Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE MEDICHE GENERALI E SCIENZE DEI SERVIZI

Ciclo XXVI

Settore Concorsuale di afferenza:
06/M1 – IGIENE GENERALE E APPLICATA E STATISTICA MEDICA

Settore Scientifico Disciplinare di afferenza:
MED/01 – STATISTICA MEDICA

TITOLO TESI

Trajectories and predictors of growth and neurodevelopment in Very Low Birth Weight infants

Presentata da: Dott. Dino Gibertoni

Coordinatore Dottorato

Prof. Nicola Rizzo

Relatore

Prof.ssa Maria Pia Fantini

Esame finale anno 2014

Il faut se tromper, être imprudent. Les hommes prudents sont des infirmes.

(Jacques Brel)

CONTENTS

ACKNOWLEDGMENTS

Several are the people to whom I owe a debt for helping me carry out this study. First of all drs. Silvia Vandini and Silvia Savini, who provided me data, feedbacks on the manuscript drafts and most of all who were patient enough to fulfill all my requests of data revisions and of explanations on the clinical and psychological processes of preterm children. For data correction, a big acknowledgement goes to dr. Mariangela D'Antuono. Profs. Giacomo Faldella, Alessandra Sansavini and Luigi Corvaglia supported from the beginning the challenge to apply some new statistical techniques to their workfield. Last but not least, prof. Maria Pia Fantini and dr. Paola Rucci supervised the study and succeeded in make things going – not only for this final dissertation but for all the Doctorate's three years.

ABSTRACT

Neurodevelopment of preterm children has become an outcome of major interest since the improvement in survival due to advances in neonatal care. Many studies focused on the relationships among prenatal characteristics and neurodevelopmental outcome in order to identify the higher risk preterms' subgroups. The aim of this study is to analyze and put in relation growth and development trajectories to investigate their relationship.

346 children born at the S.Orsola Hospital in Bologna from 01 January 2005 to 30 June 2011 with a birth weight of <1500 grams were followed up in a longitudinal study at different intervals from 3 to 24 months of corrected age. During follow-up visits, preterms' main biometrical characteristics were measured and the Griffiths Mental Development Scale was administered to assess neurodevelopment. Latent Curve Models were developed to estimate the trajectories of length and of neurodevelopment, both separately and combined in a single model, and to assess the influence of clinical and socio-economic variables.

Neurodevelopment trajectory was stepwise declining over time and length trajectory showed a steep increase until 12 months and was flat afterwards. Higher initial values of length were correlated with higher initial values of neurodevelopment and predicted a more declining neurodevelopment. SGA preterms and those from families with higher status had a less declining neurodevelopment slope, while being born from a migrant mother proved negative on neurodevelopment through the mediating effect of a being taller at 3 months. A longer stay in NICU used as a proxy of preterms' morbidity) was predictive of lower initial neurodevelopment levels.

At 24 months, neurodevelopment is more similar among preterms and is more accurately evaluated. The association among preterms' neurodevelopment and physiological growth may provide further insights on the determinants of preterms' outcomes. Sound statistical methods, exploiting all the information collected in a longitudinal study, may be more appropriate to the analysis.

INTRODUCTION

The decline of mortality in preterm infants led many researchers to focus their studies on neurodevelopmental disabilities, that remain as a great burden on infants' families and on health care systems. The incidence and the factors associated with neurodevelopmental disabilities have been investigated from a multiplicity of different viewpoints. Earlier studies,[1] focusing on the role played by nutrition on weight gain during hospitalization, showed that enriched formula milk fed preterms had both faster weight gain and improved cognitive outcomes during infancy and school-age compared with preterms who were term formula fed. The relationship of faster weight gain and head growth during the NICU stay with higher cognitive scores was found also by Ehrenkranz et al.[2] The link among nutrition, weight gain and neurodevelopment was explained by the evidence that preterms who had inadequate nutrition in their early days were exposed to higher risks of infection and comorbidities, which in turn further delayed the achievement of an optimal nutrient level. The result is a reduced child's overall health and energy level, that may lead to a lower neurodevelopment.[2]

These studies were generally conducted using as predictors clinical and anthropometrical variables collected only during the NICU stay and as outcomes the neurodevelopmental scores assessed at 18-22 months of life or later at school age. More recently, Belfort et al.[3] pinpointed that a limitation of those studies was the wide temporal window among the outcomes and the explanatory variables, therefore ignoring the effect of mediating factors that may intervene during the post-NICU preterms' development. The same authors importantly underline that, as a consequence, this leads to ignoring the association among development and socio-economic factors that may play a role mainly after discharge, when preterms' caretakers are mostly members of their family environment.

Hence, the availability of data spanning the initial months of life of preterms after hospital discharge is not sufficient to ensure a comprehensive understanding of the mechanisms underlying the developmental process. Appropriate study design and statistical methods may actually provide a considerably better insight into this fundamental phase of preterms' life.

This study uses Latent Curve Model analysis, a methodology that fully utilizes all available data, to attain three objectives:

- to obtain a model describing preterms' neurodevelopment trajectory from 3 to 24 months of corrected age;
- to obtain a model describing preterms' height growth trajectory from 3 to 24 months of corrected age;
- to combine neurodevelopment and growth trajectories into a single combined model, describing the relationships existing among the two patterns of change.

The models outputs include the shape and parameters of these growth trajectories, the effects of clinical and socio-economic predictors on the baseline levels and slopes of the two outcomes and lastly the degree of association among the height and neurodevelopment trajectories.

This study uses only height as a growth indicator but its replication substituting height with weight or cranial circumference is straightforward.

MATERIALS AND METHODS

Study population

The study population included ELBW (≤1000 g) and VLBW (≤1500 g) infants or infants born at less than 32 weeks of gestational age admitted at birth to the Neonatal Intensive Care Unit (NICU) of S.Orsola University Hospital, Bologna (North Eastern Italy) from 1/1/2005 to 06/30/2011, and enrolled in a follow-up program. Follow-up visits were made at 3, 6, 9, 12, 18 and 24 months corrected age; these specific time-points were chosen because they coincide with important milestones in the process of acquiring cognitive and functional abilities.

Written informed consent to participate in the study was obtained from parents. Data were anonymized prior to data analysis and the study protocol was approved by the local Ethics Committee.

Outcome measures of growth and neurodevelopment

The growth of newborns and specifically of preterm newborns has been quite extensively studied. A great effort has been made to set up longitudinal studies aimed to determine the standard trajectories of growth pattern for weight, length and cranial circumference. In these studies, different analytical and statistical criteria to summarize the results as well as different time points for the follow-up were considered.

A first distinction must be made among *growth* and *catch-up growth*. The term "catch-up growth" was introduced by Prader[4] and Tanner[5] in 1963 and was usually intended for height growth[6]; it describes the period of rapid linear growth in children that followed a period of growth inhibition, whose effect is to reconduce the children to their expected preretardation growth curve[7]. For preterm infants, it is referred to the early quick acceleration usually observed in SGA newborns. De Wit et al.[6] argued that the correct measure for height catch-up growth is the standardized deviation score (SDS) and its change over time observed well beyond the first year life, because in that period this measure may be highly variable and still very dependent on the birth height. They

gathered information from previous studies showing that 80%-85% of SGA newborns recover in a normal height range in the first year of life, and that a similar result has been provided for preterms before the age of 3 years.

In a very comprehensive paper, Sullivan et al.[8] stratified a sample of 194 infants into five subgroups defined by SGA, term/preterm condition and presence/absence of comorbidities in preterms applying mixed effects linear models to test the differences in the group trajectories. Comparing infants on z-scores of weight, height and BMI over a long follow-up period until 12 years of age and using a set of biological and socio-cultural predictors to investigate the determinants of catch-up, they found that preterms had generally a steeper growth in the first 18 months but only some preterms groups reached the same growth as full-terms at 12 years of age. Therefore there is some evidence that a catch-up of length may be found not only for SGA but for VLBW preterms as well. However, since the aim of this study is not specifically focused on differences among SGA and non-SGA preterms, from now onwards the gain in length over time will be simply described as growth.

Neurodevelopment was evaluated using the revised Griffiths Mental Development Scale (GMDS-R, 0–2 years version).[9] This scale consists of 276 dichotomous items that explore five functioning domains: Locomotor, Personal-Social, Hearing and Language, Eye and Hand Coordination, Performance. The assessment of these five separate domains allows to understand whether a delay in neurodevelopment may be due to some specific cognitive area, thus allowing to obtain a detailed cognitive profile for each preterm. Raw and standardized scores for each domain and a composite raw (RGQ) and standardized General Quotient (GQ) were calculated. Raw scores are the number of items appropriate for the infant's age that were met by the preterm at each administration; RGQ is the sum of the five raw subscales' scores. In the absence of normalized scores for the preterm infant Italian population,[10,11] standard scores were obtained using the tables of standardized scores for the English infants population.[12] For each domain, standardized scores have mean=100 and sd=16, while GQ has mean=100.5 and sd=11.8. Comparison with the standard values allows to evaluate whether preterms' competences are different, though the absence of normalized scores on the Italian population and possible biases on first months' scores lead to a cautious approach in the interpretation of results.[13] In this study, only the composite GQ and RGQ scores were analyzed, using the raw scores for the sample description and the

standardized scores as the repeated observed measures in the latent growth models. For descriptive purposes GQ was also classified into the following categories: normal development (GQ≥88.7), mild (-1 to -2 SD, corresponding to 88.6-76.9 GQ scores), moderate (-2 to -3 SD, corresponding to 76.8-65.1 GQ scores) and severe delay (<-3 SD, corresponding to GQ≤65).[11] The test was administered by two psychologists with long-standing experience in developmental assessment.

Griffiths Scales have been used in several studies[13] on preterms that examined the relation among neurodevelopment and gestational age[11,14,15] or gender[16] or to investigate prospectively the relation among neurodevelopment and subsequent cognitive assessments, as scholar age cognitive delay[17] or intellective quotient at 42 and 66 months of age.[18] Findings are of a positive association among gestational age and neurodevelopment at 24 months, with ELBW newborns showing the worst performances; prospectively, a low development score at 24 months was found to be predictive of impairment at school age.

From the statistical viewpoint, these studies did not always rely on methods that allowed a thorough utilization of data available at each time point. Specifically, Gutbrod,[19] Rijken,[20] Leppänen[21] and Brandt[22] tested the association between clinical predictors and growth or between preterms' growth and national growth charts with separate analyses at each observed timepoint; Mercier[23] evaluated the predictors of severe disabilities at 18-24 months of age, and Ehrenkranz[2] tested the association among in-hospital growth velocity and neurodevelopmental and growth outcomes at 18 to 22 months. Other studies instead took advantage of all the available data by using MANOVA for repeated measures[11] and mixed-effects linear regression[3,8] to estimate the growth curve parameters. In our study, two different outcomes will be measured: the neurodevelopment and the length growth curves until 24 months of age. To obtain these estimates and to subsequently put them in relation the Latent Curve Models methodology was used.

Independent variables

Biometric and clinical characteristics of the preterms were collected during their stay in NICU and in the follow-up visits, as well as preterms' parents socio-economic variables. Each of these variables may then be evaluated as a potential predictor of neurodevelopment and length trajectories.

The independent variables analyzed in this study are:

Gender has been studied as a predictor of growth with contrasting results. Similar growth patterns of height among genders were found by Casey,[24] Rijken[20] and Sullivan,[8] while Hack[25] found better growth in females and Guo[26] and Morris[27] in males. As for neurodevelopment, both the studies of Ehrenkranz[2] and Mercier[23] showed an increased risk of severe disabilities for males at 18-24 months of corrected age.

Gestational age (GA) in weeks was based on the last menstrual period and first-trimester scan.

Small for gestational age (SGA) is a binary variable indicating whether newborns' birth weight was below a standardized score of -1.28, corresponding to the 10th percentile of the reference neonatal growth charts developed ad hoc for the Italian population of preterm infants.[28] While it is known that being SGA is associated with a worse outcome in extremely preterm infants,[29] its influence on neurodevelopment is controversial. A negative effect on neurodevelopment was found only when paired with insufficient postnatal growth[30] but in other studies SGA infants were found to have a better 5 and 20 months neurodevelopment than AGA infants paired by birthweight[19] or to have no effect on 5-year cognitive outcome.[21]

Length of stay in the NICU (HS) in days was used as a proxy of neonatal morbidity after conducting preliminary bivariate linear regression analyses, showing that this variable was positively and significantly ($p<0.05$) associated with the following severe postnatal conditions: mechanical ventilation, chronic lung disease (oxygen need at 36 weeks postmenstrual age), early onset and late onset sepsis (including both culture proven or clinical sepsis), necrotizing enterocolitis (requiring surgery), severe intra-ventricular hemorrhage (grade 3 and 4 as classified by Papile et al.,[31] including post-hemorrhagic hydrocephalus requiring surgery or periventricular leukomalacia, classified as the presence of periventricular cysts at any cranial ultrasound performed during

hospital stay) and severe rethinopathy of prematurity (stages 3 to 5 according to the International Committee for the Classification of Retinopathy of Prematurity).[32]

Mother's age at birth was introduced as a potential predictor of neurodevelopment; a previous study[33] showed that mother's age was significantly associated with socio-economic status. A higher mother's age, as well as other variables such as the number of siblings, may also be indicative of better caregiving skills.

Twins have usually a high incidence among preterms; a recent study[34] found that twins had a higher risk of mortality than singletons and a slightly higher long-term risk of motor and neurodevelopmental deficencies. The variable was coded as 0 for singletons and 1 for twins.

Information on **siblings** was entered in the study with a dichotomous variable, indicating whether preterms were firstborns (coded 0) or had older siblings (coded 1). In previous studies[35,36] a negative influence on neurodevelopment of having siblings was found, explained by mothers spreading their attention among more than one child, therefore allowing a lower responsiveness to children needs, when compared to firstborns' mothers.

