Long-term consequences of differences in early growth : epidemiological aspects

Euser, A.M.

3

A regression model with unexplained residuals was preferred in the analysis of the fetal origins of adult diseases hypothesis

M.G. Keijzer-Veen
A.M. Euser
N. van Montfoort
F. W. Dekker
J.P. Vandenbroucke
H.C. van Houwelingen

J Clin Epidemiol 2005; 58:1320-1324

Abstract

Background and objective

A continued controversy exists whether the assessment of the influence of low birth weight on adult blood pressure necessitates adjustment for adult weight in the analysis on the fetal origins of adult diseases hypothesis. Here we first explain the difficulty in understanding an adjusted multivariate regression model, and then propose another way of writing the regression model to make the interpretation of the separate influence of birth weight and changes in weight later in life more straightforward.

Study design and setting

We used a multivariate regression model containing birth weight (standard deviation score; SDS), and residual adult weight (SDS) to explore the effect on blood pressure (or any other outcome) separately. Residual adult weight was calculated as the difference between actual adult weight and the expected adult weight (SDS) given on a certain birth weight (SDS).

Results

The coefficients of birth weight and residual adult weight show directly the effect on the analyzed outcome variable.

Conclusions

We prefer to use this regression model with unexplained residuals when the adjusted variable is in the causal pathway in the analyses of data referring to the fetal origins of adult diseases hypothesis.

Introduction

In the literature on the fetal origins hypothesis, a continued controversy exists whether the assessment of the influence of low birth weight on adult blood pressure necessitates adjustment for adult weight.[1,2] The controversy was fueled by the meta-analysis of Huxley et al.,[3] who described little or no relation between birth weight and adult blood pressure if unadjusted for adult weight, and implied that such adjustment might even be misleading. The effect of adding adult weight as a variable in the regression of blood pressure on weight at birth is intricate: a review by Lucas et al.[4] suggested that such a regression model should in fact be interpreted as the influence of a change in weight between birth and adulthood -and no longer as the influence of birth weight. Nonetheless, the interpretation of data by this concept remains confusing.

Our objective here is first to explain the difficulty in understanding the adjusted regression for the general reader, and then to propose another way of writing the regression model to make the interpretation of the separate influence of birth weight and changes in weight later in life more straightforward. We will explain the model not only conceptually and algebraically, but also by an example on data from an ongoing study on the effect of birth weight on blood pressure. Validation of the model in future analysis is warranted.

The adjusted regression analysis

Originally the association between birth weight and adult blood pressure was analyzed mainly without adjustments for additional variables.[5] Later, it was shown that subjects born with low birth weight tended to gain more weight compared with subjects born with a normal birth weight. Weight gain alone was also associated with an increased risk for high blood pressure. Therefore, adult weight was seen as a potential confounder in the analysis, and adjustment for it became more common.[6] Some studies, however, found a significant association between birth weight and adult diseases only after adjustment for adult weight.[7] Therefore, the need for a multivariate regression model incorporating the effects of both birth weight and adult weight seemed to be the most promising statistical approach. Still, the interpretation of what was achieved by this adjustment remained unclear.

Lucas et al.[4] outlined the consequences of adjustment for adult weight (or length) in a multivariate regression analysis. They proposed using four regression models to analyze the data (Table 1), and stated that in the adjusted models the early and later size of the subjects can no longer be interpreted as stand-alone variables: adjusting early size for later size is a measure of change in size between the earlier and later measurement. In their terminology,

the *early model* describes the relation between early size (i.e., birth weight, or bw) and outcome ($Y = \alpha_1 + \beta_1 X_{bw}$). In the *late model*, the relation between later size (i.e., adult weight, or aw) and outcome is studied ($Y = \alpha_2 + \gamma_2 X_{aw}$). The *combined model* (adding later size to the early model) can be interpreted as describing the relation between change in size and outcome ($Y = \alpha_3 + \beta_3 X_{bw} + \gamma_3 X_{aw}$), as argued by Lucas et al.[4] (see Table 1). Adding the interaction term of early and later size yields the *interaction model*, allowing exploration of whether early size affects the relation between later size and outcome ($Y = \alpha_4 + \beta_4 X_{bw} + \gamma_4 X_{aw} + \delta_4 X_{bw} X_{aw}$)[4]. Note, however, that the changing coefficients (in size and direction) in the combined and the interaction models compared to the early model result in a complicated interpretation. Indeed, the effect of later size is codetermined by the effect of early size on outcome, because adult weight is determined in part by birth weight, which influences the coefficients in the combined model. This also implicitly assumes a quadratic relation between birth weight and outcome in the interaction model, at least under the assumption that birth weight and adult weight are linearly related (Table 1).

