Cohort Profile: The Dutch Perined-Lifelines birth cohort

Nastaran Salavati[1]*, Marian K. Bakker[1,2], Eline M. van der Beek[3,4], JanJaap H. M. Erwich[1]

1 Department of Obstetrics and Gynecology, University Medical Centre of Groningen, University of Groningen, Groningen, The Netherlands, **2** Department of Genetics, EUROCAT Registration Northern Netherlands, University Medical Centre of Groningen, University of Groningen, Groningen, The Netherlands, **3** Department of Pediatrics, University Medical Centre of Groningen, University of Groningen, Groningen, The Netherlands, **4** Danone Nutricia Research, Utrecht, The Netherlands

* n.salavati@umcg.nl

Abstract

Background

Maternal nutrition status (e.g. dietary/nutrient intake) during pregnancy has been associated with pregnancy outcomes including birth weight, infant survival and metabolic health of the offspring during later life. During the past few years, maternal dietary intake, at least three months before conception, has been shown to affect pregnancy outcomes also. However, literature investigating this link is still scarce. The studies that have looked at preconception dietary intake in relation to pregnancy outcome were either animal studies, had small sample sizes or focused on only selected macronutrient intake rather than complete (macro) nutrient composition or dietary intakes (e.g. food groups). Therefore, we aim to investigate the association between preconception diet and pregnancy outcomes in a linked birth cohort. The main objective of this manuscript is to describe the methodology of establishing this birth cohort and to describe both the characteristics of the study population included as well as the representativeness in terms of dietary intake.

Methods

We created the birth cohort by linking two existing databases; a large population-based cohort study in the Netherlands (The Lifelines Cohort study) and the Dutch national birth registry (Perined), through a 'trusted third party'. The birth cohort contains information on maternal dietary intake during preconception as well as pregnancy outcomes.

Results and discussion

In the Lifelines Cohort study, 3,418 pregnancies were available for linking with Perined. In total, 2,368 pregnancies (86.9%) were linked with Perined, resulting in the birth cohort. With this linked cohort we are able to provide insights on the associations between dietary intake before conception and pregnancy outcomes. Such data could potentially improve nutritional care for women of childbearing age. Lifestyle changes in the period preceding pregnancy may be most effective in improving pregnancy outcomes. A focus on this window of

regulations for study participants. Hence, other research groups can reproduce our analysis and investigate new research questions.

opportunity may provide both sufficient time, as well as a period when women are potentially motivated to adopt health optimizing behaviours.

Introduction

The central role of nutrition and metabolism in pregnancy for health and well-being of pregnant women, pregnancy outcomes, and long-term health and development of the offspring has been generally recognized. Maternal nutrition during pregnancy has been linked to birth outcomes including fetal growth, gestation length, congenital anomalies, as well as long term health effects for the offspring through intra-uterine programming with impact on later susceptibility to for instance cardiovascular diseases and Type 2 Diabetes Mellitus [1,2]. Congenital anomalies, preterm birth and low birth weight are some of the adverse pregnancy outcomes that were influenced by dietary deficiencies, and these contribute greatly to neonatal mortality and the global burden of disability later in life [3,4]. There is increasing recognition that also a woman's nutritional status before pregnancy affects maternal and child outcomes. Animal studies suggested that diet may influence oocyte quality during the preconception period as well as placenta and early embryonic development during the first trimester of pregnancy, and thereby the proficiency of the fetus, fetal development and finally health throughout life [5]. Formation of most organs and important physiological changes in the mother occurs mostly between the third and seventh week after the last menstrual period, when many women are not yet aware of being pregnant. Consequently, several teratogenic effects have been associated with dietary and/or environmental exposures during this time [6]. Therefore, it is important to optimize the woman's dietary intake in the periconceptional period to ensure optimal outcomes. De-Regil et al. [7] discussed the current guidelines on preconceptional nutrition interventions for girls and women and highlighted the gaps in evidence and guidelines. They found that, with the exception of vitamin A, folic acid, iodine and a few other nutrients [8], literature about the link between nutritional status before conception and birth outcomes was scarce. Therefore, further research needs to be done to study possible associations between preconceptional diet, fetal growth and pregnancy outcomes in large cohorts. Although there are already some existing birth cohorts (e.g. Southampton Womens Study [9]), most research in this area was focused on nutrition and lifestyle factors during (early) pregnancy, rather than the preconception period. In this manuscript we describe how we have used available dietary data from a large population-based cohort study in the Northern Netherlands (Lifelines) and linked it to the available data in the Dutch Perinatal Registry (Perined) to create a Perined-Lifelines linked birth cohort that can be used to investigate the possible association between dietary intake in the preconception period and pregnancy outcomes. In this paper, we will describe the characteristics of the Perined-Lifelines linked birth cohort and investigate to what extent this cohort is a representative sample of the Northern Netherlands population in terms of dietary intake.

Materials and methods

The Perined-Lifelines linked birth cohort consisted of female participants of childbearing age participating in the Lifelines Cohort Study, who became pregnant during their enrollment in Lifelines. Since information about the course of the pregnancy and pregnancy outcomes were not routinely collected in Lifelines, information on these pregnancies and pregnancy outcomes was derived from the Dutch Perinatal Registry (Perined).

Overall design Lifelines Cohort Study

The Lifelines Cohort Study is a prospective population-based cohort study examining the health and health-related behaviors of 167,729 persons living in the Northern Netherlands. The overall design and rationale of the study has been described in detail elsewhere [10]. In brief, between 2006 and 2013, inhabitants aged between 25 and 50 living in the northern part of the Netherlands were invited by their general practitioners (GP) to participate in Lifelines. Individuals with severe psychiatric or physical illness, limited life expectancy (<5 years) or insufficient knowledge of the Dutch language to complete a Dutch questionnaire were not considered eligible by the GP and not invited to participate. In addition, individuals who were not invited by their GP and did not meet above mentioned excluding criteria, were able to enroll themselves in the study via the Lifelines website.

Individuals who were interested received an informed consent form. After this informed consent was signed, the participants received a baseline questionnaire and an invitation for a health assessment at the Lifelines research site. During this visit, participants were asked whether family members would also be willing to participate, thereby contributing to the three-generation design. Overall, 49% of the participants (n = 81,652) were invited through their GP, 38% (n = 64,489) via participating family members and 13% (n = 21,588) self-registered via the Lifelines website. It was shown that the Lifelines population can be generalized to the general population and recruitment strategy was not subject to selection bias as it had minor effect on the level of representativeness [11].

