

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The association between neighbourhood deprivation, fetal growth, small-for-gestational age and preterm birth: a population-based prospective cohort study.
AUTHORS	Gootjes, Dionne; Posthumus, Anke G; Jaddoe, Vincent; Steegers, E.A.P.

VERSION 1 – REVIEW

REVIEWER	Brown, Hillary University of Toronto at Scarborough, Health & Society
REVIEW RETURNED	04-Mar-2021

GENERAL COMMENTS	<p>This is an interesting study looking at the impact of neighbourhood deprivation on fetal growth and other adverse pregnancy outcomes using data from the Generation R Cohort Study in the Netherlands. The paper is well written and the study well executed. I have the following comments:</p> <p>Introduction:</p> <ol style="list-style-type: none">1. I thought the Introduction should have included some more information on why and how neighbourhood deprivation could impact pregnancy outcomes above and beyond the effects of individual level SES, etc. <p>Methods:</p> <ol style="list-style-type: none">1. Please define the acronym CRL (crown rump length) at its first use (page 5).2. I understand that ethnicity was measured as a covariate, but was it accounted for in the calculation of fetal growth? See for example: https://www.jogc.com/article/S1701-2163(16)35159-3/abstract3. Why were preterm birth and SGA the only measures of adverse pregnancy outcomes? Given the label, I expected a broader range of outcomes. A rationale for the choice of outcomes would be useful – e.g., are they most likely to be affected by neighbourhood deprivation vs. other complications such as gestational hypertension, gestational diabetes, etc.4. Why did the covariates include complications in a previous pregnancy and not the current pregnancy?5. The authors chose to use a logistic regression model, but interpret the ORs in terms of risk. I understand that most of the outcome rates are < 10%; however, SGA is 12.2% in the exposed group, and some of the other outcomes also hover around 10%. I would recommend not interpreting the OR as indicating risk but
-------------------------	---

	<p>rather as odds – or to use a method (e.g., modified Poisson regression, binomial regression, etc) that allows direct estimation of relative risk.</p> <p>Discussion:</p> <p>6. The data are from the Generation R Cohort Study which was conducted on births between 2002 and 2006. The most recent data are therefore over 15 years old. Could the authors provide more information on whether the findings are still relevant and how time may or may not affect their interpretation? Further detail on this should be added to the Discussion.</p> <p>7. I was surprised that the focus of the implications appeared to be on targeting health behaviours of residents of deprived neighbourhoods. If the effect of neighbourhood deprivation remained after controlling for (some of) these individual level factors, I think an important implication relates to population-level interventions such as policies that promote healthier environments (e.g., more walkable neighbourhoods, universal basic income, etc) – not interventions that rely on individual behavior change. Could the authors address this?</p> <p>8. The authors may also want to emphasize further that the magnitude of the findings is somewhat small (though statistically significant). What impact does this have on interpretation in terms of expected effect of interventions on neighbourhood deprivation?</p> <p>Figures and tables:</p> <p>9. Please add the n (%) with each outcome to the tables for the different exposure groups to give a sense of the public health importance of the finding – e.g., in Table 2.</p>
--	---

REVIEWER	VENDITTELLI, FRANÇOISE Academic Hospital of Clermont-Ferrand, Gynécologie Obstétrique
REVIEW RETURNED	21-Mar-2021

GENERAL COMMENTS	<p>Reviewer's comments on manuscript bmjopen-2021-049075</p> <p>Comments to the authors:</p> <p>1) Synthesis of the article</p> <p>The article assesses the association between neighborhood deprivation and fetal growth during the first second and third trimester of pregnancy, and adverse pregnancy outcomes. The authors did not find an association between neighborhood deprivation and the first trimester growth. The found a negative association between neighborhood deprivation and the fetal growth in the second and third trimester of pregnancy, and with a risk of SGA and preterm birth.</p> <p>2) General comments</p> <p>Thank you for your efforts to publish your work but this article needs several clarifications to improve its reading and internal and external validity.</p> <p>3) Detailed comments</p>
-------------------------	--

3.1) Title

It should include the study design: "neighborhood deprivation, fetal growth and adverse pregnancy outcomes: a population based cohort study; the Generation R Study.

Could you explain what is the signification of "R" of Generation R Study?

3.2) Introduction

The justification of this study is not well explained. We know that low economic status is associated with adverse pregnancy outcomes. We have several publications even during wars.

Why did you realize a new study as we have yet some good studies on the topic? You explained to the readers "It is however unknown whether this potentially modifiable factor is also associated with an early fetal development". Could you give to the readers some animal or physiological studies to support a research hypothesis?

The objectives are not well introduced at the end of the introduction "therefore..." and we do not know what the main objective is and what the secondary objectives are.

3.3) Methods

- Describe the materials section in a separate chapter of the study design.

- Materials: Why did you exclude the spontaneous and late abortion, and the intra-uterine deaths? They are part of adverse pregnancy outcomes and are pertinent variables for the research topic. Could you specify the years of the studied births. What was the beginning of the enrollments and the end of the follow-up? Did you exclude the in utero fetal anomalies? Did the women should speak Dutch?

- Patient and public involvement statement: you are in contradiction with line 363.

- Neighborhood deprivation: for me it is the description of the exposition status. This is a geographic score, not an individual score. Could you define what are in your country a low income and a low educational level. What is the black work in your country (if you have a high level of black work you may have an information bias, at minimum this point has to be discuss in the discussion chapter). Your data are old, why? The recent economic crisis may have increased the proportion of women living in a neighborhood deprivation context.

- Covariables:

• I do not understand the rational of your choice to work on the last menstrual period. Even with a regular cycle the best way to have

the last calculated menstruation is to have a good first trimester ultrasounds. If you think that the first growth is not uniformed among women or among ethnicity groups then you will not controlled this fact by using last menstrual period. This is a big weakness of this work.

If I understand, the pregnancy dating process was not the same among all the women included in the study (line 105-106) which is also an epidemiological problem.