Table 1. Interpretation of the multivariate regression model of Lucas et al.[4]

Model description	Equation
Early model, regression analysis of birth weight (bw) to outcome measure	$Y=\alpha_1+\beta_1 X_{bw}$
Late model, regression analysis of adult weight (aw) to outcome measure	$Y=\alpha_2+\gamma_2 X_{aw}$
Combined model, adding later size to early model	$Y=\alpha_3+\beta_3 X_{bw}+\gamma_3 X_{aw}$
Interaction model, adding the interaction of early and adult size to the combined model	$Y=\alpha_4+\beta_4 X_{bw}+\gamma_4 X_{aw}+\delta_4(X_{bw}X_{aw})$
Interaction model, with subtraction of the means	$Y=\alpha_4+\beta_4 X_{bw}+\gamma_4 X_{aw}+\delta_4[(X_{bw}-\bar{X}_{bw})(X_{aw}-\bar{X}_{aw})]$

Variables: X_{bw} birth weight; X_{aw} adult weight; $X_{bw}X_{aw}$ interaction of birth weight and adult weight; Y expected outcome; α intercept; β, γ and δ coefficients.

Which analysis meets the researcher's concerns?

Whether later size (e.g., adult weight) is a confounder in the analysis of early size (e.g., birth weight) and adult diseases, such as hypertension, or is rather a factor in the causal pathway is an ongoing debate in the literature. Adjustment for adult weight might not be justified after all.[1-3] Whatever the causal explanation, birth weight is positively correlated with adult weight and adult weight is correlated with adult blood pressure; therefore, we do first of all expect that any positive relation between birth weight and adult blood pressure will be attenuated upon adding adult weight to the model (the coefficient of birth weight will become closer to zero). Next, according to Lucas et al.,[4] it might be those who grew more than expected (i.e., attained greater adult weight for a given birth weight) who would develop the higher blood pressures. This would reverse the already attenuated relation with birth weight into a negative relation.

As researchers, we remain interested in the separate contribution of birth weight (reflecting prebirth influences) and change in weight from birth to adulthood (reflecting early life influences). Thus, we want to have an estimate of both. We want first an estimate of the effect of birth weight alone, and second, what we really want to know is the effect of someone growing more in weight than would be expected from a given birth weight. In a statistical analysis this can be accomplished in a single model by first calculating the expected adult weight, or eaw, based on birth weight ($X_{eaw} = \alpha_0 + \beta_0 X_{bw}$), and then subtract expected adult weight from actual adult weight - which is in effect the calculation of a residual ($X_{res} = X_{aw} - X_{eaw}$) (Table 2). Adding this residual increase in weight in a regression model of blood pressure on birth weight has three advantages. First, it leaves the coefficient of birth weight unchanged (because the effect of birth weight on adult weight is already taken out of the residual). Second, it gives us an insight into the additional influence of growing more in weight than expected upon the adult blood pressure. Third, the two variables in the regression model (birth weight and the residual increase in weight) are now independent, because the residual cannot be predicted from birth weight. Therefore, the interaction model does not assume a quadratic relation anymore. Li et al.[8] earlier described this model in the analyses of a Guatemalan study in which the association between prenatal and postnatal growth and adult body composition was studied; however, no algebraic explanation of this model was shown.

The proposed technique is not unique to the problems of interpreting regression in the fetal origins of adult diseases hypothesis. It has been used in social sciences literature under the name of *residualized gain score*.[9,10]

It should be noted that algebraically the combined model of Lucas et al.[4] is the same as the combined model using unexplained residuals (Appendix A); however, the effect of birth weight and residual postnatal growth is directly shown by the coefficients of the proposed unexplained residual regression model. In both models, for the interaction model we suggest to multiply not just the two variables, but first subtract the mean of that variable. In the model of Lucas et al.,[4] this becomes $(X_{bw} - \overline{X}_{bw})(X_{aw} - \overline{X}_{aw})$; in the proposed model this becomes $(X_{bw} - \overline{X}_{bw})(X_{res} - \overline{X}_{res})$. As the mean of a residual is zero, this can be rewritten in $(X_{bw} - \overline{X}_{bw})X_{res}$.

