these steps are crucial for adequate data analysis to achieve reliable conclusions [19,20].

Technological advances can minimize errors and duplicated information, reduce costs, as well as provide the almost immediate construction of a database using

bioinformatic tools [20,21]. Multiple and diverse software for data analysis tailored to specific questions are available, but very few are designed specifically to investigate the role of early-life events and daily habits.

The Internet is a promising tool for overcoming some of the logistic and financial feasibility constraints in research, and is gaining considerable popularity [21]. In contrast to other fields – such as in psychological and marketing research – in epidemiology, less than 1% of the reported studies are internet-assisted [14]. Successful examples of large-scale studies in which enrolment and follow-up occur over the Internet are the Millennium Cohort Study [22], the Nurses and Midwives e-Cohort Study [23], and the Danish Web-based Pregnancy Planning Study [24]. In nutritional epidemiology, there is the NutriNet-Santé Study, a web-based prospective study launched in France in 2009 [25].

In middle-income countries, similar initiatives using the Internet as a tool for data collection are scarce, and when considering the inclusion of health professionals they are inexistent. The development of web-based surveys able to identify environmental risk factors in developing countries with wide geographic area and limited resources would be highly desirable. As far as we know, the use of self-administered online questionnaires for this kind of population has not yet been reported.

The purpose of the current study was to describe the methodology of the Nutritionists' Health Study (NutriHS), a prospective Web-based study developed at the School of Public Health, University of São Paulo (FSP/USP), Brazil, to gather information on early-life events, daily habits, emergent cardiometabolic risk factors, and health outcomes of a specific subset of the Brazilian population – undergraduates and graduates in Nutrition. The outcome of major interest is excessive body weight and obesity-related morbidities.

METHODS AND ANALYSIS

Implementation of the NutriHS study included the following steps: 1) development of the web-based system; 2) collection of data related to early-life events, lifestyle habits, anthropometry, and body composition; 3) assessment of the cardiovascular profile based on traditional and emerging risk factors; 4) analysis of the association of

early-life events and lifestyle habits with the gut microbiota composition, and cardiovascular risk factors.

- **The Web-based system**

With the support of the FSP/USP Informatics Center, a dedicated Web-based online self-administered system (http://www.fsp.usp.br/nutrihs/index.html) was developed, named *e*-NutriHS, to collect health-related data based on previously validated printed questionnaires. Details on its development have been described elsewhere [26]. In brief, the Web-based system was built using free programming languages: HyperText Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript used to create a working dynamic website with friendly user's interface; the Hypertext P reprocessor (PHP), server-side scripting language, used to transfer of data to the NutriHS database; and Structured Query Language (SQL) is the database programming language used at the database management system, MySQL.

The colors for NutriHS website – green and orange – refers to health, nutrition and youth. The logotype and slogan call attention to the need of increase knowledge on the role of nutrition for health.

The development also included skip patterns to hide questions (depending on the person's answer to a prior question), randomly organized questions, and validation checks providing additional information to simplify and help responding, pointing out questions that needed attention. The *e*-NutriHS data can be transferred to Excel format, allowing further analyses. Since the database is located at the FSP/USP server, its security was guaranteed.

Prior to releasing the *e*-NutriHS to participants, a prototype was created (26). Twenty volunteers, graduates at the FSP/USP who have been working in the health area, aged 25 to 60 years, tested the system and checked its ease-of-use, data entry, and website design. Based on their opinions, a user-friendly version of *e*-NutriHS was released.

The website provides details on the study purposes, the team involved, and contact information, and also represents an opportunity for learning with recent research reports from international literature. In order to increase awareness of the NutriHS, a social network page on the Internet was created

(https://www.facebook.com/nutrihs) with a direct link to the *e*-NutriHS homepage. This website plays an important role in stimulating dialogue among researchers and participants.

- **Study design, recruitment, and inclusion**

Given the growing recognition of the importance of the life course approach for understanding the determinants of NCCDs, the NutriHS was designed to be a cohort study in which retrospective and prospective data are being collected. The study was designed to include all undergraduates and graduates from Brazilian Nutrition Colleges, aged ≥18 years. Undergraduates from Nutrition courses and nutritionist volunteers were selected based on their potential skill to navigate the Internet and on the expected high quality of data provided. Exclusion criteria are pregnancy and no access to the Internet.