Migrant condition of the mother is a dichotomous variable (mother of Italian nationality vs. other nationalities) that was included as an indicator of preterms' socio-cultural environment. Migration to Bologna and to Italy is a quite recent phenomenon, originating mainly from developing countries and Eastern Europe and comprising by a large amount non-specialized workers. For these reasons, migrants often live in conditions of material deprivation (associated to poverty) and social deprivation (involving isolation and low levels of social cohesion) that are linked to a higher risk of a preterm birth.[37] These conditions of deprivation may likely affect also the post-hospitalization phase, when preterms usually live with their family, whose care-providing ability may be impaired by factors like poor housing conditions and low fluency in Italian language.

Hollingshead Index (HI) is a well-established index of socioeconomic status that takes into account the educational and occupational status of the preterms' parents.[38] A higher HI is directly related to higher educational and occupational status, that is believed to reflect acquired knowledges and skills. HI ranges from 0,

when both parents have no formal education and are unemployed or retired, to 66 when both parents share the highest educational (graduate degree) and occupational (higher executives or major professionals) levels. Widely used in psychometry, HI was employed in studies on preterm infants[8] as a stratification variable.

Diet at discharge and at 3 months of corrected age was coded as maternal (own mother's raw milk, either given by bottle or directly from the breast), mixed and exclusive formula milk. Fortification of bottle-administered human milk was routinely done during hospitalization and recommended after discharge until the weight of 3.5 kg was achieved. When needed, "preterm formula" was used during hospitalization and "post-discharge formula" was recommended after discharge until the weight of 3.5 kg was achieved.

Sustained human milk feeding until 3 months of age is a dichotomous indicator that takes value 1 when a preterm was fed maternal or mixed milk both at discharge and at 3 months corrected age. The beneficial effects of raw human milk feeding on neurodevelopment are well-known; however in a recent study on preterms (Gibertoni) a significant positive association with neurodevelopment at 24 months corrected age was found only for preterms who were fed human milk until 3 months of corrected age.

Nursery school attendance is a dichotomous variable observed at 12, 18 and 24 months of preterms' age stating whether they were attending a nursery school. It was hypothesized that preterms' attendance of an environment external to the family could be associated to a different neurodevelopment level (AMPLIARE).

Statistical analysis

The study sample characteristics were summarized with descriptive statistics that included mean ± standard deviation and median with interquartile range for continuous variables, absolute and relative frequencies for categorical variables. Bivariate relationships among predictors were analyzed using t-tests or chi-square tests or linear correlations, depending on the level of measurement of the analyzed variables.

To determine whether length of NICU stay (HS) is a proxy of newborns' severity of illness, a series of bivariate linear regressions were carried out, where HS was the outcome and each complication was in turn the only predictor.

Latent Curve Models[1]

As Preacher et al.[39] clearly state in the Introduction of their book, "early approaches to investigating change [over time] were very limited in that (a) they focused exclusively either on group-level or on invidual-level growth and (b) they addressed only two occasions of measurement, resulting in data too impoverished to allow examinations of some of the most basic and interesting hypotheses of change over time". The afore-mentioned shortcomings may be overcome by using Latent Curve Models (LCM), a class of statistical methods belonging to the Structural Equation Models (SEM) family. LCM are designed to deal with longitudinal data collected on a sample of individuals, allowing to make inferences both on the interindividual change and on the intraindividual change over time, furthermore investigating on the predictors of change. LCM possess all the advantages of SEM, such as the ability to deal effectively with missing data and comprehensive measures of model fit to the data. Maybe the greatest advantage of SEM and LCM, linking the statistical analysis to the theoretical speculation, is the extreme flexibility in model design.

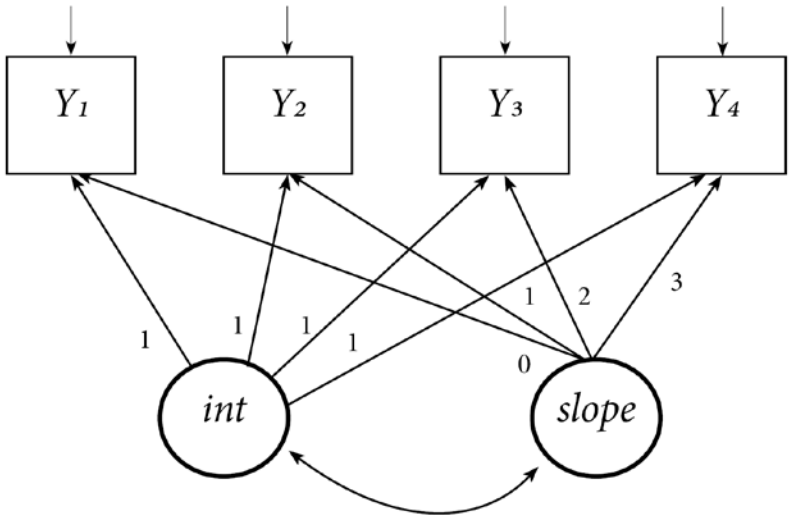
Several types of LCM of increasing complexity may be set up, depending on the hypotheses that can be made on the type of change. The simplest LCM is the **Unconditional Linear Latent Curve Model** (ULLCM), which considers the series of the observed values Y_i as an expression of two latent growth factors, the starting level of the outcome measure (the *intercept)* and its linear growth rate (the *slope).* These factors are sometimes referred as the *true* initial measure and slope,[40] because they represent estimates of the unknown corresponding values in the population. Since the growth curve estimated with an ULLCM is a straight line, this parameterization is suitable for outcomes that are supposed to change at a constant rate over time. In the ULLCM, the loadings from the latent intercept to the observed variables are all set to 1, because the intercept equally influences each observation; the loadings from the slope factor are set to a sequence of values that are proportional to the distance between the time points at which the observed variable was measured. By setting to 0 the loading from the slope to the first repeated measure, the intercept is assumed to estimate the mean value of y at the first

[1] General references for the methodology of Latent Curve Models are the books by: Bollen and Curran[69]; Preacher, Wichman, MacCallum and Briggs;[39] Wang and Wang.[70]

assessment period. Each observed measure has an estimated residual, that is the part of variance that was not explained by the two growth factors. The parameters estimated by an ULLCM are:

- the mean intercept μ_α, which represents the mean initial level of the analyzed outcome; inference is made on μ_α to test the hypothesis that the mean intercept is different from 0;

Figure 1 –Diagram of an Unconditional Linear Latent Curve Model



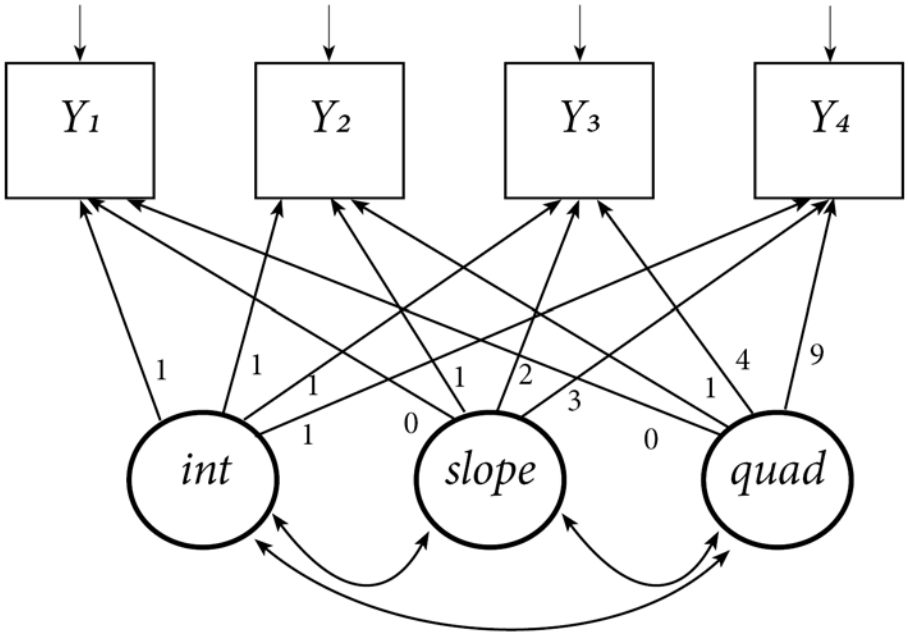
the mean slope μ_β, which represents the mean rate of change of the analyzed outcome per unit of time as defined by the lags on the loadings; inference is made on μ_β to test the hypothesis that the mean growth rate is different from 0: should the null hypothesis be rejected, a significant increase or decrease was assessed;

- the variance of the mean intercept ψ_α, which represents the variability of individual initial levels of the outcome; inference is made on ψ_α to test the hypothesis that the individuals share the same initial level of the measured outcome;
- the variance of the slope ψ_β, which represents the variability of individual rates of change in the outcome; inference is made on ψ_β to test the hypothesis that the individuals share the same slope;

- the correlation between intercept and slope $\psi_{\alpha\beta}$, which represents the covariation among the growth factors; inference is made on $\psi_{\alpha\beta}$ to test the hypothesis of a relation among starting values of the outcome and the growth rate.

When a linear curve does not represent the data adequately, nonlinear models may be a better solution than ULLCM. **Unconditional Latent Quadratic Curve Models** (ULQCM) are an upgrade over ULLCM designed to produce a growth curve in a quadratic form. This type of curves are identified by two components: the linear component, that corresponds to the mean rate change, and the quadratic component, that corresponds to the acceleration (or deceleration if its parameter is negative) of the linear slope. Thus, ULQCM are well-suited to represent outcomes that have an initial steep increase followed by a stabilization as well as outcomes that have a later increase after an initial slow change. The curvilinear pattern is defined adding to the linear model a third latent variable that corresponds to the quadratic term and by setting its loadings on the observed values to the squares of the linear slope loadings.

Figure 2 –Diagram of an Unconditional Quadratic Latent Curve Model

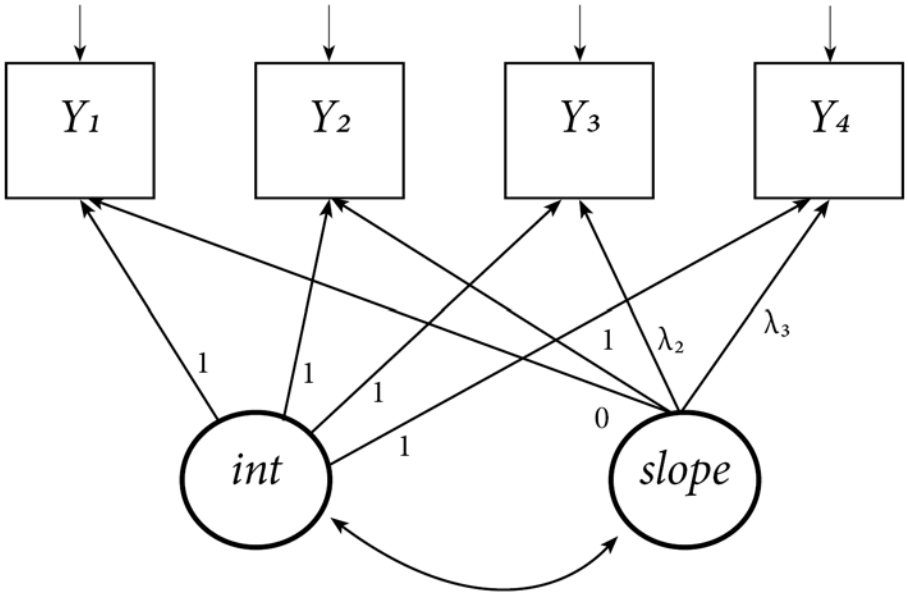


Four more parameters need to be estimated in an ULQCM:

- the mean quadratic component μ_γ, which represents the degree of curvature rate in the trajectory; inference is made on μ_γ to test the hypothesis that the mean curvature rate is different from 0: should the null hypothesis not be rejected, then the model is equivalent to a linear model;
- the variance around the mean quadratic component ψ_γ, which represents the variability of individual curvature rates; inference is made on ψ_γ to test the hypothesis that the individuals share the same degree of curvature;
- the correlations between the mean quadratic component and the intercept ($\psi_{\alpha\gamma}$) and between the mean quadratic component and the slope ($\psi_{\beta\gamma}$), which represent the covariations among curvature and the other growth factors; inference is made on these correlations to test the hypothesis of a relation among starting values of the outcome, growth rate and the curvature.

Another way to obtain nonlinear growth curves is to assume that the growth progression is not given a priori by setting each loading to a parameter corresponding to time lag between observations, but instead it is unknown and must be estimated by the model. This type of model is defined **Unconditional Completely Latent Curve Model** (UCLCM)[41,42] and is more exploratory than the previous models, because the researcher may gain insight into what trajectory might be the more appropriate to fit the data. Therefore an UCLCM may be well-suited for irregular trajectories that are of neither linear or quadratic shape and need to be evaluated point by point. To obtain an UCLCM (Fig.3) only two loadings from the slope to the observed values need be fixed: choosing to constrain the first to 0 and the second to 1 a metric of change was set and consequently the other loadings' estimates reflect the cumulate proportion of change experienced until the corresponding timepoint compared to the change occurred between the first two observations.[43] In an UCLCM the same parameters of an ULLCM needs to be estimated, plus the unknown loadings from the third (λ_3, corresponding to the 9-months observation in this study) to the last observed measure (λ_6, corresponding to the 24-months observation in this study).