- Growth parameters: We need to know if you made an audit to assess the quality of the fetal measurements according to your national guidelines. If the clinicians were each responsible of the ultrasounds assessment, even in a research center, you may have an information bias.

- Adverse pregnancy outcomes: I do not understand why you did not assess others adverse pregnancy outcomes which are well known to be associated with individual low incomes or neighborhood deprivation (Large for gestational age, gestational diabetes). Why did you choose the 10th percentile like endpoint for SGA?

We have an international definition of the small-for-gestational age advisory board consensus statement of 2001 which is not the 10th [Lee PA, Chernausk SD, Hokken-Koelega ACS, Czernichow P for the International SGA Advisory Board: International Small for Gestational Age Advisory Board consensus development conference statement: management of short children born small for gestational age, April 24-October 1, 2001. Pediatrics 2003, 111: 1253-61].

FGR is different of SGA (Peter et al in 2003 (international small for gestational age advisory consensus development conference statement.. » SGA refers not to fetal growth but to the size of infant at birth. The term intra uterine growth retardation (IUGR) suggests a diminished growth velocity in the fetus. You have a prospective cohort and so you should be able to classify FGR and SGA. I also think that, maybe, more the social deprivation is more the SGA is severe. It could be interesting to have this information in the descriptive tables.

- Clarify the chapter: all the endpoints in a same chapter (growth parameters and selected pregnancy outcomes).

- We have a few explanations about the questionnaires. What was the language of the questionnaires?

3.4) Results

We have only one descriptive table "baseline characteristics" but we do not have a descriptive table for the pregnancy follow-up, the delivery and the neonatal characteristics. How many women had a hypertension or gestational diabetes? That is a problem because these diseases are prognostic factors which are linked with your endpoints (term birth and birth weight)

The use of means reduces the information and does not help to look at the covariables which may be potentially included in the multilevel analyses. Give the age and parity and BMI in a categorical fashion (age for example: ≤ 18 , >18 and ≤ 35 , >35 ; BMI: <18.5 , ..., >25 ..).

How many ultrasounds per women in the quartiles of table 1? How many hospitalizations?

Give the birth weight and gestational age at delivery in classes in the 4 groups.

Finally, we do not know what was the main outcome and we do not have a posteriori power which should be calculate.

Some covariates are explained (lines 137-142) but are not available in the tables for the readers (complications in previous pregnancy, mode of delivery...). Why did you not adjusted for hypertension for the endpoint SGA?

3.5) Discussion

We do not have in the descriptive tables all the medical risk factors so it is difficult to be sure that the association between neighborhood deprivation and impaired development and adverse pregnancy outcomes remained emphasizing an isolated role for neighborhood deprivation as risk factor for pregnancy.

The chapter should at the beginning give the main results, secondly the limitations and strengths of the study, and then discuss the results with published articles.

I am not surprised by your results. In physiology, we know that embryos and fetus take all the need to growth from the pregnant women. During the first trimester it is, generally, easy for the embryos to have nutriments from the mother but if the mother does not have enough nutriments it is a problem for the fetus.

You explain to the readers that one of the strength of your work is the large number of the women, but we do not have an a posteriori power for a main objective. The limitations of the study that should be discussed in this section are why in a prospective cohort do not you have an anxiety questionnaire at the beginning of pregnancy, a dietary habit questionnaire? An individual risk score is better than a geographic one in term of causal link. In my country, social mobility is usual during a pregnancy. Are the women who did not speak Dutch were excluded of the study or did not participate more often? It does not make sense to think that an early risk assessment could permit to schedule extra ultrasounds and check-up, and be assisted to improve modifiable lifestyle risk factors. The problem for the poor women is to be able to pay good food and a comfortable flat (ie. Primary prevention); and not to improve the medical appointments and ultrasounds (ie. Secondary prevention). In term of public health, it is too late if your actions are only to improve the numbers of medical appointments. However, perhaps in your country the ultrasounds and medical appointments are not free for the pregnant women (we do not know and we do not have the number of ultrasounds and medical appointments in the 4 groups and in your guidelines for the follow-up fir low risk pregnancies).

	<p>3.6) Conclusion</p> <p>It is too long. Some sentences should be in the discussion.</p> <p>3.7) Abstract</p> <p>We do not know what the main outcome was. In the design we have the term prospective cohort study and in the text “population based prospective cohort study”. The outcomes are not well defined. Living in a deprived neighborhood is not an intervention but an exposition. Look at the logical link between the results section and the conclusion; it is strange. The results have to be expressed for the higher neighborhood deprivation score and not for the lowest score.</p> <p>3.8) Minors comments</p> <ul style="list-style-type: none"> - I saw some English errors in the text. - Strengths and limitations of the study: “this study investigates the association between neighborhood deprivation and fetal growth and...” is not a strength of the study. - All abbreviations should have been introduced in the text before their use. - Line 268 (M since maternal....).
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Hillary Brown, University of Toronto at Scarborough Comments to the Author:

This is an interesting study looking at the impact of neighbourhood deprivation on fetal growth and other adverse pregnancy outcomes using data from the Generation R Cohort Study in the Netherlands. The paper is well written and the study well executed. I have the following comments:

We thank Dr. Brown for the positive comment. Furthermore we thank dr. Brown for the elaborate suggestions that have improved our manuscript. Our answers (in bold) are as follows:

Introduction:

1. I thought the Introduction should have included some more information on why and how neighbourhood deprivation could impact pregnancy outcomes above and beyond the effects of individual level SES, etc.

Indeed, we did not address this very important point. Now, we have added information on the possible mechanisms behind neighbourhood deprivation and fetal growth and pregnancy outcomes in the introduction.

Lines 62-69: *Living in a deprived neighbourhood is known to be a risk factor for adverse health outcomes, above and beyond the association with individual risk factors such as inadequate nutrition and lifestyle behaviors. Living in a deprived neighbourhood may lead to exposure to a suboptimal environment, with higher rates of air pollution, less access to facilities such as a green environment to walk in, less health care facilities close by, and little possibility to purchase healthy food nearby. Lastly, living in a deprived neighbourhood is acknowledged as a source of chronic stress, which is associated with increased cortisol levels, and thereby acts as an independent risk factor for adverse health outcomes.*(11,12)

Methods:

1. Please define the acronym CRL (crown rump length) at its first use (page 5).

We have adjusted this accordingly.