Data collection.   Information on lifestyle factors and diet was collected at enrollment (baseline), during follow up (e.g. questionnaires), and during visits at the Lifelines research locations. The time between the baseline questionnaires and the first follow-up questionnaire was approximately two years. The time between the first follow-up questionnaire and the second follow-up questionnaire was also approximately two years. A comprehensive overview of the data collection can be found at the Lifelines catalogue at www.lifelines.nl. Self-administered questionnaires were used to collect data regarding demographics (ethnicity, education) and lifestyle (smoking, alcohol, diet). Height and body weight without shoes and heavy clothing were measured during the visit to one of the Lifelines research sites, with the SECA 222 stadiometer and the SECA 761 scale. Body mass index (BMI) in kg/m^2 was calculated. Dietary intake was retrieved through a 110 item semi-quantitative baseline food frequency questionnaire (FFQ) assessing food intake over the previous month. This FFQ was developed by Wageningen University using the Dutch FFQTOOL^{TM}, in which food items were selected based on the Dutch National Food Consumption Survey of 1997/1998 [12–14]. This FFQ was designed to include food groups that account for at least 80% of the variance and 80% of the population intake of both energy and macronutrients. Seven response categories were used to assess consumption frequency, ranging from "not this month" to "6–7 days a week". Portion size was estimated by fixed portion sizes (e.g., slices of bread, pieces of fruit) and commonly used household measures (e.g., cups, spoons). The FFQ focused on estimates of energy and macronutrients intake and food groups, including alcohol intake, using the Dutch food composition database of 2011 [15]. Reliability of reported dietary intake was assessed by the Goldberg cut-off method, which relies on the ratio of reported energy intake and basal metabolic rate [16], calculated with the Schofield equation [17]. Dietary information was available for 144,095 adults.

Design of Perinatal Registry of the Netherlands (Perined)

Information on pregnancy and pregnancy outcome (e.g. birth weight, gestational age, apgar score) was derived from Perined [18]. Perined is a national registry that contains data on

pregnancy, obstetric history and pregnancy outcomes of more than 97% of all pregnancies in the Netherlands [19]. These data were routinely collected by 94% of midwives, 99% of gynaecologists and 68% of pediatricians including 100% of Neonatal Intensive Care Unit paediatricians [19]. Perined aims to improve the quality of perinatal care through providing data for research and on audits on adverse outcomes.

Linking procedure: Involved parties and agreements

Within the Lifelines Cohort study, participant data has gone through a data verification procedure, the research data were pseudonomised and separated from personal data before it was made available for researchers [10]. Also the registry of Perined is anonymous, and does not contain a unique patient identifier available for researchers. Therefore, the available personal linking variables from both Perined and Lifelines, which were birth date and 4-digits ZIP code of the residential address of the female participants from Lifelines, and birth date of their child, were not available to the researchers.

Linking between the selection of female participants from Lifelines with Perined, was therefore done anonymously using a pseudonym with the assistance of a 'trusted third party' (ZorgTTP, Houten, The Netherlands), facilitated by Mondriaan project (UMCG)/Lygature (Utrecht, The Netherlands). 'Mondriaan' operated as a data-broker or intermediate bringing parties together offering and seeking health data for scientific research. Before the linking procedure was initiated, all the involved parties formulated and signed agreements to cover rights and obligations concerning handling and storing of the data (S1 Fig). 'ZorgTTP' was responsible for the pseudonymisation of the data, a process that included three different phases.

First, an unique pseudonym for each female participant was made, using the three personal linking variables. If one of these three variables was missing (in either Lifelines or Perined), data from the respective female participant could not be linked to Perined data. In this phase, the 'Privacy and Send Module' was used at the locations of the sources (Lifelines and Perined (performed by Research Data Support, UMCG)). The available database, at both source locations, was validated, and 'pre-pseudonyms' were created of the personal linking variables through hashing. Personal data were removed, and the remaining dataset was divided in a part with the 'pre-pseudonyms' and a part with the substantive information like variables on nutrition and pregnancy at each source location,- i.e. four datasets in total. These datasets were encrypted and send to 'ZorgTTP' using a protected connection (HyperText Transfer Protocol Secure (HTTPS)-connection).

Within the second phase, the editing phase, 'ZorgTTP' decrypted the part with the 'pre-pseudonyms' with a private key. The 'pre-pseudonyms' were subsequently encrypted for a second time. 'ZorgTTP' had no access to the part of the data with the substantive information (e.g. nutritional, pregnancy related), which was secured and could only be decrypted by the researcher.

In the last phase, the 'Aim- and Receive Module' was used, whereby the receiver (the researcher) was able to download the pseudonomised data from the server of 'ZorgTTP'. The datasets, both with the pseudonyms and with the substantive information, were decrypted by the researcher through a private key and merged subsequently. Through the corresponding pseudonyms in both Lifelines and Perined, the datasets were linked to each other, resulting a Perined-Lifelines linked birth cohort.

Design Perined-Lifelines linked birth cohort

For the Perined-Lifelines linked birth cohort, we selected female participants from the Lifelines Cohort study who indicated in their first or second follow-up questionnaire to have delivered
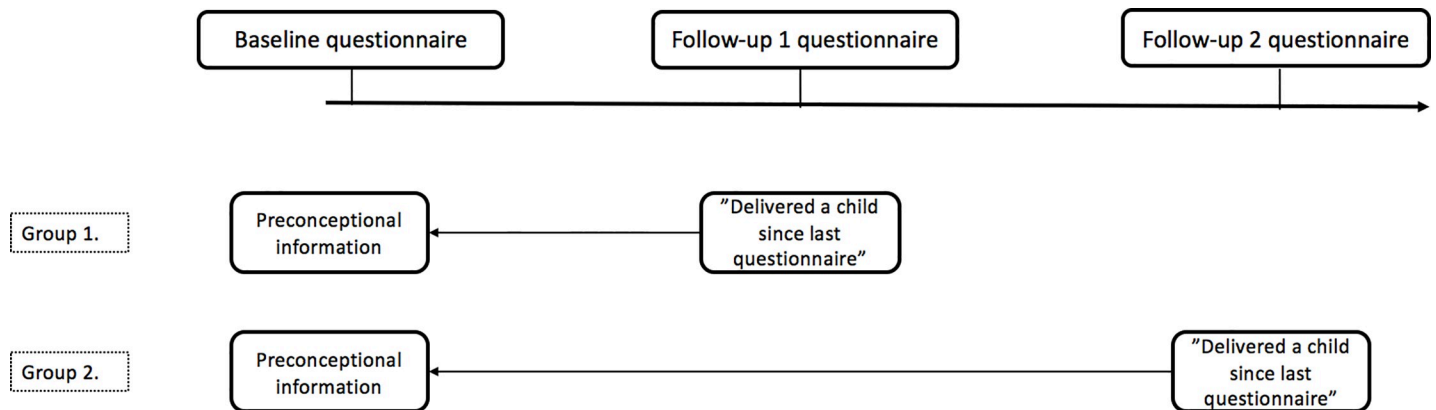
Fig 1. Overview Lifelines participants.

https://doi.org/10.1371/journal.pone.0225973.g001

a child since the previous questionnaire. The information collected at baseline (e.g. nutrient intake through a Food Frequency Questionnaire (FFQ)) was considered as the preconceptional information available for that specific pregnancy (Fig 1). The date of birth of the child was retrieved from the Dutch population register.