Figure 3 –Diagram of an Unconditional Completely Latent Curve Model

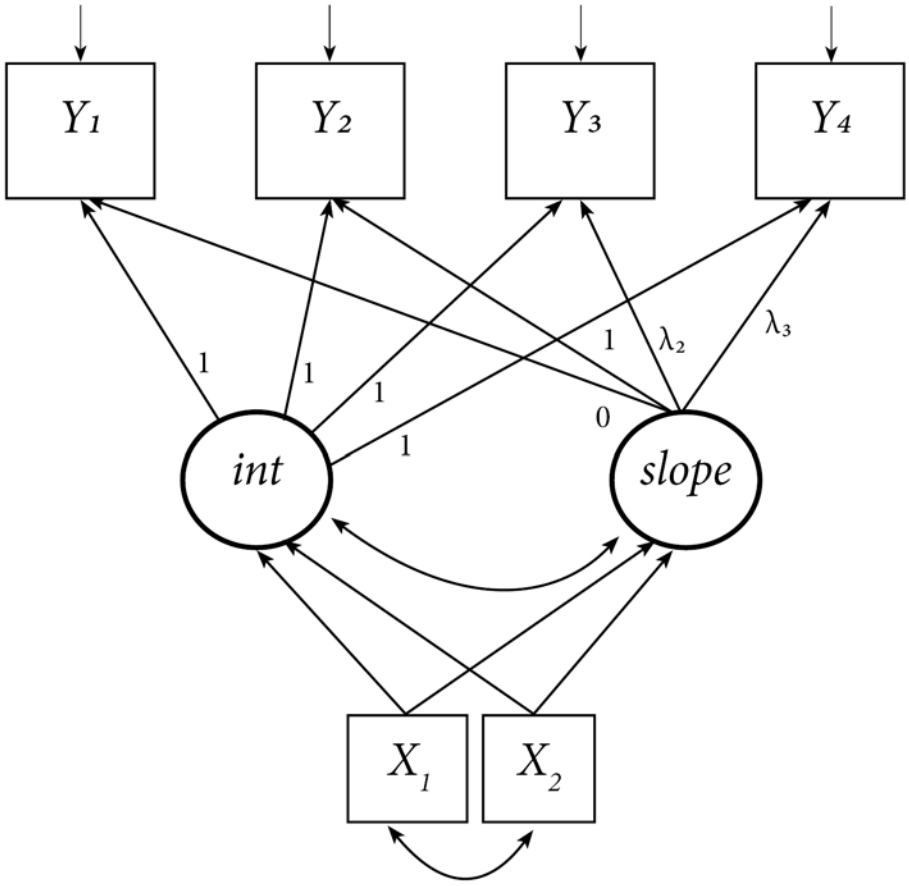


Each of the previously described models may be turned into a **conditional model** by adding at least one exogenous covariate. When the covariate does not change over follow-up time and it may theoretically be associated with at least one of the latent growth terms, then it is defined as a **time-invariant covariate**. This definition reflects that the covariate is a variable that may change among preterms but not over time and thus is supposed to influence just the mean baseline level and the mean growth rate and does not specifically affect the repeated observations. Typical time-invariant covariates are gender, gestational age or ethnicity. The association among time-invariant covariates and a latent growth factor is evaluated as a linear regression, therefore the strength of the association is measured by a regression coefficient and the usual inference on the coefficient's significance is provided. Furthermore, by introducing time-invariant covariates the latent growth factors become dependent variables and consequently the estimation of their proportion of explained variance is made possible. Additional parameters to be estimated in a conditional model are:

- Regression coefficients from covariates to growth factors;

- Residual variances of the growth factors;
- Correlations among the covariates (if specified).

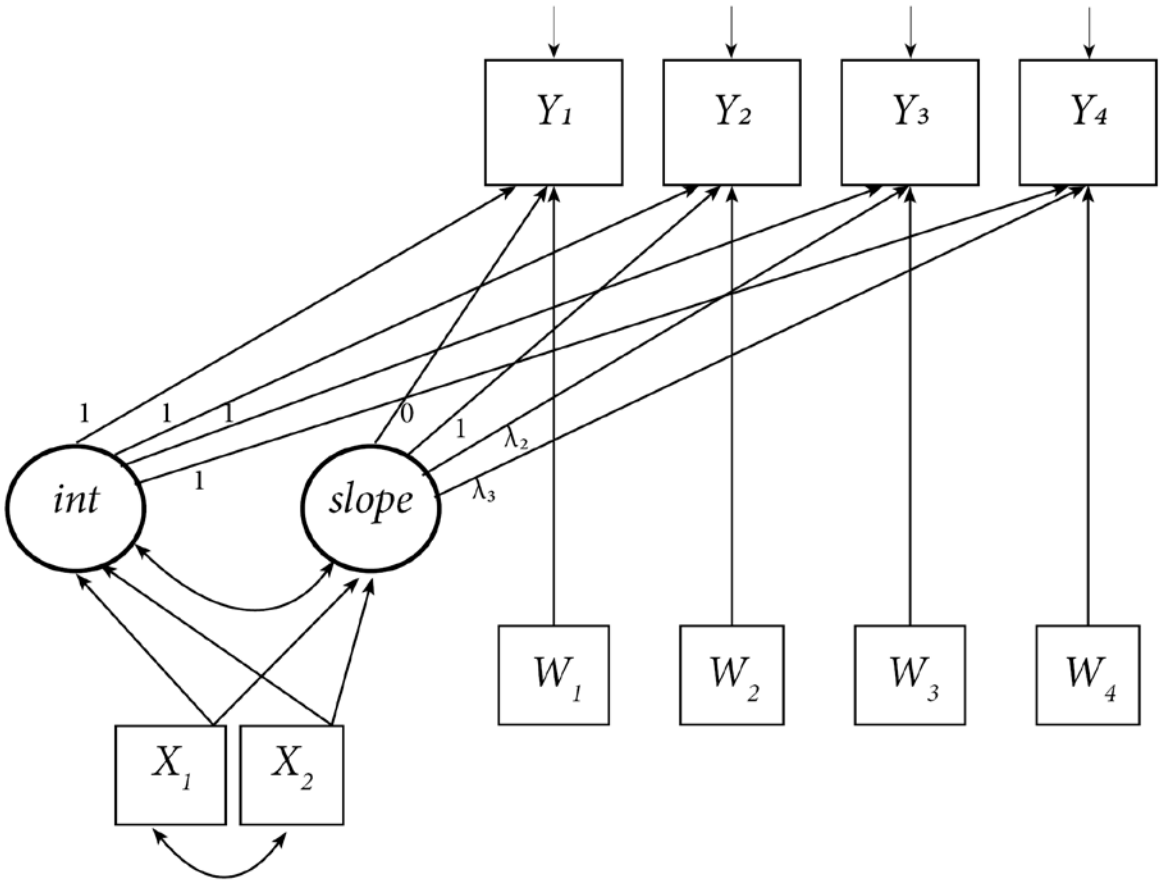
Figure 4 –Diagram of a Conditional Latent Curve Model with time-invariant covariates



Alternatively or in addition to time-invariant covariates, **time-varying covariates** (TVC) may be added to a model. TVCs are variables measured simultaneously to the observed repeated measures, such as cranial circumference or nursery school attendance, and in an LCM are used as exogenous predictors of each corresponding wave of the observed variables. In this way, the explained variation of the observed repeated measures may be substantially improved. Differently from TICs, a TVC may change its effect on the outcome

over time, therefore an interesting result is the assessment of when the effect is significant and when it is stronger. TVCs are allowed to covary with the latent growth terms and with the TICs.

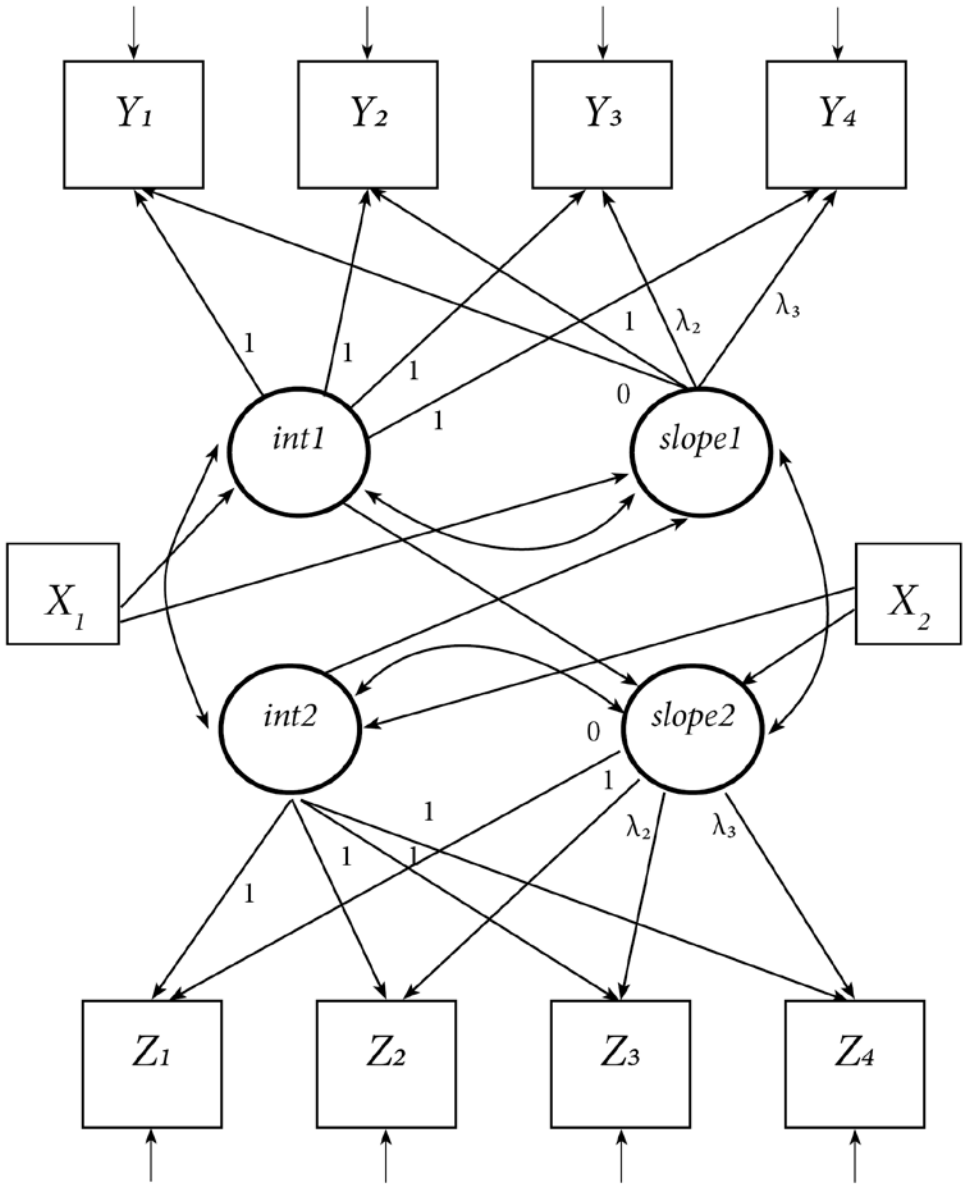
Figure 5 –Diagram of a Conditional Latent Curve Model with Time-varying covariates



Time-varying covariates are repeated measures themselves and rather than considering them as mere covariates it may be more helpful to view them as an outcome as well. In this way, a **Parallel Process Model** or **Multivariate Latent Curve Model** (MLCM) that combines two different growth curves may be defined. The advantage of MLCM is that causal relationships and covariations among the two sets of growth factors may be drawn and tested, allowing to investigate the relationships among aspects of change for different variables.[42] For the evaluation of MLCMs it is advisable to work first on the single variable growth models and to join them only

after good solutions for both have been found. Parameter estimates of the joint model should not differ from those of the single models unless for different sample sizes due to to missing data. The researcher may then decide whether the two processes are simply related by adding covariations among the latent growth terms or they are causally related, by adding regression terms among the latent factors. With this latter choice it is possible for example to assess if the slope of one process may be predicted by the other process' slope or mean initial value, and vice versa. A consequence of this choice is that the two latent intercepts may act as mediator variables between the independent covariates and the latent slopes: for instance, the effect of being SGA on the slope of neurodevelopment may be accounted as the sum of a direct effect and an indirect effect, passing through the initial level of length. It will then be possible to understand in a more accurate way the underlying processes that connect individual characteristics to the neurodevelopment curve.

Figure 6 –Diagram of a Multivariate Latent Curve Model with time-invariant covariates



Refinements of model estimation

Modifications to some of the parameters to be estimated are allowed in order to obtain a better fit or to resolve model identification issues. Means of the observed repeated measures should be constrained to lay on a straight

line (in the case of linear models) or on a parabola (in the case of quadratic models), but to improve model fit some of them may be freely estimated. This technique produces an estimated curve connecting with a straight line or a parabola each point except for those freely estimated, that will show as a bump in the curve. The estimated slope is then referred to the curve connecting all the constrained time points.

Similarly, all observed variables are constrained to have the same variance, but when actually one or some of them differ importantly, this may cause serious model identification problems due to the non-positive definite covariance matrix issue. To resolve this issue it is necessary to let the variance causing problems to be freely estimated, thus obtaining a better fit as well.

Model improvements of this kind are usually suggested by the values of model indices, that are estimates of how much model fit will improve whether a parameter is modified or added to the model. A high modification index is often a sign of model misfit. Typical modification to models are freeing a constrained parameter, adding a causal effect or adding a covariance among variables; however, these changes must always have a theoretical justification.

Another modification to data that is often necessary is related to the variance magnitudes. In LCM and more generally in every SEM model, if the ratio of the largest to the smallest variance of the variables included in the model exceeds 10, the covariance matrix is *ill-scaled* and may determine inaccurate estimates of the model fit, due to the iterative nature of the estimation process.[44] This may very likely happen in conditional models, if continuous measures are evaluated together with dichotomous variables. To overcome this issue, the variable(s) with the higher variance(s) must be divided by a constant, transforming them at a smaller scale until their variance magnitude is comparable to the other variables' variances.