Next to the model of Lucas et al.[4], other simplified models are suggested to use in the analysis of the fetal origins of adult diseases hypothesis to measure the effect of change in weight. When researchers think about the problem, they often intuitively propose to subtract adult weight (standard deviation score; SDS) and birth weight (SDS) as a measure of change in weight and add this to birth weight (SDS) in a multivariate regression model. The problem

Table 2. Interpretation of unexplained residual regression model

Model description	Equation
Early model, regression analysis of early weight to outcome measure	$Y=\alpha_1+\beta_1 X_{bw}$
Late unexplained residual model, regression analysis of residual of expected adult weight to outcome measure[a]	$Y=\alpha_2+\gamma_2 X_{res}$
Combined unexplained residual model, adding the residual of the expected adult weight to early model	$Y=\alpha_3+\beta_3 X_{bw}+\gamma_3 X_{res}$
Interaction unexplained residual model, adding the interaction the difference between birth weight and the mean birth weight and the difference between the residual and the mean residual of the expected later size to the combined unexplained residual model[b]	$Y=\alpha_4+\beta_4 X_{bw}+\gamma_4 X_{res}+\delta_4[(X_{bw}- \overline{X}_{bw})(X_{res}- \overline{X}_{res})]$ in which $\overline{X}_{res}$ is zero.

Variables: X_{bw} expected birth weight; X_{res} residual of expected adult weight, based on birth weight; $(X_{bw}- \overline{X}_{bw})(X_{res}- \overline{X}_{res})$, interaction of birth weight and residual of expected adult weight; Y, expected outcome; α, intercept; β, γ, and δ, coefficients.

[a] First, expected adult weight X_{eaw} is calculated, based on birth weight $(\alpha_0 + \beta_0 X_{bw})$. Then, the residual for expected adult weight is calculated as $X_{res} = (X_{aw} - X_{eaw})$. This leads to the equation in column 2.

[b] In the interaction unexplained residual model, $\beta_1 = \beta_3 = \beta_4$ and $\gamma_2 = \gamma_3 = \gamma_4$.

with this model is the phenomenon of regression to the mean. The relative position of subjects with low birth weight will tend to increase and that of subjects with high birth weight will tend to decrease over time. This phenomenon is not present in the unexplained residual model, because in the calculation of adult weight residuals out of birth weight we force the residuals not to be related to birth weight. The coefficient of birth weight in a linear regression model of adult weight residuals is exactly zero (with very small confidence interval and a *P*-value of exactly 1).

Second, it has also been suggested to use population-based SD scores instead of calculating the residual of expected adult weight (SDS). However, the subjects studied in research concerning the fetal origins of adult diseases hypothesis are mostly not comparable to the general population, because of an overrepresentation of the low birth weight subjects. Subjects with low birth weight have different growing patterns. Therefore, for most studies it is not recommended to use population-based SD scores to calculate expected adult weight (SDS) and weight gain (SDS). In addition, it takes about 3 years after birth before an individual will track on his or her centile, especially in low birth weight infants. If the population-based reference standards were to be used as a measure for expected adult weight (SDS), in which the mean adult weight (SDS) will be zero, low birth weight (SDS) subjects will tend to have a negative residual for adult weight, because of their suboptimal growth. Then, the residual would not reflect the correct variable to answer our second question: what is the effect of someone growing more in weight than would be expected from a given birth weight? So, calculating the residual adult weight out of birth weight should be performed with the expected adult weight from the group of subjects that are used in the study.

Conclusion, and proposal

Algebraically, the combined model of Lucas et al.[4] and our combined model with the residuals increase in weight can be rewritten in terms of each other, except for the situation where an interaction term is entered (see Appendix A). In the proposal by Lucas et al.[4], however, one needs two separate models: first estimating the coefficient from the early model, and then looking at the coefficient for attained weight in the combined model (without paying attention to the coefficient of birth weight in that combined model, because the latter has become meaningless). For this reason, we prefer the proposed model with residuals because it permits in a more straightforward way to estimate the effect of birth weight and the effect of additional weight gain in a single model. We also prefer to use the interaction model containing the unexplained residuals, because no quadratic relation is assumed and because in principle all coefficients show their own effect without mutual influence (Table 2). Therefore, the interpretation of the model with the unexplained residuals is easier. An example with numerical data from an ongoing study in the Netherlands is given in Appendix B, including Tables B1 and B2.