2. I understand that ethnicity was measured as a covariate, but was it accounted for in the calculation of fetal growth? See for example:

[https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.jogc.com%2Farticle%2FS1701-2163&data=04%7C01%7Cd.gootjes%40erasmusmc.nl%7C889a71fb51a74edc896c08d937e7c9a8%7C526638ba6af34b0fa532a1a511f4ac80%7C0%7C0%7C637602290948145164%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiIl6lk1haWwiLCJXVCi6Mn0%3D%7C1000&data=bMfrLZSWFiC1ptO17k3CUUn%2BtG7EMX28SeyCtw1ZN4%3D&reserved=0\(16\)35159-3/abstract](https://eur01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.jogc.com%2Farticle%2FS1701-2163&data=04%7C01%7Cd.gootjes%40erasmusmc.nl%7C889a71fb51a74edc896c08d937e7c9a8%7C526638ba6af34b0fa532a1a511f4ac80%7C0%7C0%7C637602290948145164%7CUnknown%7CTWFpbGZsb3d8eyJWljojMC4wLjAwMDAiLCJQIjoiV2luMzliLCJBTiIl6lk1haWwiLCJXVCi6Mn0%3D%7C1000&data=bMfrLZSWFiC1ptO17k3CUUn%2BtG7EMX28SeyCtw1ZN4%3D&reserved=0(16)35159-3/abstract)

The reviewer does address an important point regarding fetal growth, and that ethnicity may impact this outcome. Unfortunately, we were not able to open the link that was provided by dr. Brown. However, the standard deviation scores of fetal growth are calculated within this multi-ethnic cohort, thereby accounting for this factor. However, we couldn't rule out the confounding role of ethnicity, which is why we adjusted for this factors as covariate.

3. Why were preterm birth and SGA the only measures of adverse pregnancy outcomes? Given the label, I expected a broader range of outcomes. A rationale for the choice of outcomes would be useful – e.g., are they most likely to be affected by neighbourhood deprivation vs. other complications such as gestational hypertension, gestational diabetes, etc.

The reviewer points out that the rationale behind the outcome data is not well explained. The aim was to focus on growth and development in the embryo/fetus, and this comment is now added to the manuscript (Lines 116-123) as follows:

'The aim of the study was to focus on fetal outcomes, in terms of growth and development. The selected outcomes were carefully chosen from the 'Big 4 conditions', which are specifically defined conditions that precede perinatal mortality in 85 % of all cases, namely: small for gestational age (birth weight < 10th percentile for gestational age), preterm birth (birth < 37 weeks of gestation), congenital disorders, and/or low Apgar score (<7 after 5 min).(14,15) Due to the low numbers of cases with congenital disorders within the Generation R population, and susceptibility for Apgar score to be affected by the course of delivery which may confound the effect of neighbourhood deprivation during pregnancy, we selected the other 2 major morbidity factors as outcome for this study.'

Moreover, in order to specify the outcome data more clearly, we made adjustments to the title of the manuscript accordingly: 'The association between neighbourhood deprivation, fetal growth, small-for-gestational age and preterm birth: a population-based prospective cohort study.'

4. Why did the covariates include complications in a previous pregnancy and not the current pregnancy?

In sensitivity analyses, we investigated whether a higher risk of adverse birth outcomes in residents from deprived neighbourhoods was confounded by a complication in a previous pregnancy, which is a known strong predictor for adverse outcomes such as preterm birth and SGA birth in a consecutive pregnancy. (An instrument for broadened risk assessment in antenatal health care including non-medical issues. Vos, van Veen, Birnie, Denktaş, Steegers, Bonsel. *Int J Integr Care*. 2015 Mar. doi: 10.5334/ijic.1512.). No adjustments for current complications in

pregnancy were performed, since no differences in outcome were found when adjusting the analyses by complications in a previous pregnancy, and no big differences between the 5 categories of neighbourhood deprivation were found.

5. The authors chose to use a logistic regression model, but interpret the ORs in terms of risk. I understand that most of the outcome rates are < 10%; however, SGA is 12.2% in the exposed group, and some of the other outcomes also hover around 10%. I would recommend not interpreting the OR as indicating risk but rather as odds – or to use a method (e.g., modified Poisson regression, binomial regression, etc) that allows direct estimation of relative risk.

We agree with the reviewer that we need to express the outcomes as odds, instead of in terms of risks. We adjusted this accordingly.

Discussion:

6. The data are from the Generation R Cohort Study which was conducted on births between 2002 and 2006. The most recent data are therefore over 15 years old. Could the authors provide more information on whether the findings are still relevant and how time may or may not affect their interpretation? Further detail on this should be added to the Discussion.

The reviewer is correct that the data on pregnancy and birth is over 15 years old. This is due to the long-term follow-up setting for this prospective cohort study, which is set up with the aim to study pregnancy exposures and long term outcomes. With regard to relevance of the data: embryonic and fetal growth measurement guidelines have not been adjusted since these measurements.

With regard to neighbourhood SES; from detailed analyses of address changes during the perinatal period we know that the great majority of women who move stays within the same (deprivation) class of living area (Reitsma JB, Kardaun JW, Gevers E, de Bruin A, van der Wal J, Bonsel GJ. [Possibilities for anonymous followup studies of patients in Dutch national medical registrations using the Municipal Population Register: a pilot study]. *Ned Tijdschr Geneeskd* 2003;147:2286–2290.) <https://bronnen.zorggegevens.nl/Bron?naam= Sociaal-Economische-Status-per-postcodegebied>)

Moreover, the degree of deprivation is fairly stable over time. This point is addressed in lines 299-303: *Next, possible misclassification of neighbourhood deprivation may have occurred if women moved during pregnancy to a neighbourhood with a different status score from the one they moved out of. However, social mobility in pregnancy is limited and if women move, they generally tend to move to a neighbourhood with a comparable deprivation status.*(32)

7. I was surprised that the focus of the implications appeared to be on targeting health behaviours of residents of deprived neighbourhoods. If the effect of neighbourhood deprivation remained after controlling for (some of) these individual level factors, I think an important implication relates to population-level interventions such as policies that promote healthier environments (e.g., more walkable neighbourhoods, universal basic income, etc) – not interventions that rely on individual behavior change. Could the authors address this?