Estimation of LCM and assessment of normality

Maximum Likelihood (ML) is the standard method used to provide LCM estimates. It has several desirable properties, such as consistency, asymptotic unbiasedness, asymptotic normality and asymptotic efficiency. Furthermore, the Full Information Maximum Likelihood (FIML) is recognized as the preferred method to deal

with missing data. The properties of the ML estimator are maintained when the observed variables (the repeated measures of GQ for models of neurodevelopment and the repeated measures of length for models of growth) have the same multivariate kurtosis as a multivariate normal distribution.[45] Should this assumption be violated, biases may occur in the estimates of asymptotic standard errors and of significance test statistics. To verify the null hypotesis of multivariate skewness and kurtosis, tests proposed by Mardia[46,47] were evaluated. In cases of violation of the normality assumption, the robust maximum likelihood (MLR) estimator is suggested; desirable properties of the MLR estimator are the standard errors and χ^2 test statistic robust estimates provided in situations of non-normality, missing data and small to medium sample sizes.[48,49]

Model fit

Being part of the SEM family, Latent Curve Models provide an assessment of the overall model fit to the data. Model fit are summary measures that quantify the adherence of model estimated parameters to the variances, covariances and the means of the observed variables. Not only a good fit is a prerequisite for interpreting parameter estimates, but comparison of model fit is a straightforward criteria to select the more appropriate among alternative models. Several fit indices have been developed and no consensus on a single standard index was reached so far. Therefore, it is advisable to report different fit indices, since they represent different aspects of model fit to the data. For the assessment of a good model fit all reported indexes must have values comprised in the respective ranges of good or acceptable fit, while for comparison among models the best fitting model is identified when it has best values on possibly each fit index. In this study, five fit indices reported by Mplus were used to evaluate model fit:

Comparative Fit Index (CFI) compares the analyzed model with the null model which assumes zero covariances among the observed variables.[50] CFI ranges from 0 to 1 and when it reaches the cutoff value of 0.95 it indicates a good model fit.[51]

Tucker-Lewis Index (TLI) is another index usually reported along with the CFI that compares model fit to the null model fit.[52] It ranges from 0 to 1, with higher values indicating a better fit; a 0.90 cutoff is the least acceptable fit value.

Root Mean Square Error of Approximation (RMSEA) is a fit index with no baseline comparison that measures average lack of fit per degree of freedom.[53] It has no upper limit and a minimum of zero, which indicates the perfect fit. Cutoff values are 0.05 for a good fit and 0.10 for a moderate fit.[54] The advantage of RMSEA is that in addition to the point estimate it provides the 90% confidence interval around its value and a close-fit test for the null hypothesis H_0: RMSEA≤0.05. To ensure a good fit, the confidence interval should have its upper end below 0.8 and the close fit test should not be rejected (p should be >0.05).

Moreover, *Akaike Information Criterion (AIC)* and *Bayesian Information Criterion (BIC)* indexes are used to compare alternative models. These indexes are based on information theory approach and take into account both the goodness of fit and parsimony of a model. Models with smallest AIC or BIC values are those with a relatively better fit and fewer free parameters compared with competing models.

Sample size and missing values

Sample size is a critical issue in SEM and consequently in LCM, because a small sample size may lead to inaccurate estimates.[44] There is no consensus upon a minimum standard sample size, because there are many features of LCM models that need to be taken care of. However, the most followed criteria is to look at the ratio *N/q* between the number of cases and the number of free parameters that needs to be estimated. This rule can be applied when the estimation method is maximum likelihood, but it depends also upon the non-normality of data. The higher the ratio, the better; however, general rules of thumb are that a minimum ratio of 5:1 may be good for normal multivariate data,[55,56] while with strong kurtotic data the ratio should be at least in the order of 10:1[57] and an ideal ratio would be of 20:1[44]. In the results section, along with the estimates, the *N/q* ratio will be provided for each tested model.

The standard method used in statistical inferential analysis to assess sample size adequacy is *power analysis*, but its application to SEM is awkward because it would require many distinct estimations that for complex models may result very challenging.[58] Several approaches have been applied in order to adapt power analysis to SEM, one of which is the test of *not close fit*,[59] based on an inference upon the RMSEA index: if the 95% confidence interval around the index is entirely below 0.05 then the model has a good fit because the hypotesis of not close fit is rejected. With this approach, a higher power value is related to a wider sample size and on a greater number of free parameters to be estimated. Applying the formulae exposed by Hancock and Freeman,[60] an estimate of power based upon the test of not close fit has been provided for the simplest of the proposed models, taking into account that its estimated value of power should be the lower among all models. Estimates of power from 0.80 to 1 indicate an optimal sample size, while values in the range 0.60-0.80 indicate a sufficient sample size.

Some follow-up measures of preterms' GQ score were missing for a number of reason, such as: unavailability of infants and parents at the requested follow-up time, follow-up visits made at an excessively delayed or anticipated time with respect to the scheduled date, impossibility to administer the test for infants' illness. When causes of missingness do not seem to be related with the outcome, missing data may be considered missing at random (MAR) and estimation procedures can be applied to obtain a full data set.

Missing data estimation was made with the FIML, which is a highly reliable estimation method and is run simultaneously to the model estimation procedure. However, in order to limit the possible bias in missing data estimation, infants who were not seen on at least 3 out of the 6 occasions between 3 and 24 months were excluded from the analysis. To evaluate whether this criteria may cause a selection bias, a preliminar representativity analysis was carried out to compare the clinical and socio-demographic characteristics of infants who had at least 3 follow-up visits against those who had only 1 or 2 visits.

Development of Latent Curve Models

The last aim of the study was to analyze the relationship among the preterms' neurodevelopment and length trajectories; to fulfill this aim, which is accomplished by evaluating a very complex model, a series of models progressively more complex have been defined and tested. Neurodevelopment and length latent curves were first estimated separately in order to find the best fitting univariate models and only at a final stage they were combined together in a multivariate LCM. At each stage, criteria to select the best model were the fit indices values and the clinical soundness of the results. Firstly, the most adequate unconditional model was selected by comparing linear, quadratic and completely latent models. The best fitting among these three models was then accrued into a conditional model with the addition of time-invariant and time-varying covariates, chosen among those that theoretically could influence the neurodevelopment or length trajectory. The best fitting model at the end of this stage was considered the best model for the single outcome of neurodevelopment or length. The two resulting models were then combined together in a multivariate LCM, that was nonetheless subject to modifications and alternative formulations, due to the likely different trajectories shapes and interrelations among factors and predictors.

Mplus 7.11 (Muthén & Muthén, Los Angeles, California, USA) was used for the estimation of latent curve models; all the other analyses were carried out using Stata 13.1 (StataCorp LP, College Station, Texas, USA).

RESULTS

Sample characteristics

The study has been carried on preterms born in the S.Orsola Hospital's NICU starting from January, 1th 2005 until June, 30th 2011 and subsequently followed-up until June, 30th 2013, in order to have a potential 24-months follow-up interval for each preterm. The preterms recruited were 346; 22 (6.4%) were excluded from the analysis because they attended less than 3 follow-up visits. The main reasons for missing visits were newborn's illness or family's temporary unavailability. Excluded newborns were significantly different only for a larger cranial circumference at discharge (mean CC was 33.7 cm vs. 32.1 cm; t-test=2.624; p=0.009) and a higher weight at discharge (mean weight was 2359 gr vs. 2097 gr; t-test=2.978; p=0.003). However, the difference in standardized weight at discharge was not significant (mean z-score was -1.269 vs. -1.680; t-test=1.588; p=0.113). Gestational age (29.5 wks. vs. 29.1) and birthweight (1276.8 gr vs. 1161.8) were higher for the excluded newborns but without achieving statistical significance (t-test: p=0.344 and p=0.134 respectively). As a result, newborns excluded because they did not attend at least 3 follow-up visits were tendentially in better conditions but not so much as to lead to a possible selection bias.

Table 1 describes the preterms characteristics. The overall mean gestational age was 29.1 weeks, with the large majority of newborns being very preterm (28<=EG<=31 wks, 66.0%) or extremely preterm (EG<28 wks, 24.4%). The proportion of preterms who were SGA was 17.6%; SGA preterms had a significantly longer gestational age when compared to AGA/LGA preterms (t-test: t=-2.78; p=0.006), because for the sample selection criteria, which also included newborns with EG>32 weeks if they had birthweight under 1500 gr., all late preterms were SGA and 48.0% of moderately preterms were SGA. Complications and comorbidities had an incidence ranging from the 1.6% of ROP to the 20.7% of BPD. Length of stay in the NICU showed a great variability, ranging from a minimum of 6 days to a maximum of 223 days (mean stay was 58.6±34.3 days and the median stay was of 51 days). In bivariate linear regressions each complication proved to be significantly

associated with a longer NICU stay, as shown in Table 2. The type of feeding at discharge was human milk (own mother's or mixed) for around 70%, and at 3 months it was 76.6%.

Tab.1 – Characteristics of the study sample (n=324)

PERINATAL AND CLINICAL CHARACTERISTICS	**N (%)**	**Mean ± Std. Dev**	**Median ± IRQ**	**Missing data n (%)**
Gestational age (weeks)			29.1±2.4	-
Late preterm (34-36 w.)	6 (1.9)			
Moderately preterm (32-33 w.)	25 (7.7)			
Very preterm (28-31 w.)	214 (66.0)			
Extremely preterm (<28 w.)	79 (24.4)			
Twins	105 (33.0)			6 (1.9)
Females	165 (50.9)			-
SGA	57 (17.6)			-
IVH or LPV	17 (5.3)			-
BPD	67 (20.7)			1 (0.3)
Sepsis	48 (14.9)			1 (0.3)
ROP	5 (1.6)			1 (0.3)
NEC	13 (4.0)			-
Hospitalization (days)			58.6±34.3	4 (1.2)
ELBW (<1000 gr.)	111 (34.3)			-
Weight at birth (gr.)		1161.8 ± 353.0	1191 ± 558	-
Weight at discharge (gr.)		2097.3 ± 365.6	1980 ± 400	8 (2.5)
Weight at birth (z-score)		-0.206 ± 1.00	-0.180 ± 1.44	-
Weight at discharge (z-score)		-1.680 ± 1.17	-1.595 ± 1.48	8 (2.5)
Length at birth (cm.)		37.1 ± 4.4	38 ± 5.8	140 (43.2)
Length at discharge (cm.)		43.8 ± 2.7	44 ± 3	144 (44.4)
Cranial circ. at birth (cm.)		27.1 ± 2.8	27 ± 4	163 (50.3)
Cranial circ. at discharge (cm.)		32.1 ± 1.7	32 ± 2	141 (43.5)
Diet at discharge				-
Own mother's milk	109 (33.6)			
Formula milk	98 (30.2)			
Mixed milk	117 (36.1)			
Diet at 3 months				29 (8.9)
Own mother's milk	51 (17.3)			
Formula milk	226 (76.6)			
Mixed milk	18 (6.1)			

BIOMETRIC CHARACTERISTICS	Mean ± Std. Dev	Median, IQR	Missing data, n (%)
Weight at 3 months (gr.)	5451.6 ± 956.1	5530, 1205	15 (4.6)
Weight at 6 months (gr.)	7063.5 ± 1119.0	7080, 1545	25 (7.7)
Weight at 9 months (gr.)	8212.5 ± 1252.9	8290, 1675	49 (15.1)
Weight at 12 months (gr.)	9106.3 ± 1309.1	9140, 1635	55 (17.0)
Weight at 18 months (gr.)	10351.3 ± 1433.7	10400, 1755	52 (16.0)
Weight at 24 months (gr.)	11535.0 ± 1648.0	11575, 2135	45 (13.9)
Weight at 3 months (z-score)	-0.765 ± 1.48	-0.662, 1.70	15 (4.6)
Weight at 6 months (z-score)	-0.776 ± 1.29	-0.751, 1.65	25 (7.7)
Weight at 9 months (z-score)	-0.706 ± 1.23	-0.637, 1.61	49 (15.1)
Weight at 12 months (z-score)	-0.708 ± 1.18	-0.685, 1.43	55 (17.0)
Weight at 18 months (z-score)	-0.784 ± 1.13	-0.765, 1.43	52 (16.0)
Weight at 24 months (z-score)	-0.725 ± 1.15	-0.714, 1.57	45 (13.9)
Length at 3 months (cm.)	58.3 ± 3.4	58.5, 5	45 (13.9)
Length at 6 months (cm.)	65.4 ± 3.4	65.5, 4.3	46 (14.2)
Length at 9 months (cm.)	70.3 ± 3.3	70.5, 4.3	47 (14.5)
Length at 12 months (cm.)	74.4 ± 3.3	74.5, 4.5	57 (17.6)
Length at 18 months (cm.)	80.7 ± 3.6	81, 4.5	52 (16.0)
Length at 24 months (cm.)	86.3 ± 3.5	86.5, 4.2	46 (14.2)
Length at 3 months (z-score)	-1.000 ± 1.61	-0.850, 2.00	45 (13.9)
Length at 6 months (z-score)	-0.825 ± 1.48	-0.667, 2.02	46 (14.2)
Length at 9 months (z-score)	-0.452 ± 1.32	-0.292, 1.67	47 (14.5)
Length at 12 months (z-score)	-0.362 ± 1.28	-0.280, 1.60	57 (17.6)
Length at 18 months (z-score)	-0.433 ± 1.24	-0.352, 1.55	52 (16.0)
Length at 24 months (z-score)	-0.400 ± 1.12	-0.339, 1.33	46 (14.2)
Cranial circ. at 3 months (cm.)	40.2 ± 1.7	40.4, 2.2	17 (5.2)
Cranial circ. at 6 months (cm.)	43.1 ± 1.9	43.2, 2.4	46 (14.2)
Cranial circ. at 9 months (cm.)	44.9 ± 1.9	45, 2.2	48 (14.8)
Cranial circ. at 12 months (cm.)	46.1 ± 1.9	46.1, 2.4	55 (17.0)
Cranial circ. at 18 months (cm.)	47.3 ± 1.9	47.3, 2.3	55 (17.0)
Cranial circ. at 24 months (cm.)	48.2 ± 1.9	48.2, 2.3	46 (14.2)
Cranial circ. at 3 months (z-score)	-0.198 ± 1.35	0, 1.67	17 (5.2)
Cranial circ. at 6 months (z-score)	-0.272 ± 1.42	-0.167, 1.81	46 (14.2)
Cranial circ. at 9 months (z-score)	-0.299 ± 1.37	-0.154, 1.77	48 (14.8)
Cranial circ. at 12 months (z-score)	-0.273 ± 1.36	-0.231, 1.85	55 (17.0)
Cranial circ. at 18 months (z-score)	-0.457 ± 1.34	-0.357, 1.77	55 (17.0)
Cranial circ. at 24 months (z-score)	-0.517 ± 1.31	-0.407, 1.74	46 (14.2)