Inclusion and exclusion criteria. Female participants who indicated to have delivered more than one child in the period between two consecutive questionnaires were excluded for further analysis, as it would be difficult to link the correct date of birth from the Dutch population register to the corresponding child, and could result in a less reliable linked dataset. After linkage, we excluded pregnancies with preterm births (gestational age <37 weeks) and women with missing or unreliable reported dietary intake. The reliability of reported dietary intake was checked using the Goldberg cut-off method, which relies on the ratio of reported energy intake and basal metabolic rate [16] calculated with the Schofield equation [17]. The reported dietary intake of women with a ratio below 0.87 or above 2.75 was considered not reliable. In addition, intake by the women less than 500 kcal/day was considered as implausible energy intake, and therefore these women were excluded for further analyses [20,21].

Dietary intake

For the women in the final linked cohort, detailed data on dietary intake was available. This included information on intake of macronutrients (carbohydrates, proteins (total, animal, plant), fat) and energy (kcal, KJ) in grams per day. In addition, the 110 FFQ items were categorized into 22 food groups (S1 Table) [22]. Every woman contained information on the intake of every specific food group in grams per day, as well as data on protein (animal and plant), fat, carbohydrate intake contribution of the different food groups.

Ethical approval

The Lifelines Cohort Study was conducted according to the principles of the Declaration of Helsinki and is in accordance with research code University Medical Center Groningen. Within the informed consent from The Lifelines Cohort study, participants approved that their data were used for research purposes, including for linking with other existing databases. The linking procedure for the Perined-Lifelines linked birth cohort was approved as a non WMO study by the medical ethical committee of the University Medical Center Groningen, The Netherlands (METc number: 2018/506) and waived the requirement for additional informed consent. All data in this study was fully anonymized before we accessed them.

Results

Characteristics of the Perined-Lifelines linked birth cohort

In the period between 2006 up to the end of 2013, Lifelines included 167,729 participants, consisting of women, men and children (Fig 2). The total group of women in Lifelines was 97,504 (58.1% of which 48,957 are women of childbearing age (18–45 years, age at inclusion). Of these women 1,269 reported to have delivered a child between baseline and follow up-1, and 1,387 women reported to delivered a child between follow up-1 and follow up-2. In total, 381 women reported that they had one pregnancy between baseline and follow up-1, and again a pregnancy between follow up-1 and follow up-2. In total, there were 3,037 women with self-reported pregnancies between baseline and follow up-2 contributing to a total of 3,418 pregnancies.

The ZIP code of residential address and birth date of the women was available in Lifelines for all women. For 2,725 pregnancies a birth date of the child was retrieved in the Dutch population register and were therefore available for linking with Perined. In total, 2,368 pregnancies (86.9%) were linked with Perined. Of these women, 28 had two pregnancies in the linked cohort, one pregnancy between baseline and follow up-1 and one pregnancy between follow up-1 and follow-up 2. After linkage, we excluded pregnancies with pre-term births (gestational age <37 weeks; n = 110) and unknown sex of the child (n = 1). In addition, participants with unreliable or missing reported dietary intake were excluded from analyses (resp. n = 427 and n = 168), leaving 1,698 (62.3%) women in the birth cohort (Fig 2). Women with reliable dietary intake were more often higher educated, were younger at preconception, had a higher percentage of alcohol consumers, and lower percentage of smokers (S2 Table). They were more often nulliparous, and the median birth weight in this group was slightly lower compared to the excluded group with unreliable dietary intake (S2 Table).

As described, the resulting birth cohort contained data related to 1698 term pregnancies. The average maternal age at enrollment was 29 years (25th-75th percentile: 27–32 years) (Table 1). Children were on average born at 39.4 weeks of gestational age with an average birth weight of 3570 grams (25th-75th percentile: 3249–3880 grams). The majority of the children had an Apgar score of 10 at 5 min. Almost all women were of white (east/west European) ethnicity (97.8%), and the majority completed higher education (55.8%). The average maternal BMI was 23.8 kg/m^2 (25th-75th percentile: 21.7–26.6), and average energy intake amounted to 1813 kcal/day (25th-75th percentile: 1545–2141). The average time between de FFQ (baseline) and the delivery of the child was 13 months (25th-75th percentile: 11–16 months).

Representativeness of dietary intake of the Perined-Lifelines linked birth cohort