Indeed, the author is right. In the discussion we pointed out population-level interventions, but this wasn't stated clearly enough. We made adjustments accordingly.

Line 360-367 *'Although the magnitude of our findings is somewhat small, the results of this study suggest an isolated risk for living in a deprived neighbourhood. This emphasizes the importance of policies that promote healthier neighbourhoods. This could be achieved by targeted*

population-level interventions. A review has demonstrated many successful area-based initiatives that have been implemented in deprived areas across Western-Europe already.(42) Initiatives may consist of interventions that aim to tackle the various problems in deprived areas, with regard to the *psychical (more walkable neighbourhoods, increasing green environments, reducing air pollution and the reduction of litter.) and social domain (lowering crime rates, vandalism).(43)*

8. The authors may also want to emphasize further that the magnitude of the findings is somewhat small (though statistically significant). What impact does this have on interpretation in terms of expected effect of interventions on neighbourhood deprivation?

We agree that though the findings are statistically significant, the effect sizes are small, which is important to address. On an individual level, interventions may only result in small effects. However on a population level, small individual level changes may result in large wins for the total population. This is added to the manuscript.

Line 367-370 ‘Small effects of these interventions may be expected in terms of differences in fetal growth and birthweight. However even small individual effects may still have clinical and public health relevance, e.g. when they affect a large segment of the population, or when a small effect has long term implications, as is the case with birthweight.’

Figures and tables:

9. Please add the n (%) with each outcome to the tables for the different exposure groups to give a sense of the public health importance of the finding – e.g., in Table 2.

We made adjusted accordingly, and added the n(%) in the Table.

Reviewer: 2

Dr. FRANÇOISE VENDITTELLI, Academic Hospital of Clermont-Ferrand Comments to the Author:
Reviewer's comments on manuscript bmjopen-2021-049075

Comments to the authors:

1) Synthesis of the article

The article assesses the association between neighborhood deprivation and fetal growth during the first second and third trimester of pregnancy, and adverse pregnancy outcomes. The authors did not find an association between neighborhood deprivation and the first trimester growth. The found a negative association between neighborhood deprivation and the fetal growth in the second and third trimester of pregnancy, and with a risk of SGA and preterm birth.

2) General comments

Thank you for your efforts to publish your work but this article needs several clarifications to improve its reading and internal and external validity.

We thank the reviewer for the extensive advice, recommendations and questions. Our answers (in bold) are as follows:

3) Detailed comments

3.1) Title

It should include the study design: “neighborhood deprivation, fetal growth and adverse pregnancy outcomes: a population based cohort study; the Generation R Study. Could you explain what is the signification of “R” of Generation R Study?

Generation ‘R’ refers to ‘Rotterdam’; this prospective cohort study is based in the city of Rotterdam. (Line 75)

3.2) Introduction

The justification of this study is not well explained. We know that low economic status is associated with adverse pregnancy outcomes. We have several publications even during wars. Why did you realize a new study as we have yet some good studies on the topic? You explained to the readers "It is however unknown whether this potentially modifiable factor is also associated with an early fetal development". Could you give to the readers some animal or physiological studies to support a research hypothesis? The objectives are not well introduced at the end of the introduction "therefore..." and we do not know what the main objective is and what the secondary objectives are.

Thank you for this point. Indeed, we know that very low SES or extreme hunger are associated with adverse pregnancy outcomes. From different animal and human studies we know that embryonic development is already sensitive to external factors, such as maternal nutrition and smoking habits.

(Nutritional effects on oocyte and embryo development in mammals: implications for reproductive efficiency and environmental sustainability. C.J. Ashworth, L.M. Toma, M.G. Hunter. Philos Trans R Soc Lond B Biol Sci, 364 (2009), pp. 3351-3361 / Role of micronutrients in the periconceptional period. Cetin, C. Berti, S. Calabrese. Hum Reprod Update, 16 (2010), pp. 80-95 /Maternal Lifestyle Impairs Embryonic Growth: The Rotterdam Periconception Cohort. Matthijs R Van Dijk 1, Nicole V Borggrevén 1, Sten P Willemsen 1 2, Anton H J Koning 3, Régine P M Steegers-Theunissen 1 4, Maria P H Koster 1. PMID: 28884629 DOI: 10.1177/1933719117728801. Reprod Sci 2018 Jun)

Therefore, we wondered whether this is also the case for an external factor such as neighbourhood deprivation. We added more background information and hypothesis on why and how neighbourhood deprivation may affect embryonic and fetal development.

Line 53 'Recent evidence shows that other factors, such as maternal nutrition and lifestyle, already affect pregnancy from the first trimester of pregnancy onwards.(6-9) Gaining a better understanding of modifiable factors that influence pregnancy from the earliest phase onwards is important.'

lines 62-69 'Living in a deprived neighbourhood is known to be a risk factor for adverse health outcomes, above and beyond the association with individual risk factors such as inadequate nutrition and lifestyle behaviors. Living in a deprived neighbourhood may lead to exposure to a suboptimal environment, with higher rates of air pollution, less access to facilities such as a green environment to walk in, less health care facilities close, and little possibility to purchase healthy food nearby. Lastly, living in a deprived neighbourhood is acknowledged as a source of chronic stress, which is associated with increased cortisol levels, and thereby acts as an independent risk factor for adverse health outcomes.(11, 12).'

3.3) Methods

- Describe the materials section in a separate chapter of the study design.

We made changes accordingly.

- Materials: Why did you exclude the spontaneous and late abortion, and the intra-uterine deaths? They are part of adverse pregnancy outcomes and are pertinent variables for the research topic.