Volunteer recruitment was done using multiple strategies. Invitations were made to the university staffs, lectures for undergraduates or graduates were offered, news was published on the FSP/USP website and on the NutriHS social network page, and banners were placed in strategic places.

After signing the electronic informed consent, users are required to fill out personal data and to create a unique login and password. Then, their access to questionnaires is released. Once questionnaires completed, participants are instructed to schedule, on the *e*-NutriHS, a face-to-face visit at the Health Center for clinical examination and biological sampling. Recruitment started in March 2014 and this is planned to be completed by July 2016. After a 3-year interval, they will be invited to complete a similar set of questionnaires.

- **Questionnaires and variables of interests**

The baseline set of questionnaires deals with socioeconomic, early-life events, and lifestyle data, including physical activity and eating habits. The order for filling out these questionnaires is chosen by the participant. Excessive body weight is a major outcome, which is being assessed using anthropometric measurements and densitometry. Other variables will be used as exposures, predictors, potential confounders, or effect modifiers.

1. **Demographic, socioeconomic lifestyle, and clinical questionnaires**

The structured questionnaires obtained data regarding skin color, educational level, marital status, person(s) they live with, parents or guardians, family income, number of people in the house, work, work hours, use of medications, medical history, familial medical history, a restrictive diet, alcohol consumption, tobacco consumption, height, and weight.

2. Early-life events questionnaire

The contents of this questionnaire allow one to obtain retrospective information about the mothers of the participants, their pregnancies, and data related to birth and childhood, as well as maternal health and educational level during the pregnancy period. Before filling the questionnaire, instructions are provided, recommending participants to ask their parents (mainly mothers) or to look at their birth registry to get specific information regarding gestational period and early-life events.

Data collected are divided into two parts regarding mothers and participants. Mother's age at the participant's birth, educational level, occupation, height, weight before pregnancy, use of medications, pregnancy conditions, such as weight gain and health problems, are reported. Also, participants are asked to answer as to prematurity, twins, birth weight, breastfeeding, time of breastfeeding, time of introduction of formula, medications used, weight gain during childhood and adolescence, fruit and vegetable consumption, and physical activity during childhood.

As this questionnaire requires retrospective

3. Physical activity questionnaire

The short version of the International Physical Activity Questionnaire (IPAQ), validated for the Brazilian population [27], is being employed. Intensity of physical activity (walking, moderate and intense), frequency (per week), and duration are obtained. These data enable estimates of total physical activity as quantitative or categorical variables (low, moderate, and high). A sedentary pattern has been assessed by time spent sitting.

4. Dietary assessment

Diet is assessed using a validated 102-item food frequency questionnaire [28] regarding eating habits over the last year. Data inserted into the *e*-NutriHS are immediately transformed into the equivalent food amounts for one day. The system is

able to provide reports of macro- and micronutrients, based on the USDA National Nutrient Database for Standard Reference.

5. Clinical assessment

Completion of questionnaires is required in order to access the link for scheduling the face-to-face visit to the FSP/USP Public Health Center. A couple of days before the visit, an e-mail with instructions for biological sampling and body composition determination is sent to each participant. This visit represents an opportunity to double-check previously collected selected data.

- ***Anthropometry and blood pressure***

Weight is obtained on a digital scale accurate to 100 g and height was measured using a fixed stadiometer to the nearest 0.5 cm. They are used to determine the body mass index, which was calculated as weight in kilograms divided the square of the height in meters (kg/m^2). Waist circumference is taken in centimeters at the midpoint between the lowest rib and the iliac crest. Body composition is being determined by dual-emission x-ray absorptiometry (DXA Lunar GE) that allows accurate estimates of body compartments, including visceral fat. Sitting blood pressure levels are measured using an automatic device, expressed in mm Hg.

- ***Biochemical data***

Blood samples are collected for plasma glucose and lipid profile, determined using the hexokinase method and enzymatic colorimetric assay, respectively. Aliquots are frozen at -80^{o}C for further determinations of apolipoproteins, hormones, and inflammatory markers. Insulin is determined by enzyme-linked immunoenzymatic assay (ELISA) and high-sensitivity C-reactive protein by immunochemistry. ELISA kits are also used for the measurements of cytokine concentrations. Insulin resistance will be estimated by the HOMA-IR index [29]. Extra aliquots are stored for future analyses of interest to the NutriHS.

