PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Social inequality and infant health in the United Kingdom: systematic review and meta-analyses
AUTHORS	Weightman AL, Morgan HE, Shepherd MA, Kitcher H, Roberts C and Dunstan F

VERSION 1 - REVIEW

REVIEWER	David Field, University of Leicester
REVIEW RETURNED	16/02/2012

RESULTS & CONCLUSIONS	I have covered many of these points in my report below - it is the lack of contemporary clinical context which I think is missing. I would like to see a more insightful interpretation of the findings.
GENERAL COMMENTS	Thank you very much for asking me to review this paper which describes a careful systematic review and meta analysis of studies looking at social inequality in relation to various aspects of infant health in the UK. The overall approach seems to have been very thorough and a very consistent theme is identified. I also agree with the authors that what is needed now are more studies with individual level data. However there are a number of topics touched on by the authors which I feel would benefit from a more in depth review of how the findings relate to current practice / provision of service. For example (in no particular order): 1) The section on SIDS describes this as a common cause of death in infancy. For at least 30 years it has been secondary to prematurity and congenital abnormalities but in the last few years following the back to sleep campaign it is relatively rare. Although I agree it still shows a strong social gradient it contributes little to the overall difference in infant mortality between social groups. This context is missing. 2) Some studies have used babies either ≤ 32 (or ≤ 31) weeks and others ≤ 37 weeks. The rationale for this in terms of the different aetio pathogenesis between the very preterm and the late and moderate preterm is not made sufficiently clear for the reader to interpret why the results might be different. 3) The introduction makes reference to the UK's poor performance in international comparison but makes no reference to the relevance of differences in definition. Similarly in the discussion the authors refer to a durable effect of deprivation from the 1980s but do not make reference to the important definitional change that occurred during the period (in terms of stillbirth in 1992) and the extent to which the authors tried to explore its impact (did it increase or decrease the effect of inequality or was it impossible to say). 4) Should anyone still be using low birth weight as an outcome in a developed country? 5) How could increased breast feeding initiatives deal with the inequalities seen in infant mortality when these figures are

	dominated by prematurity and congenital anomalies (even though I realise it was Government policy in England for a while)?

REVIEWER	Chris Power Professor of Epidemiology and Public Health UCL Institute of Child Health UK
REVIEW RETURNED	26/02/2012

THE STUDY	The study design was described clearly, but in a few instances a little more detail on justification would be helpful. For example: - criteria for study exclusion are given: prospective cohort, case-control and retrospective cohort studies with a sample size of 200+, and record linkage analyses of routinely collected data were all included; case studies and cross sectional surveys were excluded. But it isn't clear why cross sectional surveys were excluded. - for outcomes such as birth-weight, were there studies of continuous measures and if so, is there a rationale to focus only on low weight?
RESULTS & CONCLUSIONS	On the whole the study is well-presented. However, there are some important concerns regarding interpretation of the study findings, examples as follows: - the conclusion for failure to thrive is that 'there is no clear link' with social disadvantage (p19) based on 3 study populations (Alspac, Gateshead and 'another'). The evidence base therefore contrasts with many other outcomes that are often from large studies with wide geographical coverage. Given the limited number and regional base of studies to date for failure to thrive, a more cautious conclusion is probably justified. - research is scant for some outcomes (e.g wheeze, diarrhoea) and yet a main conclusion of the study (p18) is that 'it can be strongly argued that no more epidemiological research needs to be carried out in the UK to address this general effect of area and individual measures of social deprivation. Further research should seek to explore the factors that are more proximal to material and infant health and may help to throw light on the most appropriate times to provide support and the form(s) that such support should take.' This conclusion is difficult to justify given the weak evidence base for some outcomes, ongoing need to monitor trends and understand the causes of social inequalities. - past literature on social inequality has debated whether there are 'area' effects separate from 'individual' effects. The manuscript by Weightman and colleagues includes both area- and individual-based social indicators yet their discussion is ambiguous on the interpretation of different indicators, i.e. whether they represent different or the same phenomena. - the discussion emphasises proximal factors, such as 'stress' as underlying explanations for social inequalities, with a firm recommendation to include measures of maternal and infant stress levels in future studies (p20). Yet the literature on the intergenerational reproduction of inequalities is relevant in highlighting the importance of events/ exposures in previous generations e.g. for birthweight (see for example, work by Emmanuel). Arguably, intergenerational factors might also be considered in the interpretation of study findings. - some discussion of heterogeneity in study findings might be worthwhile. If area is important and studies are from different areas might we expect to see some heterogeneity?