We agree with the reviewer that abortions (n=29) and fetal deaths (n=75) are relevant outcomes, however the number of these outcomes are low and therefore less suitable as outcome measurements. Additionally, in order to study our outcomes related to fetal growth, we wished to exclude these since the possible underlying causes for abortion or fetal death may also affect

fetal growth patterns. Moreover, our aim was to investigate growth trajectories and outcomes of ongoing singleton pregnancies.

Could you specify the years of the studied births. What was the beginning of the enrollments and the end of the follow-up?

Mothers with a delivery date between April 2002 and January 2006 were eligible, enrollment took place between these data. The 9778 mothers enrolled in the study gave birth to 9749 live born children. Of these mothers, 91% (n=8879) were prenatally enrolled in the study, giving birth to 8976 children. For the period of investigation during the adolescence period (12-16 years), 7968 children participated. The follow-up is still ongoing, with children being up to 19 years old already.

Did you exclude the in utero fetal anomalies?

Yes: with regard to congenital malformations, we have reliable information in pregnancies. In Table A, we reran the analyses excluding all pregnancies complicated by congenital malformations. Results did not materially change from those in the manuscript.

Did the women should speak Dutch?

No, the women did not have to speak Dutch. There is support for non-Dutch speaking participants: all study materials such as questionnaires, newsletters, website, and information folders are available in three languages (Dutch, English, and Turkish). Furthermore, staff from different ethnic backgrounds is available and verbally translate these materials into Arabic, French and Portuguese. As such, the study staff is able to communicate with all participants, allowing the selection of a multi-ethnic population.

- Patient and public involvement statement: you are in contradiction with line 363.

This is now adjusted in the newly submitted version of the manuscript.

- Neighborhood deprivation: for me it is the description of the exposition status. This is a geographic score, not an individual score. Could you define what are in your country a low income and a low educational level. What is the black work in your country (if you have a high level of black work you may have an information bias, at minimum this point has to be discuss in the discussion chapter). Your data are old, why? The recent economic crisis may have increased the proportion of women living in a neighborhood deprivation context.

Low income or minimum wage is set on a sum of € 1.701,07 of a 2-person household.

A low educational level includes the level of primary education, or the first years of high school (without Advanced Placement).

It has been estimated that almost 13% of Dutch residents do any form of black work/undeclared work, which indeed might have introduced a form of bias. (<https://nos.nl/op3/artikel/2303517-honderdduizenden-nederlanders-werken-zwart-schoonmakers-koploper/> Winden de, P., Mol, M. en Zuurbier, H. (2011), Een onderzoek naar zwartwerk onder uitkeringsgerechtigden en de totale bevolking in 2010. CBS, Voorburg/Heerlen.) It is possible that this is more common in deprived or low income neighbourhoods, however the distribution of black work/undeclared work over the different types of neighbourhood is unknown. We now addressed this point in the discussion section of the manuscript. (Lines 303-305)

Indeed, the recent COVID crisis with its economic consequences might have put more families into (more) financial troubles. Since we observe a (small) effect of living in a deprived neighbourhood on the course and outcome of pregnancy, repetition of this work within a new setting after the crisis would be interesting.

- Covariables:

• I do not understand the rationale of your choice to work on the last menstrual period. Even with a regular cycle the best way to have the last calculated menstruation is to have a good first trimester ultrasound. If you think that the first growth is not uniform among women or among ethnicity groups then you will not control this fact by using last menstrual period. This is a big weakness of this work.

If I understand, the pregnancy dating process was not the same among all the women included in the study (line 105-106) which is also an epidemiological problem.

For the large majority of the study population, pregnancy dating was performed uniformly: dating performed using the first ultrasound measurement of either the CRL (if the gestational age was below 12 weeks and 5 days and CRL measurement was smaller than 65 mm), or the biparietal diameter (BPD) (from a gestational age from 12 weeks and 5 days onwards and with a BPD larger than 23 mm) according to standard guidelines. However, for the analyses with CRL as outcome measure, dating cannot be performed with the BPD since this is measured in later phases of pregnancy, and CRL is not suitable since this is the outcome measure. Then, in women with a regular menstrual cycle and known last menstrual period, dating according to this last menstrual period is the most reliable method available.

New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. B O Verburg 1, E A P Steegers, M De Ridder, R J M Snijders, E Smith, A Hofman, H A Moll, V W V Jaddoe, J C M Witteman. Ultrasound Obstet Gynecol. 2008 Apr;31(4):388-96. doi: 10.1002/uog.5225.

• Growth parameters: We need to know if you made an audit to assess the quality of the fetal measurements according to your national guidelines. If the clinicians were each responsible of the ultrasounds assessment, even in a research center, you may have an information bias.

We agree with the reviewer that adequate assessment of quality is important. Reproducibility of fetal growth parameters was assessed and described previously.

Verburg BO, Mulder PG, Hofman A, Jaddoe VW, Witteman JC, Steegers EA. Intra- and interobserver reproducibility study of early fetal growth parameters. Prenat Diagn. 2008;28(4):323-31.

Romy Gaillard, Maria A J de Ridder, Bero O Verburg, Jacqueline C M Witteman, Johan P Mackenbach, Henriëtte A Moll, Albert Hofman, Eric A P Steegers, Vincent W V Jaddoe. Individually customised fetal weight charts derived from ultrasound measurements: the Generation R Study Eur J Epidemiol. 2011 Dec;26(12):919-26. doi: 10.1007/s10654-011-9629-7. Epub 2011 Nov 15. Added to Line 135-136: 'Reproducibility of these measurements was assessed and described previously.(22, 23)'

• Adverse pregnancy outcomes: I do not understand why you did not assess other adverse pregnancy outcomes which are well known to be associated with individual low incomes or neighborhood deprivation (Large for gestational age, gestational diabetes).

Indeed, the reviewer does address a relevant question. However, the aim of the study was to keep the focus on growth and development. Moreover, the two selected outcomes were carefully chosen from the 'Big 4 conditions', which are specific defined conditions that precede perinatal mortality in 85 % of all cases, namely: small for gestational age (birth weight < 10th percentile for gestational age), preterm birth (birth < 37 weeks of gestation), congenital disorders, and/or low Apgar score (<7 after 5 min).