- ***Metagenomic analysis of gut microbiota***

Participants are instructed to collect a fecal sample and refrigerate it before the visit. Aliquots are frozen until the metagenomic analysis. The pyrosequencing of 16S subunit ribosomal amplicons will be performed using next-generation sequencing technology. After obtaining the sequences, they are subjected to phylogenetic

analysis to characterize the fecal microbiota. Bioinformatic procedures, comparing the sequences extracted from the fecal samples with standardized databases, allow determination of the abundance of bacteria belonging to phyla and genera.

- **Data and statistical analysis plan**

The NutriHS planned to collect prospectively data on emergent lifestyle-related risk factors of a specific subset of the healthy Brazilian population. For its cross-sectional phase, data analyses included the description of the population sample to be followed. Additionally, a validation sub-study of the web-collected self-reported data is being conducted. A number of statistical analyses has been planned to test associations of knowledge in nutrition and changes in lifestyle, as well as in body composition and biochemical profile.

Analyses are being performed using IBM SPSS Statistics V. 20.0. Collected measures and derived variables will be summarized using means and standard deviation (SD), medians and interquartile range (IQR), and/or frequencies. Student *t* test or one-way analysis of variance when comparing more than two groups will be used for continuous variables with normal distribution. Pearson correlation coefficient will be employed to test correlations. Correspondent non-parametric tests will use for non-normal distributed variables. For validation studies, agreement has been analyzed by Kappa statistics and Bland-Altman plot method. For multivariate analyses, generalized linear regression models will be used or logistic regression for dichotomous variables. For all tests, a p-value < 0.05 will be considered significant.

ETHICS AND DISSEMINATION

Ethics

The NutriHS study protocol and study documentation, including all questionnaires and the informed consent, was approved by the Institutional Ethics Committee, Committee for Ethics in Research, School of Public Health, University of Sao Paulo. The substantiated opinions were registered by the numbers: Of. COEP 991.542/15 and Of. COEP 257.513/13.

When accessing the *e*-NutriHS, participants must sign an electronic informed consent before the registration process. Individual privacy is respected and data will

be stored in *e*-NutriHS database following the regulations applied to security. All publications will respect confidentiality.

Discussion and Dissemination

Analyses involving the issue cardiovascular risk factors & public health, anchored in population-based studies or specific population subsets have been conducted in developed countries. Such studies are of great relevance for proposals and formulation of public policies tailored to local situations. Cohorts of health professionals, like the Nurses' Health Study I and II (http://www.channing.harvard.edu/nhs/) and the Health Professionals Follow-up Study (https://www.hsph.harvard.edu/hpfs/), which included thousands of participants in North America, brought remarkable scientific contributions for identifying and intervening in major risk factors for NCCDs [17-18]. One strong point of studies involving professionals is the high quality of the specific data collected.

The current Brazilian epidemiological scenario of increasing longevity and nutritional transition highlights the importance of obesity-related chronic diseases for mortality. Along with the traditional cardiometabolic risk factors – hypertension and impaired metabolism of glucose and lipids – lifestyle and nutritional status in different stages of life, including the intrauterine period, have been pointed out as relevant determinants of this situation [3,8,30]. Scientific evidence has associated low birthweight, cesarean delivery, and formula milk feeding with excessive body adiposity and type 2 diabetes mellitus in adult life [31-37]. Recently, Brazil exhibited relatively high rates of undernutrition; therefore, mothers of the NutriHS participants might have come from adverse conditions. Since it has been reported that the second generation is the one that suffers the biggest health impact (38,39), our study represents an opportunity for investigating risk factors for chronic diseases or even their early outcomes.

The atherosclerotic process and insulin resistance markers, both directly associated with the immune system, and subclinical inflammation, have been considered emergent risk factors [40-42]. The NutriHS is providing promising data which may contribute towards deepening the understanding of pathophysiological links of early life events, body composition, and inflammatory risk profile. The gut microbiome is meant to have a central role in mediating immune system alterations

that predispose to a proinflammatory status and metabolic disorders [43-45]. Studies about gut microbiota have fomented the hypothesis of dysbiosis early in the cycle of life, triggering mechanisms that favor body fat accumulation and insulin resistance in adult life [46]. Associations of events during intrauterine life or at delivery with gut microbiota composition in childhood have been demonstrated [47-49]. Despite the recognition that the microbiota profile results from an interaction between environmental and genetic factors [50], to our knowledge, investigations about its association with early life events and emergent cardiometabolic risk factors in young adults have not yet been reported. The ability of eating habits to change the gut microbiota composition in the NutriHS could help to elucidate how dietary factors influence cardiometabolic risk.