DEVELOPMENTAL CHARACTERISTICS	Mean ± Std. Dev	Median, IQR	Missing data, n (%)
Griffiths raw score at 3 months	53.2 ± 7.8	54, 9	14 (4.3)
Griffiths raw score at 6 months	98.3 ± 12.9	100, 15	15 (4.6)
Griffiths raw score at 9 months	137.6 ± 13.6	139, 15	14 (4.3)
Griffiths raw score at 12 months	169.6 ± 15.6	172, 14.5	16 (4.9)
Griffiths raw score at 18 months	217.1 ± 22.0	220, 22	35 (10.8)
Griffiths raw score at 24 months	251.5 ± 19.9	256, 17	35 (10.8)
Griffiths score at 3 months	113.1 ± 10.6	115, 10	14 (4.3)
Griffiths score at 6 months	107.1 ± 12.3	109, 13	15 (4.6)
Griffiths score at 9 months	107.1 ± 12.6	108, 16	14 (4.3)
Griffiths score at 12 months	102.5 ± 12.7	104.5, 14	16 (4.9)
Griffiths score at 18 months	94.1 ± 14.5	96, 17	35 (10.8)
Griffiths score at 24 months	93.9 ± 15.1	97, 18	35 (10.8)

OTHER CHARACTERISTICS	N (%)	Mean ± Std. Dev	Missing data, n (%)
Number of siblings		0.3 ± 0.6	2 (0.6)
Age of mothers		33.8 ± 5.4	3 (0.9)
Educational level of mother			16 (4.9)
low	60 (19.5)		
intermediate	146 (47.4)		
high	102 (33.1)		
Educational level of father			23 (7.1)
low	71 (23.6)		
intermediate	135 (44.8)		
high	95 (31.6)		
Hollingshead Index		35.1 ± 10.7	15 (4.6)
Migrant mothers	78 (24.1)		1 (0.3)
Nursery school attendance at 12 months	6 (2.5)		86 (26.5)
Nursery school attendance at 18 months	32 (13.1)		80 (24.7)
Nursery school attendance at 24 months	73 (33.2)		104 (32.1)

Tab.2 Bivariate linear regressions of length of stay in NICU (days) on the presence of complications

	Constant	b	std err. (b)	p	N
Intra-Ventricular Hemorrhage / Periventricular Leukomalacia	56.604	38.043	8.29	<0.001	319
Mechanical Ventilation	45.634	50.695	3.36	<0.001	320
Broncho Pulmonary Displasia	46.635	56.246	3.49	<0.001	319
Sepsis	51.782	44.301	4.75	<0.001	319
Retinopathy of Prematurity	57.350	70.050	14.93	<0.001	319
Necrotizing Enterocolitis	56.189	59.965	9.13	<0.001	320

Patterns of change of biometric and neurodevelopmental measures

Change in the main biometric measures (weight, length and cranial circumference (CC)) from birth to discharge show a relevant increase in absolute values and a corresponding large decrease in variability (as observed with median ± IQR that are less sensible to extreme values). This was mainly caused by SGA, who gained on average 1010 gr. and whose IQR reduced from 547 to 300 gr, compared with a corresponding average weight increase of 791 gr and IQR 158 gr reduction for AGA-LGA. However, such a result may be mainly attributable to the significantly longer NICU stay of SGA newborns (71.2 vs. 56.0 days, t-test=-3.01, p=0.003).

As for the socio-economic characteristics of newborns' family environment, infants in our study were generally firstborns (77.0%) and only 4.3% had two or more elder siblings. Mean and median age of mothers was around 34 years, with a wide variation spanning from 18 to 47 years. Infants born from migrant mothers were 78, with 31 different foreign countries of origin; the more frequent of these were Eastern European countries (Romania, Moldavia, Albania), Nigeria and some Asian countries (Bangladesh, Pakistan). The proportion of preterms with a migrant mother (24.1%) is not much higher than that of newborns in the Bologna province in the years 2005-2011 (23.3%)[2]. Compared to infants born from Italian mothers, a higher proportion of preterms from migrant mothers had siblings (30.8% vs. 20.8%) and their mean mothers' age was lower (31.4 vs. 34.5 years). Education

[2] Regione Emilia-Romagna, Statistica self-service, retrieved on 15.01.2014.

level of mothers and fathers was intermediate-high, since a low-level of education was reported by about 20% of parents. Among migrant mothers there was a lower proportion of graduates (25.6% vs 33.5%) and a higher proportion of mothers whose educational level was unknown or elementary school level (14.1% vs 3.7%). The Hollingshead Index (HI) was significantly lower when the mother was migrant compared to when the mother was Italian (mean HI 29.1 vs. 37.2; t-test: $t=6.21$; $p<0.001$). There was no association among HI and firstborn condition of the infants: mean HI for firstborn's parents was 35.2 compared with 35.0 for non-firstborn's parents (t-test: $t=0.158$; $p=0.875$), while there was a moderately significant positive correlation ($r=0.295$; $p<0.001$) between HI and mother's age. Infants' attendance of nursery school began at the age of 12 months (6 infants, 2.5% of the total), with a growing proportion of attendance at 18 months (12.4%) and at 24 months (33.2%). Nursery school attendance did not differ significantly between infants born from migrant and non-migrant mothers, both at 18 months (19.2% vs. 11.5%; $\chi^2=2.13$, $p=0.145$) and at 24 months (41.3% vs. 31.2%; $\chi^2=1.66$, $p=0.197$). Similarly, HI did not differ significantly between newborns attending and not attending nursery school, both at 18 months (37.5 vs 35.1; t-test=-1.16, $p=0.246$) and at 24 months (36.6 vs. 34.4; t-test=-1.43, $p=0.155$).

Biometric characteristics measured at follow-up waves (weight, length and cranial circumference at 3, 6, 9, 12, 18 and 24 months corrected age) all show a positive growth pattern in absolute values but with regard to standardized scores the patterns were different. Weight in absolute value increases at a high rate in the first followup interval and then progressively decelarates. The individual observed trajectories of weight represented in Figure 7 follow an increasing slightly curvilinear pattern, while in the individual observed trajectories of standardized weight (Figure 8) a 'funnel' pattern from 3 to 6 months may be observed, related to a reduction in scores' variability (sd decreases from 1.48 at 3 months to values around 1.2 in all the following waves). However, preterms remain underweight during all follow-up period, with means of weight z-scores floating around -0.75.

Figure 7 –Trajectories of weight between 3 and 24 months corrected age; absolute values' trajectories

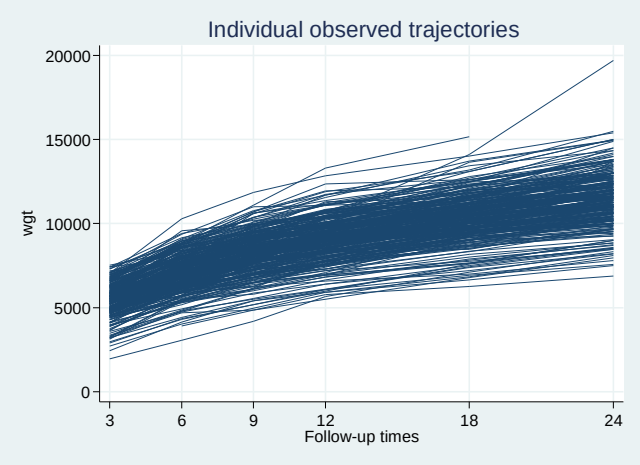
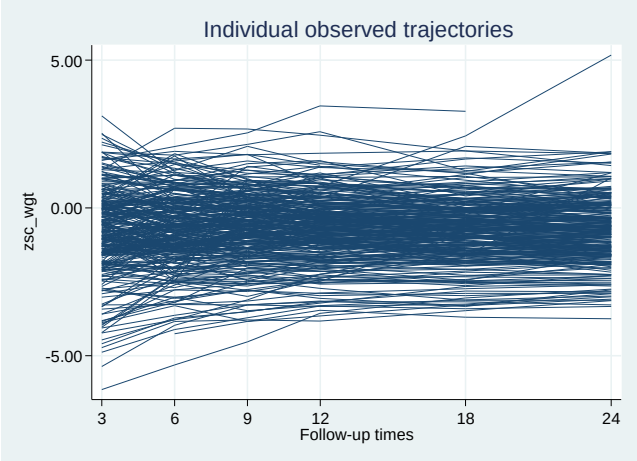


Figure 8 –Trajectories of weight between 3 and 24 months corrected age; standardized scores' trajectories



Length's growth (Figure 9) follows a pattern similar to the weight pattern for absolute values, but with an increase that looks steeper in the first months and a more marked curvilinear form. The corresponding z-scores trajectories (Figure 10) show an initial increase and a reduction in variability followed by a stabilization on negative values around -0.40 after the 9-months visit.

Figure 9 –Trajectories of length between 3 and 24 months corrected age; absolute values' trajectories

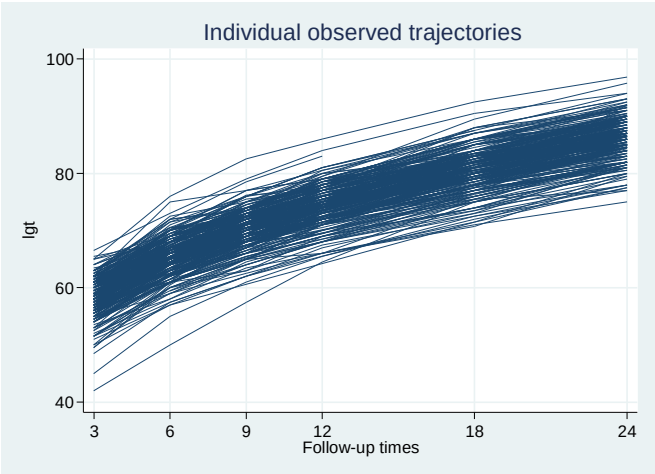
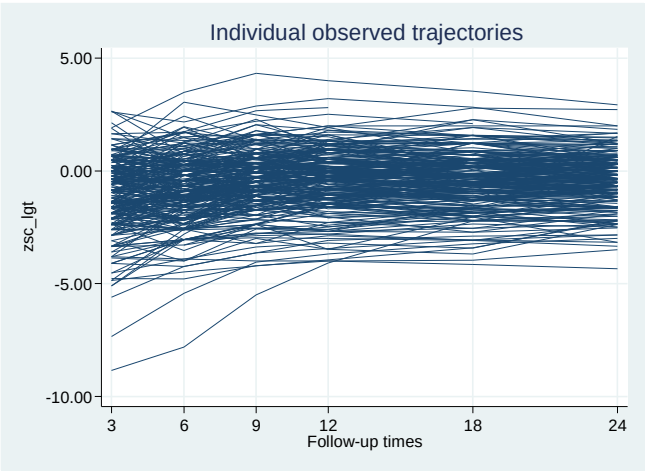


Figure 10 –Trajectories of length between 3 and 24 months corrected age; standardized scores' trajectories



Cranial circumference growth pattern was quite different (Figures 11 and 12). Until the 9-months visit its increase in absolute values was much steeper than the one found for weight and length, but afterwards it changed into a flatter one, thus defining an evident curvilinear shape. The trajectories of cranial circumference

standardized scores were slowly but constantly declining over time, with a change from a mean z-score of -0.198 at 3 months to a mean z-score of -0.517 at 24 months.

Figure 11 –Trajectories of cranial circumference between 3 and 24 months corrected age; absolute values' trajectories

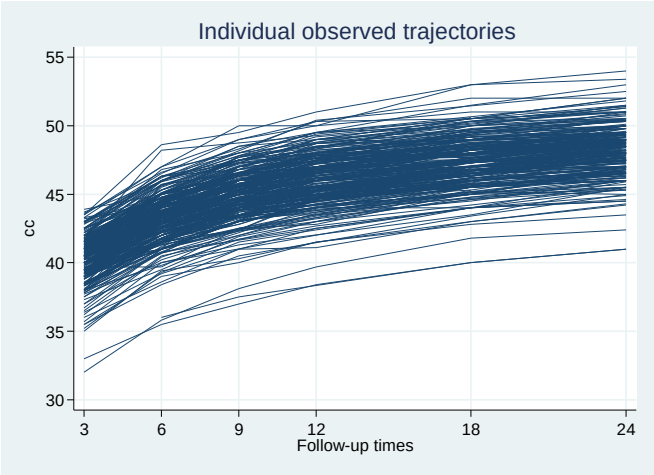
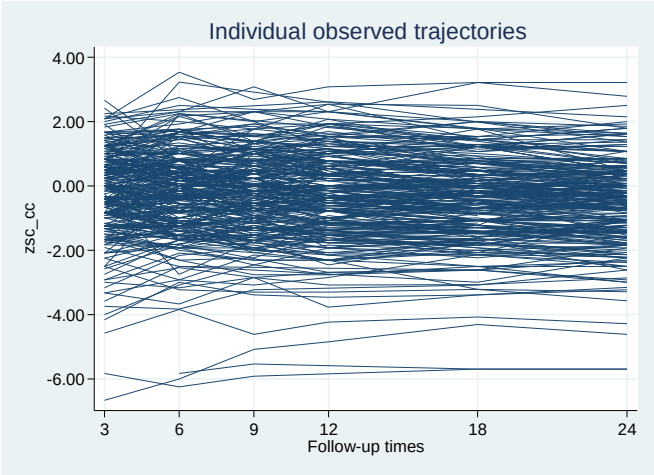


Figure 12 –Trajectories of cranial circumference between 3 and 24 months corrected age; standardized scores' trajectories



Neurodevelopment measured by the Griffiths raw scores (Figure 13) had a constant and slightly curvilinear increase, with most preterms following a similar pattern and a few of them remaining at lower values with a flatter curve. Looking at standard scores (Figure 14), the pattern is that of a constant decrease until 18 months.