To assess to what extent our study population was a representative selection of the population in terms of dietary intake we first compared intake of energy and macronutrients with the recommended reference values from the Dutch Dietary Reference Intakes (DRI) (Table 2) [23]. 'Dietary reference intakes' is a collective term of reference values that are quantitative estimates of nutrient intakes to be used for assessing diets of healthy people, including 'estimated average requirement' (EAR), 'recommended dietary allowance' (RDA), 'adequate intake' (AI) and tolerable 'upper intake level' (UL). Given a requirement with a normal distribution, the EAR was the level of nutrient intake estimated to meet the requirement of half of the healthy individuals in a group. This EAR was given for energy intake (in kcal/day) in Table 2. As shown, almost all female participants (n = 1695, 85.5%) consumed lower energy intake than that provided by the Dutch DRI (EAR: 2318–2437) (Table 2). The number of women with an intake below the DRI
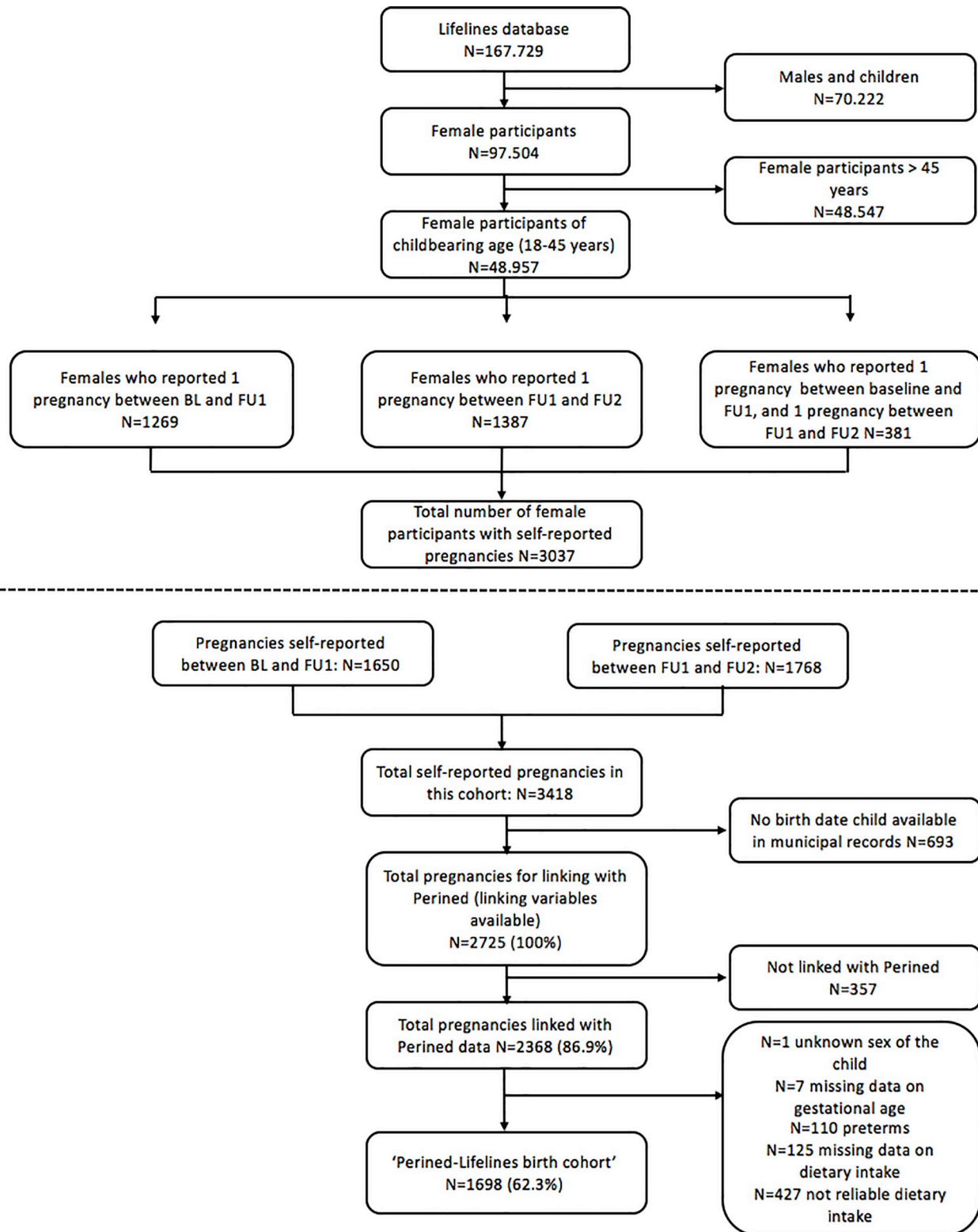
Fig 2. Flow chart of the Perined-Lifelines linked birth cohort.

https://doi.org/10.1371/journal.pone.0225973.g002

Table 1. Characteristics of the Perined-Lifelines linked birth cohort (N = 1698).

	Total cohort (n = 1698, 100%)	
Demographics		
Age at enrollment (years)	29	(27–32)
Ethnicity		
White, East/West European Ethnicity	1661	(97.8)
Other	37	(2.2)
Education[a]		
Low	105	(6.3)
Moderate	637	(37.9)
High	936	(55.8)
Missing	20	
Urbanization level by category [b]		
1	403	(24.2)
2	133	(8.0)
3	118	(7.1)
4	200	(12.0)
5	812	(48.7)
Missing	32	
Diet		
Energy intake (kcal/day)	1813	(1545–2141)
Percentage energy from: [c]		
Carbohydrates	46.8	(37.4–471.1)
Mono and Di saccharides	24.8	(22.4–284.3)
Polysaccharides	29.3	(21.3–292.9)
Protein	14.8	(15.3–128.3)
Animal protein	8.6	(4.0–84.7)
Plant protein	6.0	(2.9–65.8)
Fat	26.9	(21.1–304.8)
Lifestyle		
BMI [d] (kg/m^2)	23.8	(21.7–26.6)
BMI WHO classification		
<18.5	19	(1.1)
18.5-<25	1019	(60.0)
25-<30	469	(27.6)
≥ 30	191	(11.3)
Alcohol		
User percentage (%)	1317	(77.6)
Median consumption[e]*(g/day)*	2.7	(1.4–6.4)
Missing	2	
Smoker	212	(12.5)
Missing	3	
Pregnancy		
Maximum time between baseline questionnaire and birth child (in months)	13.0	(11.0–16.0)
Sex of the child		
Male	849	(50.0)
Gravidity		
1	692	(40.8)
2	578	(34.0)

(*Continued*)

Table 1. (Continued)

	Total cohort (n = 1698, 100%)	
3	275	(16.2)
≥4	153	(9.0)
Parity		
0	813	(47.9)
1	632	(37.2)
≥2	253	(14.9)
Birth weight (in grams)	3570	(3249–3880)
Missing	4	
Gestational age (in weeks)	39	(39–40)
Apgar-score (after 5 min)		
<10	422	(24.9)
10	1274	(75.1)
Missing	2	

Data are median (IQR) or n (%). Data were complete when there is no missing row presented.

[a] Low education: primary school, vocational and lower general secondary education; Moderate education: higher secondary education and intermediate vocational training; High education: higher vocational education and university education.

[b] Level of urbanization: 1. Very high $>$ = 2500 addresses per km^2; 2: high 1500–2500 addresses per km^2; 3: moderate 1000–1500 addresses per km^2; 4: low 500–1000 addresses per km^2; 5: rural $<$500 addresses per km^2.

[c] Energy from carbohydrates, protein and fat, relative to the sum of energy from the three macronutrients.

[d] BMI = Body mass index.

[e] Median + IQR among alcohol users. One standard drink contains 10 g alcohol.

were interpreted with caution, since, as mentioned, the EAR is the level of intake that is adequate for only half of the population [23]. In addition, the FFQ was designed to include food groups that account for at least 80% of the variance and 80% of the population intake of both energy and macronutrients. Therefore, the actual energy intake of the women could be higher than that described in the FFQ results, making a strict comparison with the DRI difficult. The RDA was calculated as the EAR plus twice the standard deviation of the requirement. This described the average daily intake level that is sufficient to meet the nutrient requirement of nearly all individuals in a group. In case the RDA cannot be determined, the AI was used, a value based on observed or experimentally determined approximations of nutrient intake by a group of healthy people. This reference value was used for protein, carbohydrates and fat intake. Finally, the UL was the highest level of daily nutrient intake that is likely to elicit no risk of adverse health effects to almost all individuals in a group. In case the intake increased above the UL the risk of adverse health effects was increased. This level was known for protein and fat intake, but not for carbohydrates. This resulted in a relative narrow DRI for carbohydrates (40–42 en% AI), resulting in 83.5% of the women consuming more carbohydrates than this 'adequate intake' in our assessment. Intake of total protein and fat was adequate according to the guideline for majority of the female participants, respectively 99.8% and 86.6%.