Young adulthood, when preventive strategies could be more effective, is a unique moment for identifying early metabolic disturbances. The NutriHS cohort, by following undergraduates and graduates from Nutrition Colleges prospectively, has a great potential to enhance knowledge regarding the role of behavioral factors, especially related to diet and lifestyle, and underlying mechanisms of NCCDs. High-quality data are expected to be obtained from this stratus of the young Brazilian population. The age range of the NutriHS participants may confer less risk of recall bias, due to proximity to infancy and to more reliable information provided by parents and guardians present in their lives.

Innovative technologies in nutritional epidemiology, such the e-NutriHS, that deal with the entire process, from data gathering to analysis, is promising to overcome some of the logistical and financial feasibility constraints that can affect the conduction of such large-scale surveys. Furthermore, the broad reach of geographic areas can enhance participation rates. However, concerns about security, duplicate records, and sampling issues (selection bias) of online surveys should be raised [14,21]. Some of these concerns could be overcome, for example, through cryptography and security management of the database, screening scripts to check implausible answers and to detect and not allow saving duplicate registries [14,19,22,25]. Recent studies have reinforced that Web-based data collection represents a valid and suitable method [51-53].

Therefore, the initiative of building a website dedicated for this study, tailored to our local situation, should be emphasized. In addition to producing reliable data, the

e-NutriHS proved to be easy to use, less costly and less time-consuming than paper-and-pencil collections. Furthermore, extra questionnaires can be included in the Web system, once the researchers decide to investigate additional concerns.

Since NutriHS implementation, users' satisfaction seems to be positive. Participants have found the system easy and practical to use. From 1,102 users (983 undergraduates and 119 graduates) the predominance of the female sex (94%) was expected and mean age of the sample is 25.1 (SD 7.4) years (unpublished data). Data on diet, physical activity, and others are ready for preliminary analysis. This scenario is in contrast with many epidemiological studies, which require a long period to get results. Another advantage of the *e*-NutriHS is related to the fact that participants receive a quick feedback about their health conditions.

Taking into consideration all these advantages, we believe that the NutriHS will contribute to elucidate the temporal sequence regarding some causal determinants and pathological effects, scarcely investigated in other cohort studies. Preliminary results (unpublished data) have fulfilled our expectation of establishing associations among events in the early stages of life, diet, gut microbiota, and circulating biomarkers, reinforcing the hypothesis of intrauterine programing interacting with environmental exposures to influence the risk profile for certain NCCDs, particularly for cardiometabolic diseases.

In summary, we conclude that the description of the NutriHS methodology, developed at FSP/USP, Brazil, is a relevant initiative to facilitate research in subsets of populations. In addition, the *e*-NutriHS proved to be a powerful and user-friendly tool for producing reliable information, which could also be used in the field of the public health. The combination of a friendly tool with the innovative purposes of NutriHS offers a remarkable resource for testing hypotheses, clarifying mechanisms of nutrition-related diseases and further planning of preventive programs.

LIST OF ABBREVIATIONS

NutriHS – Nutritionists' Health Study

NCCDs – non-communicable chronic diseases

ELISA – enzyme-linked immunoenzymatic assay

FSP/USP – School of Public Health, University of São Paulo

IPAQ – International Physical Activity Questionnaire

DXA - dual-emission x-ray absorptiometry

FUNDING

The study has not received external funding.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

LGDF, ITS, BAP and SRGF designed the study; LGDF, ITS, BAP and SRGF participated in the elaboration of the protocol. LGDF programmed the web-based system. LGDF and ITS have been participating in the acquisition of data. LGDF, ITS, BAP and SRGF drafted the manuscript.

All authors have read and have approved the final manuscript.

ACKNOWLEDGEMENTS

This study is being supported by the School of Public Health, University of São Paulo.

The authors thanks Renata Vidonsky Luthold, Angélica M. M. Valente, Briana Duarte Teixeira, Bianca Sansão Waltrick, Priscila Koritar, Angélica Almeida Obara and Marle dos Santos Alvarenga for their contribution to this project.

REFERENCES

1. Rich-Edwards JW, Stampfer MJ, Manson JE, et al. Birth weight and risk of cardiovascular disease in a cohort of women followed up since 1976. BMJ 1997; 315(7105):396-400.