Due to the low number of congenital disorders within the Generation R population, and susceptibility for Apgar score to be affected by the course of delivery which may confound the effect of neighbourhood deprivation during pregnancy, we selected the other 2 major morbidity factors as outcome for this study.

Moreover, within Generation R, maternal SES has been investigated in relation to pregnancy outcomes such as for example hypertensive disorders: *Maternal educational level and risk of gestational hypertension: the Generation R Study. L M Silva, M Coolman, E A P Steegers, V W V Jaddoe, H A Moll, A Hofman, J P Mackenbach & H Raat. Journal of Human Hypertension volume 22, pages483–492 (2008)*

This was added to lines 116-120: The aim of the study was to focus on fetal outcomes, in terms of growth and development. The selected outcomes were carefully chosen from the 'Big 4 conditions', which are specifically defined conditions that precede perinatal mortality in 85 % of all cases, namely: small for gestational age (birth weight < 10th percentile for gestational age), preterm birth (birth < 37 weeks of gestation), congenital disorders, and/or low Apgar score (<7 after 5 min).(17, 18)

Why did you choose the 10th percentile like endpoint for SGA? We have an international definition of the small-for-gestational age advisory board consensus statement of 2001 which is not the 10th [Lee PA, Chernausek SD, Hokken-Koelega ACS, Czernichow P for the International SGA Advisory Board: *International Small for Gestational Age Advisory Board consensus development conference statement: management of short children born small for gestational age, April 24-October 1, 2001. Pediatrics 2003, 111: 1253-61*].

Indeed, we acknowledge that there are multiple methods to define SGA. Already for SGA within a 10 percentile range, there are adverse neonatal effects. (McCowan LM, Figueras F, Anderson NH. *Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. Am J Obstet Gynecol. 2018 Feb;218(2S):S855-S868.*

Ray JG, Park AL, Fell DB. Mortality in Infants Affected by Preterm Birth and Severe Small-for-Gestational Age Birth Weight. Pediatrics. 2017 Dec;140(6)

Malin GL, Morris RK, Riley R, Teune MJ, Khan KS. When is birthweight at term abnormally low? A systematic review and meta-analysis of the association and predictive ability of current birthweight standards for neonatal outcomes. BJOG. 2014 Apr;121(5):515-26)

Next, we choose the definition of the 10th percentile for comparison purposes with other Generation R studies. Moreover, the 10th percentile is more frequently used in literature; this meta-analysis identified 26 commonly cited reference charts which include the 10th percentile cut point to define SGA status: *Katz J., Wu L.A., Mullany L.C., Coles C.L., Lee A.C., Kozuki N. Prevalence of small-for-gestational-age and its mortality risk varies by choice of birth-weight-for-gestation reference population. PLoS ONE. 2014;9:e92074.*

FGR is different of SGA (Peter et al in 2003 (international small for gestational age advisory consensus development conference statement.. » SGA refers not to fetal growth but to the size of infant at birth. The term intra uterine growth retardation (IUGR) suggests a diminished growth velocity in the fetus. You have a prospective cohort and so you should be able to classify FGR and SGA. I also think that, maybe, more the social deprivation is more the SGA is severe. It could be interesting to have this information in the descriptive tables.

Indeed, the reviewer is right that SGA refers to an endpoint-like outcome, compared to FGR or IUGR. For this manuscript, we chose to model fetal growth trajectories and investigate the association between neighbourhood deprivation and these trajectories, and as a hard endpoint, to investigate the association with SGA. With our approach, we investigate both the association between neighbourhood deprivation and fetal growth as a dynamic parameter, and both SGA as a static parameter of fetal growth. Through this approach, we did not have to limit the data of FGR as a deflection of growth of more than 20%.

• Clarify the chapter: all the endpoints in a same chapter (growth parameters and selected pregnancy outcomes). **We have made adjustments accordingly.**

- We have a few explanations about the questionnaires. What was the language of the questionnaires? **All study materials such as questionnaires, newsletters, website, and information folders are available in three languages (Dutch, English, and Turkish). Furthermore, staff from different ethnic backgrounds is available and verbally translate these materials into Arabic, French and Portuguese. As such, the study staff is able to communicate with all participants. (lines 142-145)**

3.4) Results

We have only one descriptive table “baseline characteristics” but we do not have a descriptive table for the pregnancy follow-up, the delivery and the neonatal characteristics. How many women had a hypertension or gestational diabetes? That is a problem because these diseases are prognostic factors which are linked with your endpoints (term birth and birth weight) The use of means reduces the information and does not help to look at the covariables which may be potentially included in the multilevel analyses. Give the age and parity and BMI in a categorical fashion (age for example: ≤ 18 , > 18 and ≤ 35 , > 35 ; BMI: < 18.5 , ..., > 25 ..).

How many ultrasounds per women in the quartiles of table 1? How many hospitalizations?

Give the birth weight and gestational age at delivery in classes in the 4 groups.

As suggested, we added categories for age and BMI in Table 1, as well as the outcomes maternal hypertension, gestational diabetes, birth weight and gestational age at delivery. With regard to hospitalizations we do not have the accurate data available.

Finally, we do not know what was the main outcome and we do not have a posteriori power which should be calculate.

Our main outcome was the fetal growth, in terms of head circumference, length and weight. Post-hoc power for 0.1 SD difference in fetal growth with an alpha of 0.05 for a study group of 8000 (this study population 8617) participants is 99.4%. This is added to the manuscript’s description of the statistical analyses (lines 187-190).

Some covariates are explained (lines 137-142) but are not available in the tables for the readers (complications in previous pregnancy, mode of delivery...).

We have now added covariates to Table 1.

Why did you not adjusted for hypertension for the endpoint SGA?

We thank the reviewer for this relevant question. We added sensitivity analyses in which we adjust for maternal hypertension (Table B). This did not materially change the outcome of the analyses.