	Minor points: - some explanation of what is meant by 'calculated at the somewhat smaller lower super output area' (p11, lines 9-10) would be helpful. - please provide a definition of IMD

VERSION 1 – AUTHOR RESPONSE

Reviewer: David Field
University of Leicester, Health Sciences

Thank you very much for asking me to review this paper which describes a careful systematic review and meta analysis of studies looking at social inequality in relation to various aspects of infant health in the UK.

The overall approach seems to have been very thorough and a very consistent theme is identified. I also agree with the authors that what is needed now are more studies with individual level data. However there are a number of topics touched on by the authors which I feel would benefit from a more in depth review of how the findings relate to current practice / provision of service. For example (in no particular order):

1) The section on SIDS describes this as a common cause of death in infancy. For at least 30 years it has been secondary to prematurity and congenital abnormalities but in the last few years following the back to sleep campaign it is relatively rare. Although I agree it still shows a strong social gradient it contributes little to the overall difference in infant mortality between social groups. This context is missing.

<< It is true that the incidence of SIDS has decreased and it could be argued that it is less important than other causes of infant mortality. However the decrease is not dramatic. In Wales in 1999-2003 (Paranjothy unpublished) there were 109; this reduced to 93 in 2004-8 and so it is still an important issue, representing about 30% of post neonatal deaths. We have added a comment to the text (p.15).

2) Some studies have used babies either ≤ 32 (or ≤ 31) weeks and others ≤ 37 weeks. The rationale for this in terms of the different aetio pathogenesis between the very preterm and the late and moderate preterm is not made sufficiently clear for the reader to interpret why the results might be different.

<< In the context of preterm birth the studies we found used a variety of different categories of preterm and this may have been a cause of the heterogeneity that was found there. The fact that the odds ratios were broadly comparable, whether studies used <37 weeks or <33 weeks, was interesting. Comments in the text have been added (pp.9,10,18).

3) The introduction makes reference to the UK's poor performance in international comparison but makes no reference to the relevance of differences in definition. Similarly in the discussion the authors refer to a durable effect of deprivation from the 1980s but do not make reference to the important definitional change that occurred during the period (in terms of stillbirth in 1992) and the extent to which the authors tried to explore its impact (did it increase or decrease the effect of inequality or was it impossible to say).

<< All our studies on stillbirths occurred after 1992, including the data collection, and so the change in definition did not impact on these results. This has been clarified in the text (p.10).

4) Should anyone still be using low birth weight as an outcome in a developed country?

<< Possibly not – but they do! We have reported the outcomes as defined by the authors of each included study.

5) How could increased breast feeding initiatives deal with the inequalities seen in infant mortality when these figures are dominated by prematurity and congenital anomalies (even though I realise it was Government policy in England for a while)?

<< The point is well made and the text has been changed to remove references to breast feeding initiatives (p.18).

Reviewer: Chris Power
Professor of Epidemiology and Public Health
UCL Institute of Child Health
UK

The study design was described clearly, but in a few instances a little more detail on justification would be helpful.
For example:
- criteria for study exclusion are given: prospective cohort, case-control and retrospective cohort studies with a sample size of 200+, and record linkage analyses of routinely collected data were all included; case studies and cross sectional surveys were excluded. But it isn't clear why cross sectional surveys were excluded.

<< A 'best evidence' approach was adopted by using data from longitudinal and record linkage studies covering an extended period of time. Cross sectional studies providing data from a single time point only were excluded in favour of these more robust designs. This has been clarified in the methodology section (p.3).

- for outcomes such as birth-weight, were there studies of continuous measures and if so, is there a rationale to focus only on low weight?

<< All the studies we saw used low birth weight (or very low), rather than comparing means. We have clarified in the text that none of the studies included continuous measures (p.7).
A possible explanation for this is that it is hard to interpret comparisons of means as a higher mean might be a good outcome or it might mean that there was an excess of high birth weight, which is not such a good outcome.

On the whole the study is well-presented. However, there are some important concerns regarding interpretation of the study findings, examples as follows:

- the conclusion for failure to thrive is that 'there is no clear link' with social disadvantage (p19) based on 3 study populations (Alspac, Gateshead and 'another'). The evidence base therefore contrasts with many other outcomes that are often from large studies with wide geographical coverage. Given the limited number and regional base of studies to date for failure to thrive, a more cautious conclusion is probably justified.

<< The text has been changed to acknowledge this point (Abstract [p.2] and p.18).