Since our group of female participants was a selection of the Lifelines cohort participants, we also wanted to assess whether the diet quality scores from the linked Lifelines-Perined birth cohort were comparable to the complete Lifelines cohort. A recent paper assessed diet quality in the complete Lifelines cohort (adult males and females) using a Lifelines Diet Score (LLDS) [22] based on the 110 FFQ food items. The LLDS was formulated to represent relative diet quality. To account for differences in energy intake between individuals, intake of food groups

Table 2. Number of female participants meeting dietary reference intake for macronutrients (protein, carbohydrates and fat) and energy.

	DRI [a]	Intake below DRI		Intake according to DRI		Intake above DRI	
		Mean (SD)	N (%)	Mean (SD)	N (%)	Mean (SD)	N (%)
Energy (kcal/day)	2318–2437 (EAR [b])	1727 (328.5)	1452 (85.5)	2374 (32.6)	72 (4.3)	2737 (287.0)	174 (10.2)
Protein (En%)	9–25 (AI [c]-UL [d])	8.8 (0.07)	3 (0.2)	14 (2.1)	1695 (99.8)	-	-
Carbohydrates (En%)	40–42 (AI)	37.5 (2.9)	145 (8.5)	41.2 (0.6)	136 (8.0)	48.1 (3.8)	1417 (83.5)
Fat (En%)	20–40 (AI-UL)	19.0 (0.8)	2 (0.1)	33.8 (3.7)	1471 (86.6)	42.4 (2.2)	225 (13.3)

[a]DRI = Dutch dietary Reference Intakes; reference intakes for Dutch non-pregnant women between 22–50 years of age [23]

[b]EAR = Estimated Average Requirements; an estimated of the average requirements of energy or nutrient needed by a group of people.

[c]AI = Adequate Intake.

[d]UL = Upper Limit.

https://doi.org/10.1371/journal.pone.0225973.t002

was expressed in grams per 1000 kilocalories (kcal) instead of grams per day. For each food group, intake was divided into quintiles to score individual consumption relative to the complete study population (n = 129.369). The quintiles ranged from 0 to 4, with 4 points being awarded to the highest quintile of consumption for positive food groups, and to the lowest quintile of consumption for negative food groups. The sum of the 12 component scores resulted in a LLDS score, of which LLDS quintiles were made. With the absolute cutoffs from the LLDS quintiles formulated in the complete Lifelines cohort study, we compared our birth cohort in terms of diet quality. Indeed, in our cohort also approximately 20% of the female participants fell in every quintile with absolute cutoffs used comparable to the Lifelines Cohort.

Discussion

Using data from the Dutch Perinatal Registry (Perined) linked to the Lifelines Cohort Study, we have created a new birth cohort consisting of women from childbearing age who delivered a child after their enrollment in Lifelines, providing information on dietary intakes during the period of preconception.

In this manuscript we have given an overview on the methodological approach and the development of the Perined-Lifelines linked birth cohort and described the characteristics of this group. Due to the detailed data collection on diet at lifestyle factors before pregnancy, the Lifelines Cohort offers a unique opportunity to increase the knowledge on the possible associations between specific habitual dietary intake data in the preconception period and pregnancy outcomes. Since information about the course of the pregnancy and pregnancy outcomes are not routinely collected in Lifelines, information on these pregnancies and pregnancy outcomes was derived from the Dutch Perinatal Registry (Perined). The current study set up allows us to focus on diet characteristics in the preconception period (defined as the period between baseline, when dietary intake was assessed, and follow-up questionnaires in which the participant indicated to have delivered a child since her enrollment in Lifelines). Linking these Lifelines data to the data from Perined, allowed us to investigate possible associations to pregnancy outcome.

Most of the published literature that examines the relationship between nutrition and pregnancy outcomes, has focused on data collection during the second and/or third trimester of pregnancy. However, the women's nutritional status before conception, and/or during early pregnancy, when women are mostly unaware of their pregnancy, may influence conception as well as early placenta, embryonic and fetal development and consequently pregnancy outcomes. Literature on the relation between diet in the periconceptional period and pregnancy outcomes is still scarce.

During early pregnancy the formation of most organs as well as placental development takes place and is directly influenced by maternal health and nutritional status [24]. Animal studies have previously shown that both maternal undernutrition as well as overnutrition, before and during pregnancy, reduces placental-fetal blood flow and consequently impacts fetal growth. If the impact of nutritional status of the women could be optimized before pregnancy, outcomes for both mother and infant could be improved. Reproduction may motivate women to adapt health optimizing behaviours for the well-being of their unborn child [25]. However, several intervention studies during pregnancy showed that (long-term) beneficial lifestyle changes are difficult to maintain, and if effective, it results in only modest, short term changes [26,27]. Instead, the preconception period may be a more optimal period to intervene [28–30].

A few studies have investigated the association between preconceptional nutrient intake and pregnancy outcomes. Observational studies have suggested that pre- and periconceptional intake of vitamins and mineral supplements are associated with a reduced risk of preterm delivery, as well as delivering offspring with low birth weight. A recent systematic review concluded however that the overall quality of the evidence was low for most outcomes with the exception of the benefit of maternal preconception folic acid intake and the risk of neural tube defects [31]. Several cohort studies have suggested that dietary patterns up to three years before pregnancy, characterized by fruit, vegetables, legumes, nuts and fish and low intake of red and processed meat, are associated with reduced risk of gestational diabetes [32–35], hypertensive disorders of pregnancy [36,37], and preterm birth [38]. Since few people will plan a pregnancy three years in advance, this highlights the need for interventions on population level in future. In addition, in most cases the preconceptional period is most often determined in retrospect, underlying the importance of increasing awareness early in women of childbearing age.

The linked Perined-Lifelines birth cohort demonstrates the feasibility of linking (healthcare) databases through a pseudonymised linking procedure. Our method is generalizable to linkage of administrative data in other contexts where data is only given out anonymously to the researcher. Errors occurring during linkage (missed-matches and false matches) can result in substantially biased results: false-matches can bias associations towards the null and missed-matches can lead to a selection bias [39]. Our linkage was done on three personal identifiable variables; 4-digits ZIP code of the residential address, the birth date of the female participants from Lifelines and birth date of their child, making the chances of errors (in terms of missed-matches and false matches) low, and resulting in a high linkage rate of 86% when information for all variables used for linkage are complete. To facilitate such a linking procedure, strict access arrangements and secure data transfer processes were established. In addition, secure physical locations were created with restricted network and no internet access. In this linking procedure, several parties were involved (S1 Fig), and agreements had to be made and signed between all the different parties, making it a time consuming process.