3.5) Discussion

We do not have in the descriptive tables all the medical risk factors so it is difficult to be sure that the association between neighborhood deprivation and impaired development and adverse pregnancy outcomes remained emphasizing an isolated role for neighborhood deprivation as risk factor for pregnancy.

As suggested, Table 1 is now extended with more information such as hypertensive disorders of pregnancy, birthweight and gestational age at birth.

The chapter should at the beginning give the main results, secondly the limitations and strengths of the study, and then discuss the results with published articles.

We have adapted this according to the suggested order.

I am not surprised by your results. In physiology, we know that embryos and fetus take all the need to growth from the pregnant women. During the first trimester it is, generally, easy for the embryos to have nutriments from the mother but if the mother does not have enough nutriments it is a problem for the fetus.

Indeed, in objective amounts, the need for nutrients may be less in the first trimester. However, from other studies we know that embryonic growth is already susceptible to changes in its environment. This is why we took this phase of pregnancy already into account as outcome variable.

You explain to the readers that one of the strength of your work is the large number of the women, but we do not have an a posteriori power for a main objective.

Indeed, as discussed under point 3.4, we have now addressed this point.

The limitations of the study that should be discussed in this section are why in a prospective cohort do not you have an anxiety questionnaire at the beginning of pregnancy, a dietary habit questionnaire?

Within the Generation R study there is information on anxiety and depression. However, research on the mechanisms underlying the association between neighbourhood deprivation and fetal growth and pregnancy outcomes is beyond the scope of this study. However for future studies, this could provide us with the data to investigate this. For now, the first step was to investigate the association of neighbourhood deprivation as a possible composite measure for different environment(al) factors.

Second, indeed within the Generation R study, there is information on a semi quantitative self-administrated food frequency questionnaires, obtained at study enrollment. However, this FFQ is only validated in pregnant women with a Dutch ethnic background, which would have diminished the power and external validity of the study. Thereby, we adjusted for alcohol use and smoking. There is a high correlation between smoking and alcohol use and dietary habits.

Heppe DH, Medina-Gomez C, Hofman A, Franco OH, Rivadeneira F, Jaddoe VW. Maternal first-trimester diet and childhood bone mass: the Generation R Study. Am J Clin Nutr. 2013;98:224–232. 18.

Voortman T, Steegers-Theunissen RPM, Bergen NE, Jaddoe VWV, Looman CWN, Kiefte-de Jong JC, Schalekamp-Timmermans S. Validation of a semi-quantitative food-frequency questionnaire for Dutch pregnant women from the general population using the method or triads. Nutrients. 2020;12:1341.

An individual risk score is better than a geographic one in term of causal link.

An individual risk score does provide a good measuring tool on the degree of individual deprivation. However, our aim was to investigate the effect of geographical/neighbourhood

deprivation, namely the possible (composite) and/or (indirect) effect of air pollution, stress and green space availability . In order to enable extensive adjustment, we did not make use of an individual risk score, but added the factors independently to the models.

In my country, social mobility is usual during a pregnancy.

Indeed the place of residence was recorded at the time of delivery; in case women moved during pregnancy misclassification of neighborhood deprivation may have occurred. From detailed analyses of address changes during the perinatal period we know that the great majority of women who move stays within the same (deprivation) class of living area. (Reitsma JB, Kardaun JW, Gevers E, de Bruin A, van der Wal J, Bonsel GJ. [Possibilities for anonymous follow-up studies of patients in Dutch national medical registrations using the Municipal Population Register: a pilot study]. *Ned Tijdschr Geneesk* 2003;147:2286–2290.)

Are the women who did not speak Dutch were excluded of the study or did not participate more often?

No, they were not excluded. All study materials such as questionnaires, newsletters, website, and information folders are available in three languages (Dutch, English, and Turkish). Furthermore, staff from different ethnic backgrounds is available and verbally translate these materials into Arabic, French and Portuguese. As such, the study staff is able to communicate with all participants.

It does not make sense to think that an early risk assessment could permit to schedule extra ultrasounds and check-up, and be assisted to improve modifiable lifestyle risk factors. The problem for the poor women is to be able to pay good food and a comfortable flat (ie. Primary prevention); and not to improve the medical appointments and ultrasounds (ie. Secondary prevention). In term of public health, it is too late if your actions are only to improve the numbers of medical appointments. However, perhaps in your country the ultrasounds and medical appointments are not free for the pregnant women (we do not know and we do not have the number of ultrasounds and medical appointments in the 4 groups and in your guidelines for the follow-up for low risk pregnancies).

Indeed, we did not emphasize the importance of population-level interventions enough. For women with an insurance, medical appointments are free. However, also for residents with an illegal immigrant status, obstetric care is always accessible.

Now, we've made adjustments accordingly.:

Line 350-370 *'Although the magnitude of our findings is somewhat small, the results of this study suggest an isolated risk for living in a deprived neighbourhood. This emphasizes the importance of policies that promote healthier neighbourhoods. This could be achieved by targeted population-level interventions. A review has demonstrated many area-based initiatives that have been implemented in deprived areas across Western-Europe already.(42) Initiatives may consist of interventions that aim to tackle the various problems in deprived areas, with regard to the psychical (more walkable neighbourhoods, increasing green environments, reducing air pollution and the reduction of litter) and social domain (lowering crime rates, vandalism).(43) Small effects of these interventions may be expected in terms of differences in fetal growth and birthweight. Though some very small individual effects may still have clinical and public health relevance, e.g. when they affect a large segment of the population, or when a small effect has long term implications, as is the case with birthweight.*

However, we also do stress the importance of early risk assessment and strict and good quality medical check-ups.

3.6) Conclusion

It is too long. Some sentences should be in the discussion.

Indeed, a conclusion section was not inserted yet. Now this is added to the manuscript.
Lines 414-417: In conclusion, our obtained insights on the association between neighbourhood deprivation and fetal growth and prematurity emphasize the need for a comprehensive research, care and policy approach from the preconception phase onwards, to mitigate the risk of adverse pregnancy outcomes due to deprivation.