Figure 13 –Trajectories of neurodevelopment between 3 and 24 months corrected age; raw scores' trajectories

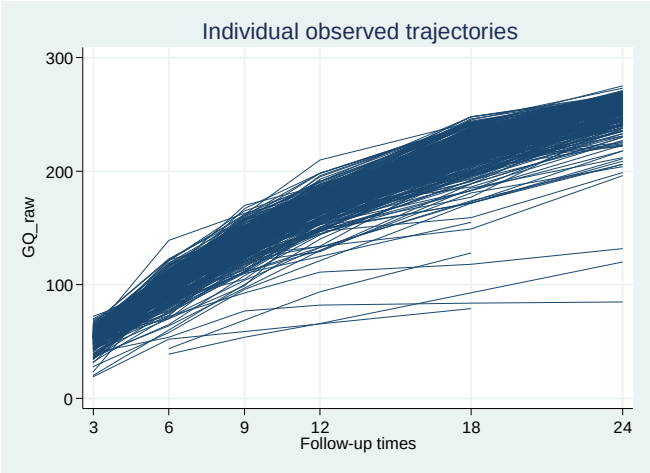
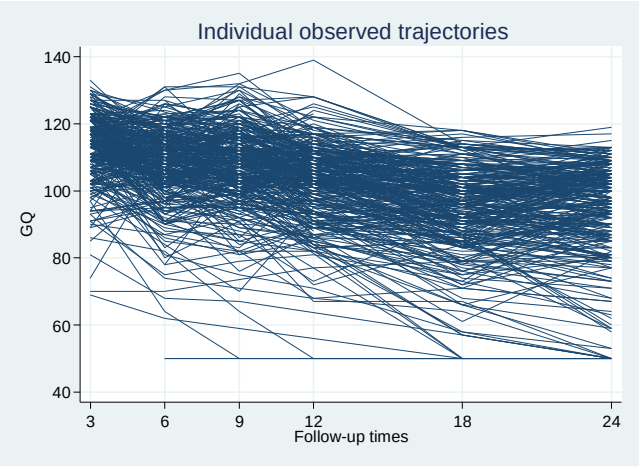


Figure 14 –Trajectories of neurodevelopment between 3 and 24 months corrected age; standard scores' trajectories



Latent curve models

Models based on repeated measures of neurodevelopment (GQ)

The following models are estimated using repeated measures of neurodevelopment taken at 3, 6, 9, 12, 18 and 24 months of age. All neurodevelopment observed variables (GQ3 to GQ24) were rescaled dividing by 10, in order to solve an ill-scaled variances issue. As a consequence, the estimates of the latent intercept should be multiplied by 10 to return to the original GQ scale when interpreting the results.

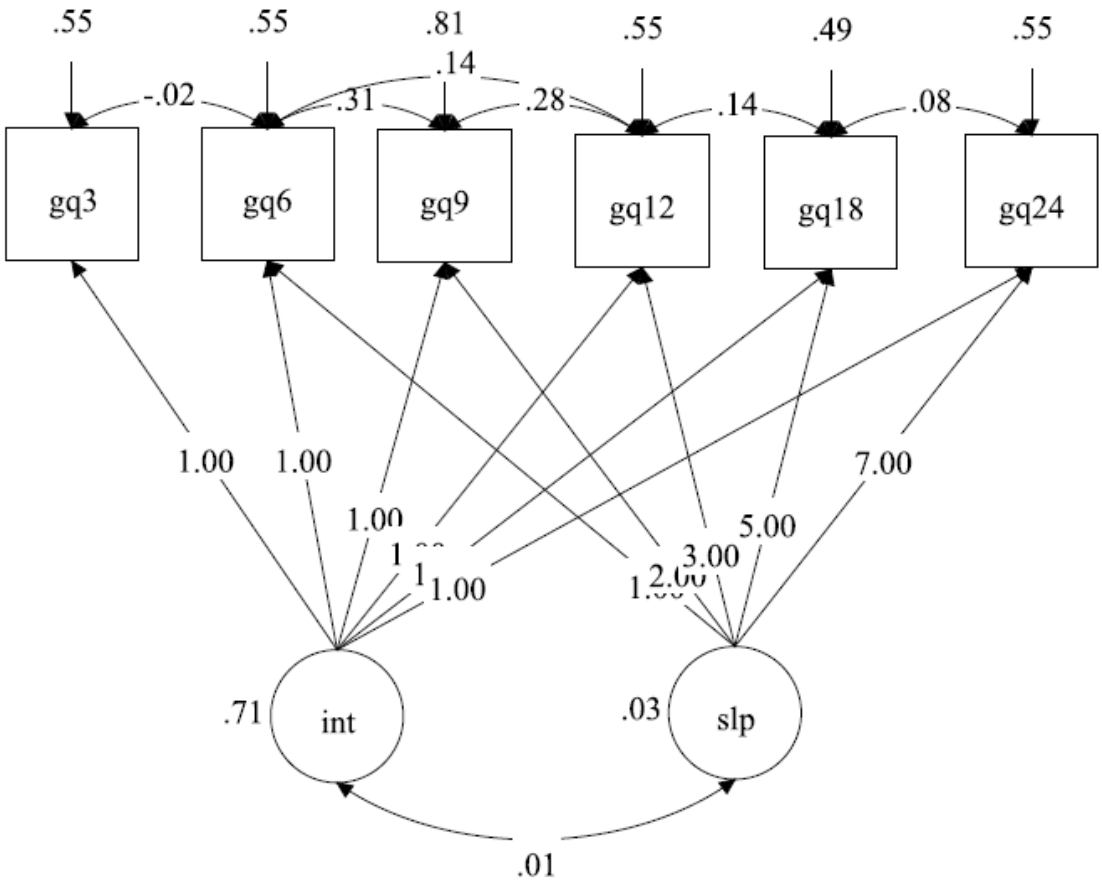
Assessment of normality assumption

The assessment of multivariate normal assumption using Mardia's multivariate skewness and kurtosis tests resulted in the rejection of the null hypotesis (for both $p<0.001$), meaning that for the observed repeated measures of neurodevelopment the assumption of normality was violated. To cope with the possible bias on significance test statistics caused by non-normality, estimates were carried out using the robust maximum likelihood estimator (MLR).

GQ-M1 – Unconditional linear curve model

The simplest LCM model estimated on the repeated measures of neurodevelopment is the unconditional linear model (Fig.15). The GQ-M1 model was defined with six repeated measures of GQ, from 3 to 24 months of age, assuming that observations were associated only with a latent intercept (labelled *int*) and a latent slope (*slp*). The linearity of the model was obtained by setting the loadings from the slope to the observed measures to values proportional to the time interval among measures (time unit is 3 months). Covariances on adjacent repeated measures were taken into account; the variances of GQ9 and of GQ18 were freely estimated to resolve a non-definite positive covariance matrix issue; the intercepts of GQ6 and GQ24 were freely estimated to obtain a better fit.

Figure 15 –Diagram of GQ-M1 model



Model GQ-M1 had a fair fit (RMSEA=0.077; CFI=0.984; TLI=0.979), indicating that a linear model is barely sufficient to explain the neurodevelopment pattern of growth. It gave estimates of the intercept (11.333) and of the slope (-0.386) means both significant at $p<0.001$, as well as the estimates of their variances; the covariance among the intercept and the slope was not significant ($r=-0.075$; $p=0.504$). Explained variance of each of the observed measures was very high, ranging from 0.516 for GQ9 to 0.801 for GQ24.

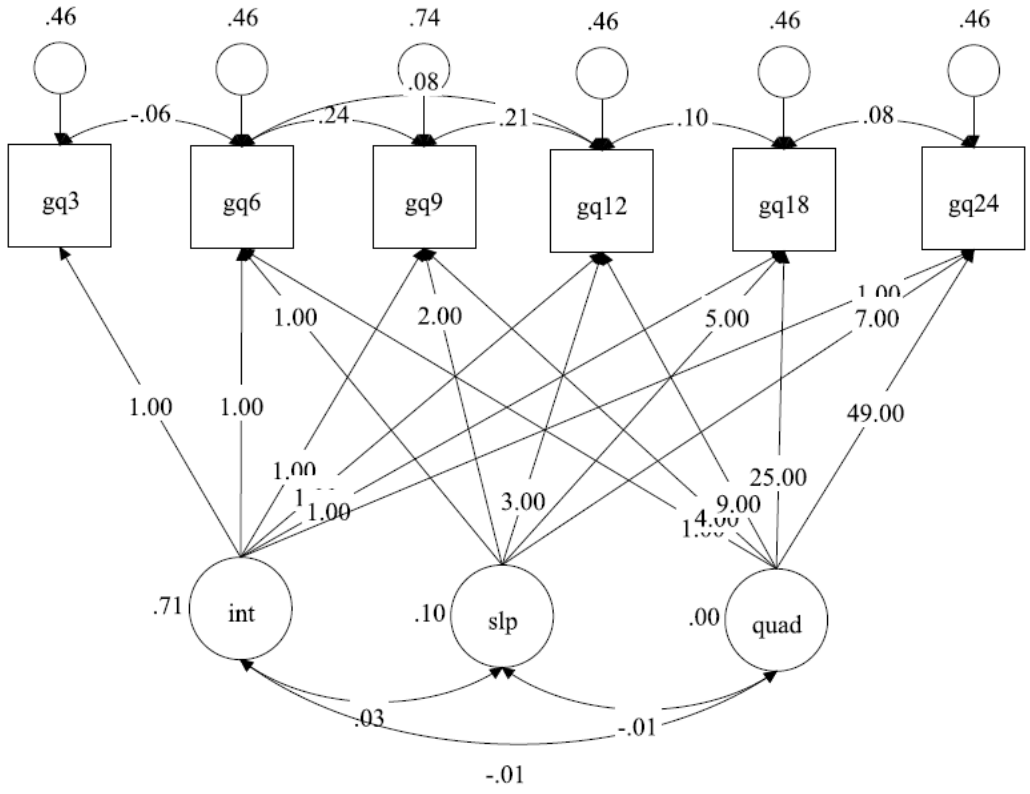
GQ-M2 – Unconditional quadratic curve model

The individual observed trajectories represented in Fig.13 and Fig.14 show that both standardized and raw neurodevelopment scores follow a curvilinear pattern; therefore, the first improvement on GQ-M1 model may be obtained by changing the curve pattern from linear to quadratic. In GQ-M2 model the quadratic term was named

quad; the intercepts of the observed measures at 6 and 18 months were freely estimated to obtain a better fit and GQ18 variance was allowed to be freely estimated in order to resolve a non-positive definite covariance matrix issue.

The fit of the GQ-M2 Model (Fig.16) only slightly improved over the GQ-M1 Model, thus remaining at only a sufficient level (RMSEA=0.072, CFI=0.990, TLI=0.982). The quadratic latent variable had a significant mean (0.021, $p<0.001$), a non-significant variance (ψ_γ=0.001, p=0.088), was significantly associated with the slope (r=-0.860, $p<0.001$) and had no relation with the intercept (r=-0.201, p=0.476). The R^2 of repeated measures increased, from a minimum of 0.591 for GQ9 to 0.812 for GQ24.

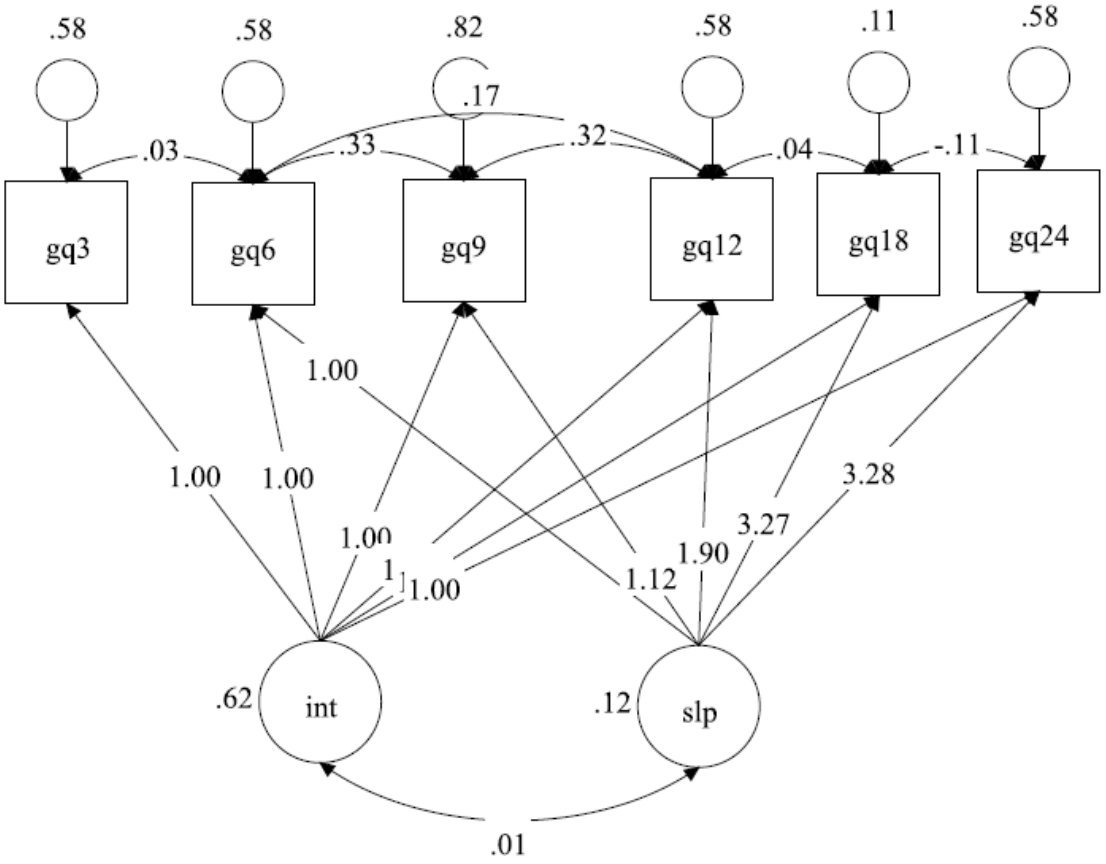
Figure 16 –Diagram of GQ-M2 model



GQ-M3 – Unconditional completely latent trajectory model

The unsatisfactory fit of models GQ-M1 and GQ-M2 indicates that the trajectory of neurodevelopment may be neither linear or quadratic. In fact it may be seen from Fig.14 and from the series of GQ means in Tab.1 that decrease in time of GQ was not constant but occurred especially among 3 and 6 months and among 9 and 18 months; two periods of stability among 6 and 9 months and among 18 and 24 months concurred to define a quite irregular trajectory. Starting from this evidence, a completely latent trajectory model was designed in order to find out a trajectory of GQ where the shape of the longitudinal trend was estimated instead of being specified a priori. The loading from the latent slope to GQ3 was set to 0 and the loading to GQ6 was set to 1 in order to set the metric of change and all the other time points had free loadings. As a consequence, the estimated loadings were interpreted as the amount of change from GQ3 to each time point, scaled relative to the change that was observed between the first two periods. GQ9 and GQ18 variances were freely estimated to resolve non-positive definite covariance matrix issues.