A limitation of linking administrative or electronic healthcare data is the imperfect nature of data for research purpose, as often some information important and interesting for research may be missing [40]. For example, information on gestational weight gain was not available in this study. However, the large sample size and representativeness of this linked cohort offers a cost-effective alternative to traditional birth cohort studies, providing valuable information on dietary intake during preconception and pregnancy outcomes in a representative group of pregnant women from the Netherlands. It has previously been described that the complete Lifelines study population is broadly representative for the adult population of the north of the Netherlands [11]. Although selection bias is a risk in a population-based cohort study, the recruitment strategy of the Lifelines participants had no effect on the representativeness of the cohort, indicating that the risk of selection bias was low and that the results that will be found

in the Lifelines cohort can be generalized to the general population. In terms of demographic variables our birth cohort was comparable with the complete Lifelines cohort. Recent literature showed that within the complete Lifelines cohort, participants were more often Dutch and that less women had a low educational attainment [22]. The overall alcohol user percentage and intake however was much lower in our cohort, compared to other studies [22], likely due to the fact that we have a cohort of only women, and that women tend to drink less and less often than men [41,42]. In addition, the women were of childbearing age and consequently might be more conscious to preserve a healthy lifestyle also anticipating a possible (future) pregnancy. Also in terms of dietary intake our cohort seemed to be representative both compared to the complete Lifelines cohort [22], as well as compared to the Dutch dietary guidelines [23]. However, as described, 85.5% of the women had an energy intake below the Estimated Average Requirement (EAR), which is possibly due to the fact that this DRI describes adequate requirements for around 50% of the women. Therefore, women consuming less energy intake than the EAR could still meet their individual daily energy requirement. In addition, the FFQ was not designed to describe 100% of the individuals energy intake, but to account for at least 80% of the variance of dietary intake and 80% of the population intake of both energy and macronutrients. Therefore it may be possible that the actual intake of energy was actually higher than described in the results from the FFQ. The reference values used for the intake of macronutrients (e.g. protein, fat) are way more robust as they apply for all individuals in a group. The fact that the intake of these nutrients was within the normal limits when compared to the DRI, shows that the dietary intake of our cohort is likely to be representative in comparison with the DRI.

Dietary intake was assessed via a FFQ, which was validated [43] and contained questions concerning dietary intake over the previous month. The advantage of recording dietary intake retrospectively, was is that the process of recording will not alter awareness of food intake and will not improve accuracy in completing the questionnaire, something which is the case in prospectively retrieved diet records [44]. Previous studies have shown that energy intake reported by FFQ showed, on average, very good agreement with actual energy intake during controlled feeding trials, while body weights were kept stable [13]. According to the literature, accurate reporting of energy intake was influenced by several factors including sex, age, educational level, BMI, psychosocial factors and lifestyle [13]. On most of these factors we have data available in our cohort, and are therefore able to investigate its possible effect and we can correct for it. Based on the Goldberg cut-off method, using reported energy intake and basal metabolic rate (14), and calculated with the Schofield equation (15), we excluded all women with reported unreliable dietary intake. We compared these women in terms of demographic variables with women having a reliable dietary intake reported (S2 Table). Women with reliable dietary intake were more often higher educated, slightly younger at preconception, the percentage of alcohol consumers was higher and percentages of the percentage of smokers was lower. We consider these differences logical and do not expect selection bias to have played a role. Therefore we do not expect this to influence our results. In addition, women with reliable data were more often nulliparous, and consequently had a slightly lower mean birth weight (3570 grams vs 3640 grams) but we do not consider this as a clinical relevant difference.

For all women, we used the FFQ data collected at baseline as their preconception' dietary intake. The time between baseline, and the moment the woman delivered her child varied among women included in our birth cohort between a minimum of 5 and a maximum of 37 months. In general, it has been shown that diet tends to be quite stable over time [44], and that changes in dietary habits after conception tend to be modest and mostly reflect intake before conception [45].

Preterm births were excluded from our cohort as the causes of preterm birth are complex and the pathophysiology that triggers it, is largely unknown [46,47]. This made it difficult to adjust for underlying pathophysiology, potentially leading to results confounded by this pathophysiology, and therefore difficult to interpretate.

As described, we noticed a discrepancy in the number of self-reported deliveries by female participants from Lifelines (n = 3,418), and the registered children in the Dutch population register (n = 2,725) in the time between baseline and FU2. This could be explained by several reasons. First, it is plausible that female participants that have indicated at FU2 that they delivered a child since the last questionnaire (FU1), have misreported the pregnancy and that this pregnancy was already reported before at follow-up 1. Such a self-reported delivery will, unjustified, be counted twice in the Lifelines cohort, but is only recorded once in the Dutch population register. In addition, only babies born after at least 24 weeks of gestational age are registered in the Dutch population registers. Women who delivered a baby with a gestational period less than 24 weeks could report child delivery, but this child may not have been registered in the Dutch population register, although this number was probably very low. Clearly, we cannot explain discrepancy based on the Dutch population register since it has been shown that 98% of the addresses and persons in the Dutch Population register were recorded correctly [48].

Conclusions

Using data from the national birth registry in the Netherlands, we established a large, linked cohort of female participants from Lifelines, that is representative of the population in north Netherlands and will enable us to investigate dietary intake during the period prior to conception and pregnancy outcomes. With this cohort we may provide insights on the influence of dietary intake before conception to pregnancy outcomes. Such data could inform the development of recommendations to improve nutritional care for women of childbearing age, aiming to improve health of women and future generations.

Supporting information

S1 Fig. Overview involved parties and agreements. 1. Data Processing Agreement; 2. Data transfer Agreement; 3. General Agreement; 4. Data Use and Access Agreement; 5. Data Services and Processor Agreement. UMCG = University Medical Centre of Groningen, TTP = Trusted Third Party.
(DOCX)

S1 Table. Food groups.
(DOCX)

S2 Table. Characteristics stratified by reliability dietary intake. Data are median (IQR) or n (%). Data were complete when there is no missing row presented. [a]Low education: primary school, vocational and lower general secondary education; Moderate education: higher secondary education and intermediate vocational training; High education: higher vocational education and university education. [b]Level of urbanization: 1. Very high $> = 2500$ addresses per km^2; 2: high 1500–2500 addresses per km^2; 3: moderate 1000–1500 addresses per km^2; 4: low 500–1000 addresses per km^2; 5: rural <500 addresses per km^2. [c]Energy from carbohydrates, protein and fat, relative to the sum of energy from the three macronutrients. [d]BMI = Body mass index. [e]Median + IQR among alcohol users. One standard drink contains 10 g alcohol. [1] Willet-Schofield. [2]Two sided p-value; Mann Whitney U test for continuous

characteristics or Pearson Chi-Square for categorical characteristics.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Nastaran Salavati, Marian K. Bakker, Eline M. van der Beek, JanJaap H. M. Erwich.

Formal analysis: Nastaran Salavati.