- research is scant for some outcomes (e.g wheeze, diarrhoea) and yet a main conclusion of the study

(p18) is that ‘it can be strongly argued that no more epidemiological research needs to be carried out in the UK to address this general effect of area and individual measures of social deprivation. Further research should seek to explore the factors that are more proximal to material and infant health and may help to throw light on the most appropriate times to provide support and the form(s) that such support should take.’ This conclusion is difficult to justify given the weak evidence base for some outcomes, ongoing need to monitor trends and understand the causes of social inequalities.

<< We have modified the text to clarify where there is strength in evidence (Abstract [p.2] and p.18). The point we were making is that the evidence for associations between general measures of deprivation and, for example, birth outcomes and infant mortality is sufficiently strong that further studies looking just at that are probably unnecessary. Therefore studies which try to identify the causal pathways so that the mechanisms by which deprivation affects this, and other, outcomes would be more useful.

- past literature on social inequality has debated whether there are ‘area’ effects separate from ‘individual’ effects. The manuscript by Weightman and colleagues includes both area- and individual-based social indicators yet their discussion is ambiguous on the interpretation of different indicators, i.e. whether they represent different or the same phenomena.

<< As the reviewer noted, there have been many studies which attempted to look at both individual and area effects. I think a reasonable summary is that both matter but individual effects are usually more important.
In these studies the investigators generally chose one of these – either an area measure, such as the Townsend, or an individual one, such as socio-economic class. One might expect that the individual measures have stronger effects but in fact the effect sizes were quite similar with area and individual measures. This might be due to the fact that social class is not the same as deprivation, though is certainly related to deprivation.
There is not enough detail in the studies to comment further on study specific aspects but we have added comments to the discussion about area and individual effect sizes (p.18).

- the discussion emphasises proximal factors, such as ‘stress’ as underlying explanations for social inequalities, with a firm recommendation to include measures of maternal and infant stress levels in future studies (p20). Yet the literature on the intergenerational reproduction of inequalities is relevant in highlighting the importance of events/ exposures in previous generations e.g. for birthweight (see for example, work by Emmanuel). Arguably, intergenerational factors might also be considered in the interpretation of study findings.

<< The text has been changed and a recent reference added noting that intergeneration effects that are independent of economic conditions may have an effect on low birth weight (Collins 2011) (p.19).

- some discussion of heterogeneity in study findings might be worthwhile. If area is important and studies are from different areas might we expect to see some heterogeneity?

<< Agreed. Heterogeneity is clearly an important issue and that we have added a few sentences to the text in both the results and discussion sections (pp.8,9,10,18).

Minor points:
- some explanation of what is meant by ‘calculated at the somewhat smaller lower super output area’ (p11, lines 9-10) would be helpful.

<< The point we were trying to make was that since postcode sectors are larger than LSOAs, area measures in them might be expected to affect the situation of individuals living there less strongly

than in smaller areas. We have reworded the text (p.11) and provided more context in the methodology (p.6).

- please provide a definition of IMD

<< We give this, and a reference, on p.6

VERSION 2 – REVIEW

REVIEWER	David Field, University of Leicester
REVIEW RETURNED	03/04/2012

GENERAL COMMENTS	I reviewed the original version of this manuscript and believe this version is much better and have no major concerns to raise. However I do feel that one of the main points in the discussion has got confused. The discussion has this sentence: Based on the results of this review, it can be strongly argued that no more epidemiological research needs to be carried out in the UK to address this general effect of area and individual measures of social deprivation on birth and infant mortality outcomes. However the discussion continues in the next paragraph: Such studies could usefully build on recent research examining behavioural change interventions regarding the known intermediate determinants of infant health; for example maternal nutrition and overweight68,69,70, and smoking68,71. The impact of teenage pregnancy, which is high in the UK relative to other countries, is also strongly associated with social disadvantage72. My feeling was that these were essentially proxy measures for individual deprivation and hence the two sections conflicts. It may be this is a misunderstanding on my part but even so the relevant sections could benefit from greater clarity in terms of the authors views on how the work should proceed. David Field University of Leicester COI: I am an active researcher and hold grants relating to this topic area.

VERSION 2 – AUTHOR RESPONSE

We thank the reviewer for this additional helpful comment.

The point we were making was that further research might usefully look at the effect of interventions on the proximal determinants rather than further epidemiological investigations, but the context was affected when the point about specific area affects (on epidemiology) was added in and we failed to correct this.

The wording in this paragraph has now been amended for clarification.