2. Rich-Edwards JW, Colditz GA, Stampfer MJ, et al. Birthweight and the risk for type 2 diabetes mellitus in adult women. Ann Intern Med 1999; 130: 278–284.
3. Lucas A, Fewtrell MS, Cole TJ. Fetal origins of adult disease - the hypothesis revisited. BMJ. 1999; 319(7204):245-9.
4. Newsome CA, Shiell AW, Fall CH, et al. Is birth weight related to later glucose and insulin metabolism? – A systematic review. Diabet Med 2003; 20:339–348.
5. Barker DJP. Fetal and infant origins of adult disease. London: BMJ Publishing; 1992.
6. Paneth N, Susser M. Early origin of coronary heart disease (the ‘‘Barker hypothesis’’). BMJ 1995; 310: 411–412.
7. Kramer MS, Joseph KS. Enigma of fetal/infant origins hypothesis. Lancet 1996; 348: 1254–1255.
8. Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. Science 2004; 305(5691):1733-6.
9. Godfrey K. The developmental origins hypothesis: Epidemiology. In: Gluckman PD, Hanson MA, editors. Developmental Origins of Health and Disease. Cambridge: Cambridge University Press; 2006. p. 6–32.
10. Bäckhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci USA 2004; 101(44):15718-23.
11. Ley RE, Turnbaugh PJ, Klein S, et al. Microbial ecology: human gut microbes associated with obesity. Nature 2006; 444(7122):1022-3.
12. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature 2011; 473: 174-180.
13. Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science 2011; 334(6052):105-8.
14. van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? Am J Epidemiol 2010; 172(11):1292-8.
15. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. Nat Rev Cancer 2005; 5(5):388-96.
16. Wilson KM, Giovannucci E, Stampfer MJ, et al. Dietary acrylamide and risk of prostate cancer. Int J Cancer 2012; 131(2):479-87.

17. Satija A, Yu E, Willett WC, et al. Understanding nutritional epidemiology and its role in policy. Adv Nutr 2015; 6(1):5-18.
18. Malik VS, Willett WC, Hu FB. Global obesity: trends, risk factors and policy implications. Nat Rev Endocrinol 2013; 9(1):13-27.
19. Callegaro M, Manfreda KL, Vehovar V. Web Survey Methodology. London: Sage; 2015.
20. Touvier M, Méjean C, Kesse-Guyot E, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. Eur J Epidemiol 2010; 25(5):287-96.
21. Couper MP. Web surveys – A review of issues and approaches. Public Opin Q 2000; 64(4):464-94.
22. Schmidt WC. World-Wide Web survey research: benefits, potential problems, and solutions. Behav Res Meth Instrum Comput 1997; 29. p.274–9.
23. Turner C, Bain C, Schluter PJ, et al. Cohort profile: the Nurses and Midwives e-Cohort Study—a novel electronic longitudinal study. Int J Epidemiol 2009; 38(1):53-60.
24. Mikkelsen EM, Hatch EE, Wise LA, et al. Cohort profile: the Danish Web-based Pregnancy Planning Study-'Snart-Gravid'. Int J Epidemiol 2009; 38(4): 938–943.
25. Hercberg S, Castetbon K, Czernichow S, et al. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. BMC Public Health 2010; 10:242.
26. Folchetti LD, Silva IT, Almeida-Pititto B, et al. Development of Web-based System for a Cohort Study - The Brazilian e-NutriHS. In: Proceedings of the International Conference on Health Informatics; 2015. p. 574-579.
27. Craig CL, Marshall AL, Sjöström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003; 35(8):1381-95.
28. Fisberg RM, Colucci ACA, Morimoto JM, et al. Food frequency questionnaire for adults from a population- based study. Rev Saúde Pública 2008; 42(3):550-554.

29. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985; 28:412–419.
30. Silveira PP, Portella AK, Goldani MZ, et al. Developmental origins of health and disease (DOHaD). J Pediatr (Rio J) 2007; 83(6):494-504.
31. Curhan GC, Willett WC, Rimm EB, et al. Birth weight and adult hypertension, diabetes mellitus, and obesity in US men. Circulation 1996; 94:3246–3250.
32. Barghava SK, Sachdev HS, Fall CH, et al. Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood. N Engl J Med 2004; 350:865-75.
33. Zhao M, Shu XO, Jin F, et al. Birthweight, childhood growth and hypertension in adulthood. Int J Epidemiol 2002; 31:1043-51.
34. Li HT, Zhou YB, Liu JM. The impact of cesarean section on offspring overweight and obesity: A systematic review and meta-analysis. Int J Obes (Lond) 2013; 37(7):893-9.
35. Darmasseelane K, Hyde MJ, Santhakumaran S, et al. Mode of Delivery and offspring Body Mass Index, Overweight and Obesity in Adult Life: A Systematic Review and Meta-Analysis. PLoS One 2014; 9(2):e87896.
36. Prior E, Santhakumaran S, Gale C, et al. Breastfeeding after cesarean delivery: A systematic review and meta-analysis of world literature. Am J Clin Nutr 2012; 95(5):1113-35.
37. Victora CG, Aquino EML, Leal M do C, et al. Maternal and child health in Brazil: Progress and challenges. Lancet 2011; 377(9780): 1863–1876.
38. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. N Engl J Med 1976; 295(7):349-53.
39. Neel JV. Diabetes Mellitus: A "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 1962; 14(4): 353–362.
40. Libby P. Inflammation and cardiovascular disease mechanisms. Am J Clin Nutr 2006; 83(suppl):456S– 60S.
41. Gilstrap LG, Wang TJ. Biomarkers and cardiovascular risk assessment for primary prevention: an update. Clin Chem 2012; 58(1):72-82.
42. Magnussen CG, Niinikoski H, Juonala M, et al. When and how to start prevention of atherosclerosis? Lessons from the Cardiovascular Risk in the

Young Finns Study and the Special Turku Coronary Risk Factor Intervention Project. Pediatr Nephrol 2012; 27(9):1441-52.

43. Wallace TC, Guarner F, Madsen K, et al. Human gut microbiota and its relationship to health and disease. Nutr Rev 2011; 69(7):392-403.
44. Bäckhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci USA 2004; 101(44):15718-23.
45. Moraes ACF, Silva IT, Almeida-Pititto B, et al. Intestinal microbiota and cardiometabolic risk: mechanisms and diet modulation. Arq Bras Endocrinol Metab 2014; 58(4):317-27.
46. Maynard CL, Elson CO, Hatton RD, et al. Reciprocal interactions of the intestinal microbiota and immune system. Nature 2012; 489:231-41.
47. Cabrera-Rubio R, Collado MC, Laitinen K, et al. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 2012; 96:544-51.
48. Azad MB, Konya T, Maughan H, et al. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. CMAJ 2013; 185:385-94.
49. Jakobsson HE, Abrahamsson TR, Jenmalm MC, et al. Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by Caesarean section. Gut 2014; 63(4):559–566.
50. Ussar S, Griffin NW, Bezy O, et al. Interactions between gut microbiota, host genetics and diet modulate the predisposition to obesity and metabolic syndrome. Cell Metab 2015; 22(3):516-30.
51. Touvier M, Méjean C, Kesse-Guyot E, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. Europ J Epidemiol 2010; 25(5), 287-96.
52. Pursey K, Burrows TL, Stanwell P, et al. How accurate is web-based self-reported height, weight, and body mass index in young adults? J Med Internet Res. 2014; 16(1):e4.

53. Lassale C, Péneau S, Touvier M, et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Santé Study. J Med Internet Res. 2013; 15(8):e152.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1 and 2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4 and 5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6 and 7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7 to 10
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7 to 10
Bias	9	Describe any efforts to address potential sources of bias	11 to 13
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	10
		(*b*) Describe any methods used to examine subgroups and interactions	N/A
		(*c*) Explain how missing data were addressed	N/A
		(*d*) If applicable, explain how loss to follow-up was addressed	N/A
		(*e*) Describe any sensitivity analyses	10
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	N/A
		(b) Give reasons for non-participation at each stage	N/A

		(c) Consider use of a flow diagram	N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	13
		(b) Indicate number of participants with missing data for each variable of interest	N/A
		(c) Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	Report numbers of outcome events or summary measures over time	N/A
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A
		(*b*) Report category boundaries when continuous variables were categorized	N/A
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	11 to 13
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12 and 13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12 and 13
Generalisability	21	Discuss the generalisability (external validity) of the study results	12 and 13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

*Give information separately for exposed and unexposed groups.