Methodology: Nastaran Salavati, Marian K. Bakker.

Project administration: Nastaran Salavati.

Supervision: Marian K. Bakker, Eline M. van der Beek, JanJaap H. M. Erwich.

Validation: Nastaran Salavati.

Visualization: Nastaran Salavati.

Writing – original draft: Nastaran Salavati.

Writing – review & editing: Nastaran Salavati, Marian K. Bakker, Eline M. van der Beek, Jan-Jaap H. M. Erwich.

References

1. Painter RC, Roseboom TJ, Bleker OP. Prenatal exposure to the Dutch famine and disease in later life: An overview. Reprod Toxicol. 2005; 20: 345–352. https://doi.org/10.1016/j.reprotox.2005.04.005 PMID: 15893910

2. Vaag A, Brøns C, Gillberg L, Hansen NS, Hjort L, Arora GP, et al. Genetic, nongenetic and epigenetic risk determinants in developmental programming of type 2 diabetes. Acta Obstet Gynecol Scand. 2014; 93: 1099–1108. https://doi.org/10.1111/aogs.12494 PMID: 25179736

3. WHO | World Health Statistics 2015. WHO. World Health Organization; 2016; Available: http://www.who.int/gho/publications/world_health_statistics/2015/en/

4. De-Regil LM, Fernández-Gaxiola AC, Dowswell T, Peña-Rosas JP. Effects and safety of periconceptional folate supplementation for preventing birth defects. De-Regil LM, editor. Cochrane database Syst Rev. Chichester, UK: John Wiley & Sons, Ltd; 2010; CD007950. https://doi.org/10.1002/14651858.CD007950.pub2 PMID: 20927767

5. Kind KL, Moore VM, Davies MJ. Diet around conception and during pregnancy—effects on fetal and neonatal outcomes. Reprod Biomed Online. 2006; 12: 532–41. Available: http://www.ncbi.nlm.nih.gov/pubmed/16790095 https://doi.org/10.1016/s1472-6483(10)61178-9 PMID: 16790095

6. Williamson CS. Nutrition in pregnancy. Nutr Bull. Wiley/Blackwell (10.1111); 2006; 31: 28–59. https://doi.org/10.1111/j.1467-3010.2006.00541.x

7. De-Regil LM, Harding KB, Roche ML. Preconceptional Nutrition Interventions for Adolescent Girls and Adult Women: Global Guidelines and Gaps in Evidence and Policy with Emphasis on Micronutrients. J Nutr. 2016; 146: 1461S–1470S. https://doi.org/10.3945/jn.115.223487 PMID: 27281812

8. Hanson MA, Bardsley A, De-Regil LM, Moore SE, Oken E, Poston L, et al. The International Federation of Gynecology and Obstetrics (FIGO) recommendations on adolescent, preconception, and maternal nutrition: "Think Nutrition First". Int J Gynaecol Obstet. 2015; 131 Suppl 4: S213–53. https://doi.org/10.1016/S0020-7292(15)30034-5

9. Inskip HM, Godfrey KM, Robinson SM, Law CM, Barker DJP, Cooper C, et al. Cohort profile: The Southampton Women's Survey. Int J Epidemiol. Europe PMC Funders; 2006; 35: 42–8. https://doi.org/10.1093/ije/dyi202 PMID: 16195252

10. Scholtens S, Smidt N, Swertz MA, Bakker SJL, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol. 2015; 44: 1172–80. https://doi.org/10.1093/ije/dyu229 PMID: 25502107

11. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the Life-Lines Cohort Study. Ali RI, editor. PLoS One. 2015; 10: e0137203. https://doi.org/10.1371/journal.pone.0137203 PMID: 26333164

12. Molag ML, de Vries JHM, Duif N, Ocké MC, Dagnelie PC, Goldbohm RA, et al. Selecting informative food items for compiling food-frequency questionnaires: comparison of procedures. Br J Nutr. 2010; 104: 446–456. https://doi.org/10.1017/S0007114510000401 PMID: 20374682

13. Siebelink E, Geelen A, de Vries JHM. Self-reported energy intake by FFQ compared with actual energy intake to maintain body weight in 516 adults. Br J Nutr. 2011; 106: 274–81. https://doi.org/10.1017/S0007114511000067 PMID: 21338536

14. Streppel MT, de Vries JHM, Meijboom S, Beekman M, de Craen AJM, Slagboom PE, et al. Relative validity of the food frequency questionnaire used to assess dietary intake in the Leiden Longevity Study. Nutr J. 2013; 12: 75. https://doi.org/10.1186/1475-2891-12-75 PMID: 23758629

15. RIVM/Dutch Nutrition Centre; Bilthoven TN 2011. N-T. NEVO online | RIVM.

16. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. Int J Obes Relat Metab Disord. 2000; 24: 1119–30. Available: http://www.ncbi.nlm.nih.gov/pubmed/11033980 https://doi.org/10.1038/sj.ijo.0801376 PMID: 11033980

17. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr. 1985; 39 Suppl 1: 5–41. Available: http://www.ncbi.nlm.nih.gov/pubmed/4044297

18. Perined, Perinatal care in the Netherlands 2006–2016. Utrecht; 2018.

19. Perined. Perinatal Care in the Netherlands 2016. Utrecht; 2018.

20. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997; 65: 1220S–1228S. https://doi.org/10.1093/ajcn/65.4.1220S PMID: 9094926

21. Rhee JJ, Sampson L, Cho E, Hughes MD, Hu FB, Willett WC. Comparison of Methods to Account for Implausible Reporting of Energy Intake in Epidemiologic Studies. Am J Epidemiol. 2015; 181: 225–233. https://doi.org/10.1093/aje/kwu308 PMID: 25656533

22. Vinke PC, Corpeleijn E, Dekker LH, Jacobs DR, Navis G, Kromhout D. Development of the food-based Lifelines Diet Score (LLDS) and its application in 129,369 Lifelines participants. Eur J Clin Nutr. 2018; 72: 1111–1119. https://doi.org/10.1038/s41430-018-0205-z PMID: 29895847

23. Gezondheidsraad. Voedingsnormen: energie, eiwitten, vetten en verteerbare koolhydraten. Gezondheidsraad. 2001; Available: https://www.mvo.nl/media/gezondheid/voedingsnormen_gr_2001.pdf

24. Wu G, Bazer FW, Cudd TA, Meininger CJ, Spencer TE. Maternal Nutrition and Fetal Development. J Nutr. 2004; 134: 2169–2172. https://doi.org/10.1093/jn/134.9.2169 PMID: 15333699

25. Phelan S. Pregnancy: a "teachable moment" for weight control and obesity prevention. Am J Obstet Gynecol. 2010; 202: 135.e1–8. https://doi.org/10.1016/j.ajog.2009.06.008 PMID: 19683692