3.7) Abstract

We do not know what the main outcome was.

The main outcome was the fetal growth trajectories of head circumference, length and weight. This is changed in the abstract.

In the design we have the term prospective cohort study and in the text “population based prospective cohort study”.

This has now been adjusted to a “population-based prospective cohort study”.

The outcomes are not well defined.

Now, we have defined them more specifically. (lines 28-29) : Main outcome measures: Fetal growth trajectories of head circumference, weight and length. Secondary outcomes measures: small-for-gestational age (SGA) and preterm birth (PTB).

Living in a deprived neighborhood is not an intervention but an exposition.

We have made changes accordingly; line 27: Exposition: Living in a deprived neighbourhood.

Look at the logical link between the results section and the conclusion; it is strange. The results have to be expressed for the higher neighborhood deprivation score and not for the lowest score.

Indeed, since we worked with a continuous ascending definition of neighbourhood deprivation, statistically neighbourhoods with lower scores are set as reference, and better results or positive associations with higher scores indicate worse or negative associations with lower scores. Results are again expressed for the higher neighborhood deprivation score.

3.8) Minors comments

- I saw some English errors in the text.

- Strengths and limitations of the study: “this study investigates the association between neighborhood deprivation and fetal growth and...” is not a strength of the study. **This is deleted.**

- All abbreviations should have been introduced in the text before their use. **Now, all newly introduced abbreviations have been explained.**

- Line 268 (M since maternal....). **This is changed accordingly.**

Table A. Associations between the neighbourhood status score and adverse pregnancy outcomes, excluding confirmed congenital disorders.

Study population	Model	Lowest deprivation on quartile	Second deprivation on quartile	Third deprivation on quartile	Highest deprivation on quartile	Trend	p-value for
------------------	-------	--------------------------------	--------------------------------	-------------------------------	---------------------------------	-------	-------------

n = 5184		n = 1190		n = 1068		n = 1273		n = 1653		trend	
		β / OR (95% CI)		β / OR (95% CI)		β / OR (95% CI)		β / OR (95% CI)			
Small for gestational age	Basic	14 (12.0%)	2.02 (1.55 ; 2.62)	96 (9.0%)	1.46 (1.09 ; 1.94)	10 (8.6%)	1.38 (1.05 ; 1.83)	10 (6.4%)	Reference	0.84 (0.79 ; 0.90)	<0.001
	Adjusted		1.53 (1.12 ; 2.07)		1.13 (0.83 ; 1.54)		1.23 (0.92 ; 1.64)		Reference	0.91 (0.84 ; 0.98)	0.01
Preterm birth	Basic	44 (3.7%)	1.40 (0.92 ; 2.15)	46 (4.3%)	1.65 (1.08 ; 2.51)	48 (3.8%)	1.43 (0.95 ; 2.17)	44 (2.7%)	Reference	0.90 (0.81 ; 0.99)	0.03
	Adjusted		1.43 (0.89 ; 2.31)		1.59 (1.02 ; 2.48)		1.40 (0.92 ; 2.13)		Reference	0.88 (0.78 ; 0.99)	0.04

Abbreviations: β : beta; OR: odds ratio. Values are odds ratios with the 95% CI of the data in SD-score and are based on logistic regression models. Basic model: by the use of SD scores it is automatically adjusted for gestational age. Adjusted model: basic model and additionally adjusted for maternal age, educational level, smoking, alcohol use, folic acid supplement use, ethnicity, parity, pre-pregnancy body mass index and fetal sex. p-for trend analysis with the neighbourhood deprivation as a continuous measure. Small size for gestational age (SGA) at birth was defined as a sex and gestational age adjusted birthweight below the 10th percentile (<-1.40 SD-score) in the study cohort. Preterm birth (PTB) was defined as a gestational age of <37 weeks at delivery.

Table B. Associations between the neighbourhood status score and SGA pregnancies, adjusted for maternal hypertension.

Study population	Model	Lowest deprivation quartile n = 1190	Second deprivation quartile n = 1068	Third deprivation quartile n = 1273	Highest deprivation quartile n = 1653	Trend	p-value for trend
		β / OR (95% CI)	β / OR (95% CI)	β / OR (95% CI)			β / OR (95% CI)

Small for gestational age	Adjusted	1.39 (1.09 ; 1.77)	0.0 1	1.15 (0.91 ; 1.46)	0.24	1.13 (0.89 ; 1.43)	0.31	<i>Reference</i>	0.91 (0.85 ; 0.97)	0.00 4
----------------------------------	----------	-------------------------------------	------------------------	-----------------------	------	-----------------------	------	------------------	-------------------------------------	-------------------------

Abbreviations: β : beta; OR: odds ratio. Values are odds ratios with the 95% CI of the data in SD-score and are based on logistic regression models. Basic model: by the use of SD scores it is automatically adjusted for gestational age. Adjusted model: basic model and additionally adjusted for maternal age, educational level, smoking, alcohol use, folic acid supplement use, ethnicity, parity, pre-pregnancy body mass index, fetal sex, and additionally for maternal hypertension.. p-for trend analysis with the neighbourhood deprivation as a continuous measure. Small size for gestational age (SGA) at birth was defined as a sex and gestational age adjusted birthweight below the 10th percentile (<-1.40 SD-score) in the study cohort. Preterm birth (PTB) was defined as a gestational age of <37 weeks at delivery.

VERSION 2 – REVIEW

REVIEWER	Brown, Hillary University of Toronto at Scarborough, Health & Society
REVIEW RETURNED	13-Oct-2021

GENERAL COMMENTS	I thank the authors for their careful revisions based on the previous reviews. I have no further suggestions or concerns.
-------------------------	---

REVIEWER	VENDITTELLI, FRANÇOISE Academic Hospital of Clermont-Ferrand, Gynécologie Obstétrique
REVIEW RETURNED	16-Oct-2021

GENERAL COMMENTS	I thank the authors for the quality of the revision of their manuscript. The article is now very clear for non-Dutch people.
-------------------------	--