Figure 17 –Diagram of GQ-M3 model



The GQ-M3 model had a much better fit compared to the previous models: RMSEA was 0.038 (90% C.I.: 0.000–0.078), CFI=0.997 and TLI=0.995. The estimated loadings reflected the pattern of change of GQ expressed by the series of means: λ_9=1.12 indicated a change of GQ from 3 to 9 months only 12% higher than the change from 3 to 6 months; λ_{12} was 1.90, indicating a faster change from 9 to 12 months with respect to the previous 3 months; λ_{18} was 3.27 and λ_{24} was 3.28 indicating that a great change happened from 12 to 18 months while from 18 to 24 there was substantially no change in GQ. On the whole period from 3 to 18 months GQ changed 3.27 times as much as from 3 to 6 months. Since the mean of the slope estimate was negative (-0.587; $p<0.001$), higher loadings represent a greater decline of GQ with respect to the reference interval 3-6 months and positive differences among pairs of time points indicate a decline of GQ in corresponding interval: from 9 to 12 months the difference between loadings was 1.90-1.12=0.78, that is in those 3 months GQ declined at a pace that

equalled 78% of that observed among 3 and 6 months. From 18 to 24 months, the difference of 0.01 indicated that GQ scores estimates at 18 and at 24 months were unchanged. The intercept had a mean of 11.297 (corresponding to a GQ score of 112.97) and the correlation among intercept and slope was not significant, showing that at the individual level there was no association between the starting point of GQ and the rate of change. Variances of intercept and slope were both significant, representing high individual variations in starting points and in slopes. Significant correlations were found among each pair of adjacent periods except for the two initial and final intervals; the correlation among GQ6 and GQ12 was found significant and added to the model. Explained variances of the observed values were all quite or very high, ranging from 0.493 to 0.948.

The main characteristics of the first 3 models on neurodevelopment are summarized in Tab.3. Among these unconditional models, the Completely latent (GQ-M3) model was undoubtedly the best fitting because it had the best values on each of the five fit indices. For this reason it was taken as the basis model upon which add covariates to test the effects of clinical and socio-economic characteristics of the preterms on their neurodevelopment trajectories.

Tab.3 Main characteristics of Unconditional linear models on repeated measures of neurodevelopment

		LCM UNCONDITIONAL MODELS		
		GQ-M1 Linear	**GQ-M2 Quadratic**	**GQ-M3 Completely latent**
Model fit	RMSEA	0.077	0.072	0.038
	CFI	0.984	0.990	0.997
	TLI	0.979	0.982	0.995
	AIC	4916.791	4911.125	4900.783
	BIC	4977.283	4982.959	4968.837
Means	Intercept	11.333 (p<0.001)	11.317 (p<0.001)	11.297 (p<0.001)
	Slope	-0.386 (p<0.001)	-0.426 (p<0.001)	-0.587 (p<0.001)
Variances	Intercept	0.710 (p<0.001)	0.710 (p<0.001)	0.622 (p<0.001)
	Slope	0.027 (p<0.001)	0.104 (p=0.009)	0.123 (p<0.001)
Correlation	Intercept-slope	0.075 (p=0.504)	0.102 (p=0.657)	0.029 (p=0.787)

The complete description of GQ-M3 Model statistics are reported in Tab 4:

Tab 4 Parameter estimates, asymptotic standard errors and p-values of unconditional completely latent model for neurodevelopment (GQ-M3 Model), n=324

Parameter	Estimate	Standard Error	p-value
Loadings			
λ_0	0		
λ_1	1		
λ_2	1.115	0.079	<0.001
λ_3	1.898	0.144	<0.001
λ_4	3.270	0.299	<0.001
λ_5	3.282	0.299	<0.001
Variances			
$\psi_{\alpha\alpha}$	0.622	0.140	<0.001
$\psi_{\beta 1\beta 1}$	0.123	0.027	<0.001
Covariance			
$\psi_{\alpha\beta 1}$	0.008	0.029	0.784
Means			
μ_α	11.297	0.059	<0.001
μ_β	-0.587	0.063	<0.001
Residual variances			
VAR(ε_0)	0.577	0.042	<0.001
VAR(ε_1)	0.577	0.042	<0.001
VAR(ε_2)	0.815	0.078	<0.001
VAR(ε_3)	0.577	0.042	<0.001
VAR(ε_4)	0.109	0.107	0.308
VAR(ε_5)	0.577	0.042	<0.001
Fit Statistics			
RMSEA	0.038		0.641[a]
CFI	0.997		
TLI	0.995		
AIC	4900.783		
BIC	4968.837		
Number of free parameters (q)	18		
N/q ratio	18.0		
Estimated power	0.62		

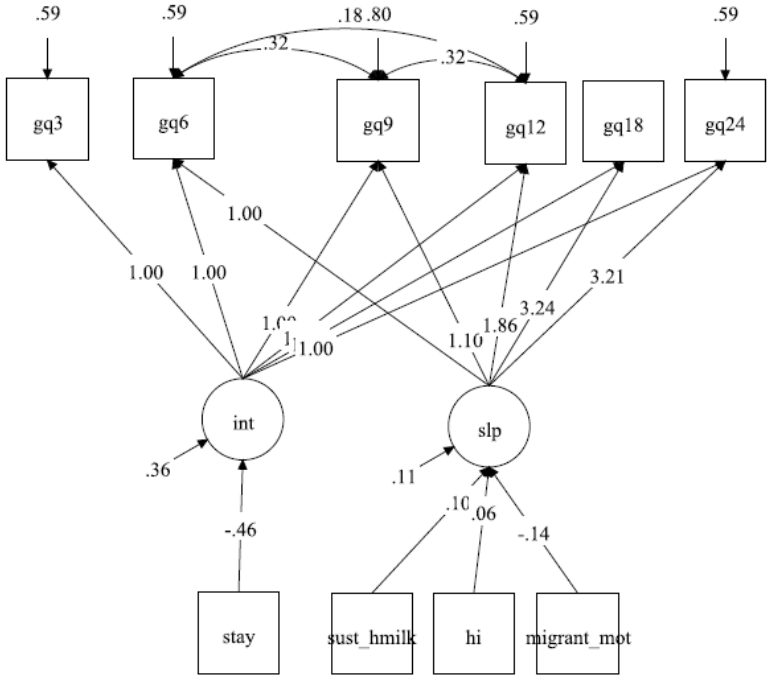
[a] Probability for RMSEA ≤ 0.05

GQ-M4 – Conditional completely latent trajectory model with time-invariant covariates

The GQ-M4 model was built upon the GQ-M3 model by adding some time-invariant covariates. In this way, the GQ-M4 model retained the completely latent trajectory that proved to be the best fitting and added some

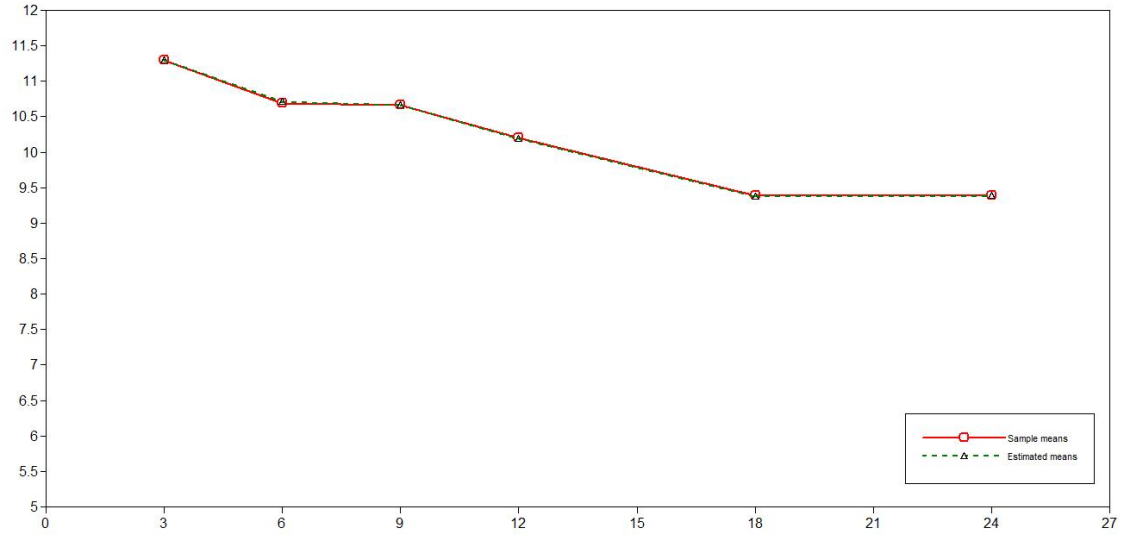
predictors that were associated with the latent intercept and/or slope. The time-invariant predictors to use in the model included ten of the variables described in the Materials and methods Section that on a clinical basis may be associated with the preterms' neurodevelopment trajectory: Gender, Gestational age, SGA, Length of stay in the NICU, Mother's age at birth, Twins, Migrant mother's condition, Hollingshead Index, Diet at discharge and Sustained human milk feeding until 3 months of age. All continuous variables (Gestational age, Length of stay, Mother's age and Hollingshead Index) were centered on their medians to obtain an estimated intercept that could be easily interpretable. Length of stay was expressed in months, by dividing the original variable by 30, and Hollingshead Index was divided by 10 to avoid the ill-scaled covariance matrix issue. As a first step, ten different models with each of these predictors as the only time-invariant covariate of latent intercept and slope were evaluated; each predictor proved to be significant on at least one of the two latent variables at $p<0.200$. Therefore, GQ-M4 Model was built including all predictors and removing one at a time, in decreasing order of p-value, those that were not significant at $p<0.05$ with each of the two latent variables. Starting from the complete model, SGA, Twins, Diet at discharge, Mothers's age at birth, Gender and Gestational age were removed in sequence and hospital stay, Hollingshead Index, Migrant mother and Sustained milk were retained. The final GQ-M4 Model, where each time-invariant covariate is significant at $p<0.05$ on the intercept and/or the slope, is shown in Fig. 18.

Figure 18 –Diagram of GQ-M4 model (only significant paths are shown)



Model fit was very good: RMSEA was 0.022, with p(RMSEA≤0.05=0.927), CFI=0.997 and TLI=0.995; the comparative AIC and BIC indexes were sensibly lower than those found in GQ-M3 model. The excellent fit is also shown in Figure 19, where the two lines connecting the estimated and observed means perfectly overlap.

Figure 19 –Observed and estimated means of neurodevelopment resulting from the GQ-M4 model

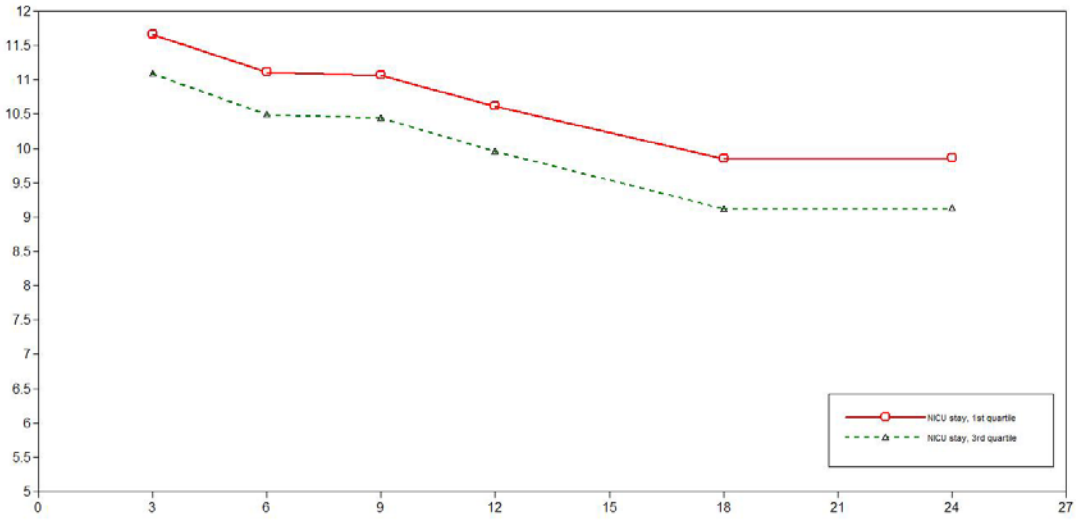


Estimated loadings of time points from 9 to 24 were a little smaller than those obtained with GQ3 Model, but they shared the same trajectory pattern, with the two flat segments among 6 and 9 months and among 18 and 24 months. The intercept latent variable had an estimated mean of 11.452 ($p<0.001$) and a residual variance of 0.378 ($p<0.001$); the intercept slope had a mean of -0.547 ($p<0.001$) and an estimated residual variance of 0.106 ($p<0.001$); the correlation among them was -0.090 and still not significant ($p=0.474$).