In conclusion, we prefer to use regression model with unexplained residuals when the adjusted variable is in the causal pathway in the analyses of data referring to the fetal origins of adult diseases hypothesis.

Acknowledgments

We thank S. le Cessie (Dept. of Medical Statistics, LUMC, Leiden, The Netherlands) and M.J.J. Finken (Department of Clinical Epidemiology, LUMC, Leiden, The Netherlands) for their comments on the manuscript.

Appendix A. Derivations

To rewrite the combination unexplained residual model in the combination model of Lucas et al.,[4] where Y is the expected outcome; α is the intercept; β is a coefficient; X_{bw} is the birth weight; X_{aw} is the adult weight; X_{eaw} is the expected adult weight, based on early size ($\alpha_0 + \beta_0 X_{bw}$); and X_{res} is the residual of expected adult weight ($X_{aw} - X_{eaw}$):

$$Y = \alpha_1 + \beta_1 X_{bw} + \gamma X_{res} \qquad \text{(the unexplained residual model)}$$

$$Y = \alpha_1 + \beta_1 X_{bw} + \gamma [X_{aw} - (\alpha_0 + \beta_0 X_{bw})]$$

$$Y = \alpha_1 + \beta_1 X_{bw} + \gamma X_{aw} - \gamma \alpha_0 - \gamma \beta_0 X_{bw}$$

$$Y = (\alpha_1 - \gamma \alpha_0) + (\beta_1 - \gamma \beta_0) X_{bw} + \gamma X_{aw}$$

$$Y = \alpha' + \beta' X_{bw} + \gamma' X_{aw} \quad \text{(the Lucas et al.[4] model)}$$

$$\alpha' = \alpha_1 - \gamma \alpha_0$$

$$\beta' = \beta_1 - \gamma \beta_0$$

$$\gamma' = \gamma$$

To add the interaction term $(X_{bw} - \overline{X}_{bw}) * (X_{aw} - \overline{X}_{aw})$ into the Lucas et al.[4] model, first suppose that X_{aw} is exactly linearly related to X_{bw}. Then, where ε is the residual:

$$X_{aw} = \alpha_0 + \beta_0 X_{bw} + \varepsilon$$

and

$$\overline{X}_{aw} = \alpha_0 + \beta_0 \overline{X}_{bw}$$

So, $(X_{aw} - \overline{X}_{aw}) = (\alpha_0 + \beta_0 X_{bw} + \varepsilon) - (\alpha_0 + \beta_0 \overline{X}_{bw})$

which can be rewritten as:

$$(X_{aw} - \overline{X}_{aw}) = \beta_0 (X_{bw} - \overline{X}_{bw}) + \varepsilon$$

Adding this to the interaction term

$$(X_{bw} - \overline{X}_{bw})*(X_{aw} - \overline{X}_{aw}),$$

the equation will be:

$$(X_{bw} - \overline{X}_{bw})*(\beta_0 (X_{bw} - \overline{X}_{bw}) + \varepsilon)$$

This can be rewritten as:

$$\beta_0 (X_{bw} - \overline{X}_{bw})^2 + (X_{bw} - \overline{X}_{bw}) * \varepsilon$$

Here, the quadratic relation between birth weight and outcome is shown.
To add the interaction term into the unexplained residuals model:

$$(X_{bw} - \overline{X}_{bw})(X_{res} - \overline{X}_{res}) = (X_{bw} - \overline{X}_{bw})X_{res}$$

In this model, X_{res} (the residual of expected adult weight) is independent of X_{bw} (birth weight). All coefficients show the unadjusted effect of the variable on the outcome variable.

$\overline{X}$

Appendix B

Example of regression analysis according to Lucas et al.[4] and the *unexplained residual model* Tables B1 and B2): In a prospective study the systolic blood pressure at adult age was measured. Birth weight standard deviation scores (BW_{SDS}) and adult weight standard deviation scores (AW_{SDS}) were known.

In Table B1, the change in estimated coefficients is shown in both the combined as the interaction model, both with and without the subtractions of means, when the model of Lucas et al.[4] is used. In the early model, birth weight (SDS) is related to blood pressure with a coefficient of 0.361. When adult weight (SDS) is added to the model the coefficient for birth weight (SDS) changed into a negative one (-0.0928). This is a result of the relation between birth weight (SDS) and adult weight (SDS). This change in the estimated coefficient is confusing for many authors; which coefficient is giving information about the relation between birth weight (SDS) and blood pressure?