26. Stephenson J, Heslehurst N, Hall J, Schoenaker DAJM, Hutchinson J, Cade JE, et al. Before the beginning: nutrition and lifestyle in the preconception period and its importance for future health. Lancet (London, England). 2018; 391: 1830–1841. https://doi.org/10.1016/S0140-6736(18)30311-8 PMID: 29673873

27. van Elten TM, Karsten MDA, Geelen A, Gemke RJBJ, Groen H, Hoek A, et al. Preconception lifestyle intervention reduces long term energy intake in women with obesity and infertility: a randomised controlled trial. Int J Behav Nutr Phys Act. BioMed Central; 2019; 16: 3. https://doi.org/10.1186/s12966-018-0761-6 PMID: 30621789

28. Altfeld S, Handler A, Burton D, Berman L. Wantedness of pregnancy and prenatal health behaviors. Women Health. 1997; 26: 29–43. Available: http://www.ncbi.nlm.nih.gov/pubmed/9525267

29. McBride CM, Emmons KM, Lipkus IM. Understanding the potential of teachable moments: the case of smoking cessation. Health Educ Res. 2003; 18: 156–70. Available: http://www.ncbi.nlm.nih.gov/pubmed/12729175 https://doi.org/10.1093/her/18.2.156 PMID: 12729175

30. Barker M, Dombrowski SU, Colbourn T, Fall CHD, Kriznik NM, Lawrence WT, et al. Intervention strategies to improve nutrition and health behaviours before conception. Lancet (London, England). Europe PMC Funders; 2018; 391: 1853–1864. https://doi.org/10.1016/S0140-6736(18)30313-1 PMID: 29673875

31. Ramakrishnan U, Grant F, Goldenberg T, Zongrone A, Martorell R. Effect of women's nutrition before and during early pregnancy on maternal and infant outcomes: a systematic review. Paediatr Perinat Epidemiol. 2012; 26 Suppl 1: 285–301. https://doi.org/10.1111/j.1365-3016.2012.01281.x PMID: 22742616

32. Zhang C, Schulze MB, Solomon CG, Hu FB. A prospective study of dietary patterns, meat intake and the risk of gestational diabetes mellitus. Diabetologia. 2006; 49: 2604–2613. https://doi.org/10.1007/s00125-006-0422-1 PMID: 16957814

33. Tobias DK, Zhang C, Chavarro J, Bowers K, Rich-Edwards J, Rosner B, et al. Prepregnancy adherence to dietary patterns and lower risk of gestational diabetes mellitus. Am J Clin Nutr. 2012; 96: 289–295. https://doi.org/10.3945/ajcn.111.028266 PMID: 22760563

34. Bao W, Bowers K, Tobias DK, Olsen SF, Chavarro J, Vaag A, et al. Prepregnancy low-carbohydrate dietary pattern and risk of gestational diabetes mellitus: a prospective cohort study. Am J Clin Nutr. American Society for Nutrition; 2014; 99: 1378–84. https://doi.org/10.3945/ajcn.113.082966 PMID: 24717341

35. Harding KB, Peña-Rosas JP, Webster AC, Yap CM, Payne BA, Ota E, et al. Iodine supplementation for women during the preconception, pregnancy and postpartum period. Cochrane Database Syst Rev. 2017; 3: CD011761. https://doi.org/10.1002/14651858.CD011761.pub2 PMID: 28260263

36. Schoenaker DA, Soedamah-Muthu SS, Callaway LK, Mishra GD. Prepregnancy dietary patterns and risk of developing hypertensive disorders of pregnancy: results from the Australian Longitudinal Study on Women's Health. Am J Clin Nutr. 2015; 102: 94–101. https://doi.org/10.3945/ajcn.114.102475 PMID: 26040639

37. Schoenaker DA, Soedamah-Muthu SS, Mishra GD. Quantifying the mediating effect of body mass index on the relation between a Mediterranean diet and development of maternal pregnancy complications: the Australian Longitudinal Study on Women's Health. Am J Clin Nutr. 2016; 104: 638–645. https://doi.org/10.3945/ajcn.116.133884 PMID: 27465377

38. Gresham E, Collins CE, Mishra GD, Byles JE, Hure AJ. Diet quality before or during pregnancy and the relationship with pregnancy and birth outcomes: the Australian Longitudinal Study on Women's Health. Public Health Nutr. 2016; 19: 2975–2983. https://doi.org/10.1017/S1368980016001245 PMID: 27238757

39. Harron K, Wade A, Gilbert R, Muller-Pebody B, Goldstein H. Evaluating bias due to data linkage error in electronic healthcare records. BMC Med Res Methodol. 2014; 14: 36. https://doi.org/10.1186/1471-2288-14-36 PMID: 24597489

40. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Med. Public Library of Science; 2015; 12: e1001885. https://doi.org/10.1371/journal.pmed.1001885 PMID: 26440803

41. Van Rossum CTM, Fransen HP, Verkaik-Kloosterman J, Buurma-Rethans EJM, Ocké MC. Dutch National Food Consumption Survey 2007–2010 [Internet]. 2007. Available: www.rivm.nl/vcp/

42. Van Rossum C, Buurma-Rethans EJM. Vennemann F, Beukers M, Brants H, De Boer E, et al. The diet of the Dutch: Results of the first two years of the Dutch National Food Consumption Survey 2012–2016. [Internet]. Bilthoven; 2016. Available: https://www.rivm.nl/bibliotheek/rapporten/2016-0082.pdf

43. Sluik D, Geelen A, de Vries JHM, Eussen SJPM, Brants HAM, Meijboom S, et al. A national FFQ for the Netherlands (the FFQ-NL 1.0): validation of a comprehensive FFQ for adults. Br J Nutr. 2016; 116: 913–923. https://doi.org/10.1017/S0007114516002749 PMID: 27452894

44. Willett WC. Nutritional epidemiology (3rd edition). Oxford University Press; 2013.

45. King JC. A Summary of Pathways or Mechanisms Linking Preconception Maternal Nutrition with Birth Outcomes. J Nutr. 2016; 146: 1437S–1444S. https://doi.org/10.3945/jn.115.223479 PMID: 27281808

46. Ananth C V., Ananth C V., Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. J Matern Neonatal Med. 2006; 19: 773–782. https://doi.org/10.1080/14767050600965882 PMID: 17190687

47. Villar J, Papageorghiou AT, Knight HE, Gravett MG, Iams J, Waller SA, et al. The preterm birth syndrome: a prototype phenotypic classification. Am J Obstet Gynecol. 2012; 206: 119–123. https://doi.org/10.1016/j.ajog.2011.10.866 PMID: 22177191

48. Wuijts F. Report–Snapshot quality Dutch Population Register 2011 (Dutch report: rapport–Momentopname Kwaliteit GBA 2011). 2011.