In the combined unexplained residuals model, these estimated coefficients do not change (Table B2) when adult weight (SDS) is added to the model. The coefficient for birth weight and residual weight gain shift slightly in the interaction model in comparison with the combined weight residual model: probably this is due to non-exact-linear correlation between birth weight and weight gain.

The δ_4 coefficient does not change much in our example. The reason is that X_{bw} is not related to systolic blood pressure. Therefore, the δ_4 coefficient in the model of Lucas et al.[4] is comparable to the δ_4 coefficient in our model. When X_{bw} would be quadratically related to blood pressure, the δ_4 coefficient would differ much in both models.

Table B1. Estimated coefficients in our example when the Lucas et al.[4] model is used, with two types of interaction

		β	γ	δ
Model of Lucas et al. [4]	α	X_{bw}	X_{aw}	$X_{bw}X_{aw}$
Early	122.943 (α_1)	0.361 (β_1)		
Late	123.908 (α_2)		2.069 (γ_2)	
Combined	123.743 (α_3)	-0.0928 (β_3)	2.096 (γ_3)	
Interaction	123.771 (α_4)	-0.0078 (β_4)	2.231 (γ_4)	0.120 (δ_4)
		X_{bw}	X_{aw}	$(X_{bw} - \overline{X}_{bw})(X_{aw} - \overline{X}_{aw})$
Interaction with subtracted means	123.710 (α_4)	-0.0766 (β_4)	2.125 (γ_4)	0.120 (δ_4)

Variables: α, intercept; β, γ, and δ, coefficients; X_{bw} expected birth weight (SDS); X_{aw} expected adult weight (SDS)

Table B2. Estimated coefficients in our example when the unexplained residual model is used

		β	γ	δ
Model unexplained residuals	α	X_{bw}	X_{res}	$(X_{bw} - \overline{X}_{bw})X_{res}$
Early	122.943 (α_1)	0.361 (β_1)		
Late	123.623 (α_2)		2.096 (γ_2)	
Combined	123.943 (α_3)	0.361 (β_3)	2.096 (γ_3)	
Interaction	123.943 (α_4)	0.361 (β_4)	2.121 (γ_4)	0.102 (δ_4)

Variables: α intercept; β coefficient; X_{bw} birth weight (SDS); X_{res} residual of expected adult weight (SDS); $\overline{X}_{res}$ equals zero in interaction term $(X_{bw} - \overline{X}_{bw})(X_{res} - \overline{X}_{res})$. Expected adult weight $(X_{eaw}) = -0.382 + 0.216\, X_{bw}$. Residual of adult weight $(X_{res}) = X_{aw} - X_{eaw}$

References

1. Kramer MS. Invited commentary: association between restricted fetal growth and adult chronic disease: is it causal? Is it important? Am J Epidemiol 2000; 152:605-608.
2. Paneth N, Susser M. Early origin of coronary heart disease (the "Barker hypothesis"). BMJ 1995; 310:411-412.
3. Huxley R, Neil A, Collins R. Unravelling the fetal origins hypothesis: is there really an inverse association between birthweight and subsequent blood pressure? Lancet 2002 Aug 31 360;659-665.
4. Lucas A, Fewtrell MS, Cole TJ. Fetal origins of adult disease-the hypothesis revisited. BMJ 1999; 319:245-249.
5. Barker DJ, Bull AR, Osmond C, Simmonds SJ. Fetal and placental size and risk of hypertension in adult life. BMJ 1990; 301:259-262.
6. Law CM, de SM, Osmond C, Fayers PM, Barker DJ, Cruddas AM et al. Initiation of hypertension in utero and its amplification throughout life. BMJ 1993; 306:24-27.
7. Fall CH, Pandit AN, Law CM, Yajnik CS, Clark PM, Breier B et al. Size at birth and plasma insulin-like growth factor-1 concentrations. Arch Dis Child 1995; 73:287-293.
8. Li H, Stein AD, Barnhart HX, Ramakrishnan U, Martorell R. Associations between prenatal and postnatal growth and adult body size and composition. Am J Clin Nutr 2003; 77:1498-1505.
9. ter Kuile MM, Spinhoven P, Linssen AC, Van Houwelingen HC. Cognitive coping and appraisal processes in the treatment of chronic headaches. Pain 1996; 64:257-264.
10. Rowlands AV, Ingledew DK, Powell SM, Eston RG. Interactive effects of habitual physical activity and calcium intake on bone density in boys and girls. J Appl Physiol 2004; 97:1203-1208.