Each of the four time-invariant predictors was significantly associated with only the intercept or the slope, therefore yielding two separate groups of covariates: HS was associated with the mean baseline level of neurodevelopment, while Sustained feeding with human milk, Hollingshead Index and Migrant condition of mothers were associated with the mean slope of variation of neurodevelopment on time. Specifically:

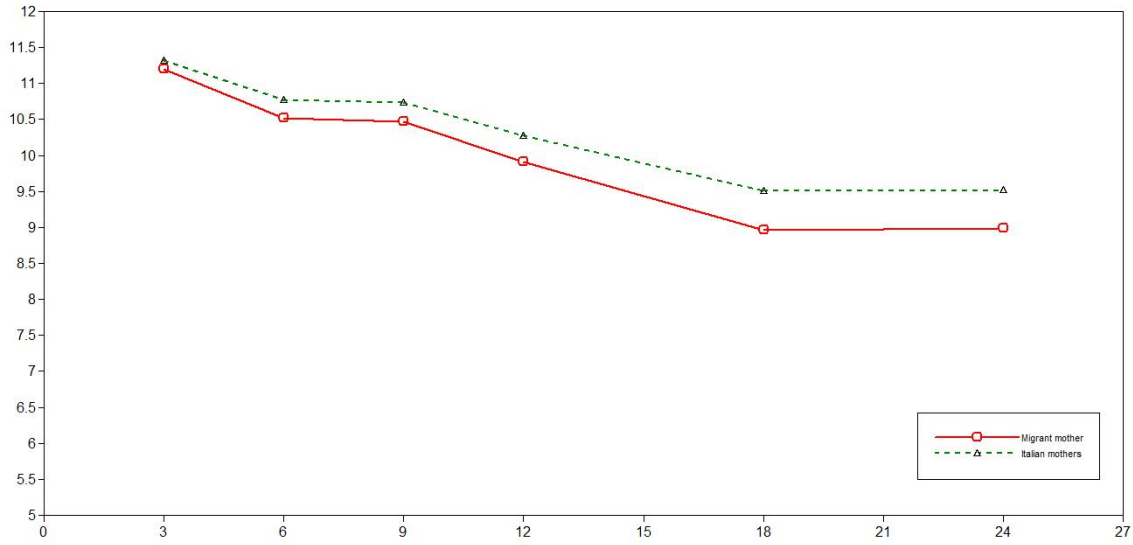
- Length of stay in NICU was negatively associated with the latent intercept, showing that each additional month of stay neurodevelopment mean baseline level score decreased by 4.47 (-0.447x10) ($p<0.001$). In Fig.20 the estimated adjusted trajectory of preterms with 1st quartile stay is compared to those with 3rd quartile stay: since length of stay in NICU affects only the mean baseline level of neurodevelopment, the two curves have a constant offset over all the follow-up interval.

Figure 20 – Comparison of the adjusted estimated neurodevelopment trajectories of preterms in the 1st and 3rd quartiles of NICU stay resulting from the GQ-M4 model



- Hollingshead Index and sustained human milk feeding were associated positively with the latent slope, indicating that in preterms whose parents had higher HI (0.068; p=0.004) or who were fed human milk until 3 months of CA (0.100; p=0.035) the decline in GQ was less steep.
- On the contrary, infants of a migrant mother had a steeper decline in GQ as the association with the slope of this condition was negative (-0.126; p=0.049). In Fig.21 the estimated adjusted trajectory of preterms with migrant mothers is compared to those with Italian mothers: migrant condition of mothers affects only the slope of neurodevelopment, thus the two curves start at the same GQ level and diverge increasingly.

Figure 21 – Comparison of the adjusted estimated neurodevelopment trajectories of preterms with migrant mothers and preterms with Italian mothers



The four covariates together explained a relevant proportion of variance of the latent intercept (R^2=0.407) and a lower proportion of variance of the latent slope (R^2=0.128), that nonetheless was significantly higher than 0 (p=0.008). The sample size was reduced to 305 for a few missing data in the covariates; the number of free parameters to be estimated in the GQ4 model was 26, thus leading to a N/q ratio of 11.7.

Tab 5 Parameter estimates, asymptotic standard errors and p-values of completely latent curve model for neurodevelopment with time-invariant covariates (GQ-M4 model), n=305

Parameter	Estimate	Standard Error	p-value
Loadings			
λ_0	0		
λ_1	1		
λ_2	1.074	0.083	<0.001
λ_3	1.906	0.153	<0.001
λ_4	3.302	0.310	<0.001
λ_5	3.283	0.312	<0.001
Variances			
$\psi_{\alpha\alpha}$	0.378	0.089	<0.001
$\psi_{\beta1\beta1}$	0.106	0.025	<0.001
Covariance			
$\psi_{\alpha\beta1}$	-0.018	0.027	0.503
Means			
μ_α	11.452	0.066	<0.001
$\mu_{\beta1}$	-0.547	0.062	<0.001
Residual variances			
VAR(ε_0)	0.583	0.043	<0.001
VAR(ε_1)	0.583	0.043	<0.001
VAR(ε_2)	0.803	0.080	<0.001
VAR(ε_3)	0.583	0.043	<0.001
VAR(ε_4)	0.139	0.108	0.196
VAR(ε_5)	0.583	0.043	<0.001
Fit Statistics			
RMSEA	0.022		0.927
CFI	0.997		
TLI	0.995		
AIC	4517.281		
BIC	4614.009		
Number of free parameters (q)	26		
N/q ratio	11.7		

[a] Probability for RMSEA ≤0.05

Tab 6 Coefficient estimates and p-values for latent variables regressed on time-invariant covariates (GQ-M4 model), n=305

Predictor variable	Intercept	Slope
Length of stay in NICU	-0.447 (p<0.001)	-0.035 (p=0.156)
Sustained human milk	-0.110 (p=0.344)	0.100 (p=0.035)
Migrant mother	-0.128 (p=0.325)	-0.126 (p=0.049)
Hollingshead Index	-0.036 (p=0.464)	0.068 (p=0.004)

GQ-M5 – Conditional completely latent trajectory model with time-invariant covariates and time-varying covariates

The last series of models tested on the neurodevelopment trajectory were derived from the GQ-M4 model with the addition of a time-varying covariate, nursery school attendance. Only the measures taken at 18 and 24 months of nursery school attendance could be introduced in the model, because at the previous follow-up visits there were almost no infants attending nursery school (6 infants at 12 months). Two different models were considered: the first had two synchronic effects from attendance at 18 months to GQ18 and from attendance at 24 months to GQ24, the second had only a diacronic effect from attendance at 18 months to GQ24. None of these two models brought a fit improvement compared to GQ-M4 model, and the observed values of neurodevelopment at 18 and 24 months were independent from infants' attendance of nursery school.

As a result, the GQ-M4 model may be considered the best model for the evaluation of the neurodevelopment trajectory, since it had a good fit to the data and the results obtained, both those related to the latent variables and those related to the covariates' effects, thoroughly permitted a clinical interpretation.

Models based on repeated measures of length (LGT)

Repeated measures of length were used to model preterms' growth with LCM. Standardized scores of length at 3, 6, 9, 12, 18 and 24 months CA were used. In each tested model, sample size shrinked to 298 cases because length was missing at each follow-up visit for 26 preterms. These preterms with missing data on follow-up lengths had a significantly shorter stay in NICU (42.2 days vs. 60.0 days; t-test=2.46, p=0.014) and a significantly longer gestational age (30.7 weeks vs. 28.9 weeks; t-test=-3.82, p<0.001), thus they may be considered as having better clinical conditions at birth. However, no differences were found on weight at discharge and on socio-economic variables.

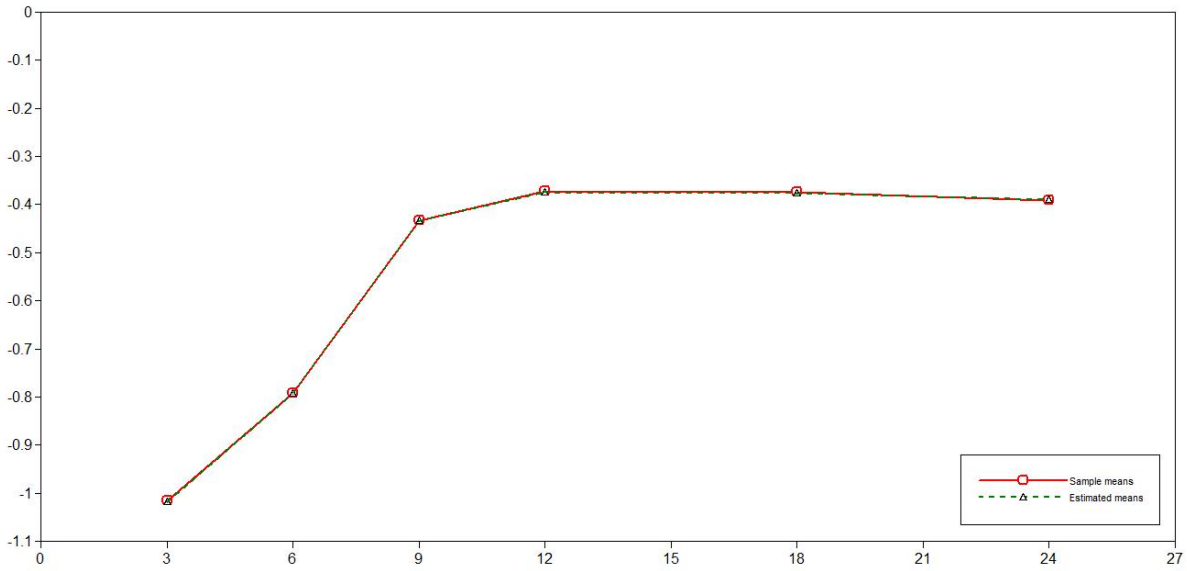
Testing the normality assumption

The assessment of multivariate normal assumption using Mardia's multivariate skewness and kurtosis tests resulted in the rejection of the null hypotesis, meaning that for the observed repeated measures of length the assumption of normality was violated. The results of the univariate tests indicate that skewness was significant at each time point except for 18 months, while kurtosis was significant until the 9 months visits; this means that at earlier time points the distribution of length showed higher randomness, resulting in a greater departure from normality of distribution. To cope with the possible bias on significance test statistics caused by non-normality, estimates were carried out using the robust maximum likelihood estimator (MLR).

LGT-M1 – Unconditional linear curve model

Looking at the trajectories' diagrams of length and standardized length (Figures 9 and 10), a linear curve may seem inadequate to represent the variation of length over time. However, by allowing the means of length at 3, 9, 12 and 24 months to deviate from a straight line, it was possible to obtain a broken line that provided a good fit to the data, as shown by Fig.22.

Figure 22 –Observed and estimated means of length (z-scores) resulting from LGT-M1 model

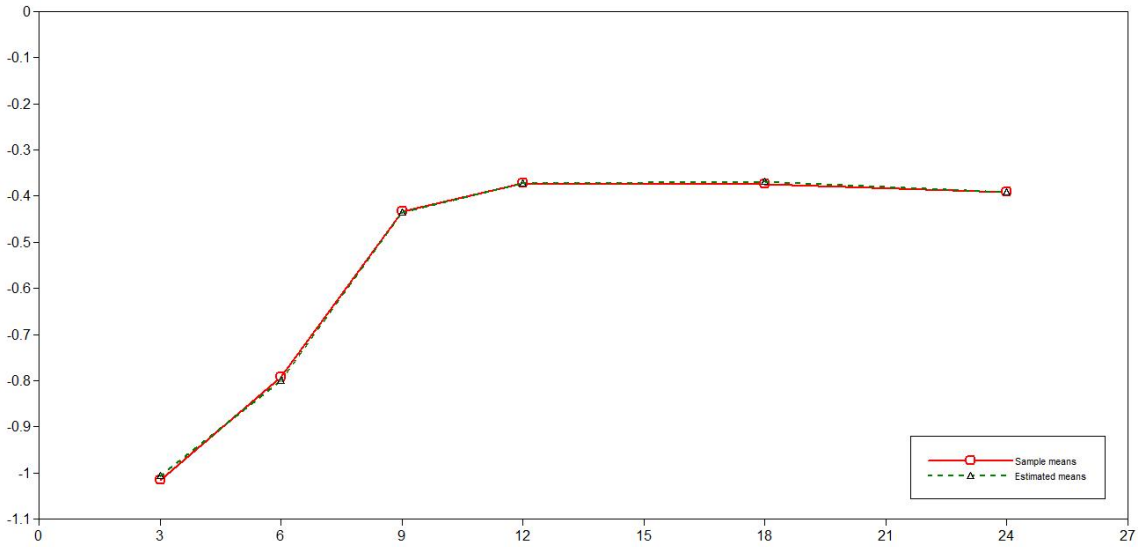


LGT-M1 model had a satisfactory fit: RMSEA=0.049 (90% CI: 0.000-0.095, p=0.450); CFI=0.997; TLI=0.994. Each parameter of the model was significant; the intercept was -0.896 (p<0.001), the slope was 0.104 (p<0.001), the covariance between intercept and slope was -0.115 (p<0.001), intercept and slope variances were respectively 2.011 (p<0.001) and 0.016 (p<0.001). As a result, in the unconditional linear model length was expressed as having a 0.104 standardized score increase every 3 months, with significant variations among individuals in starting points and in slopes and with a negative covariance indicating that those who started at a higher length had a lower growth.

LGT-M2 – Unconditional quadratic curve model

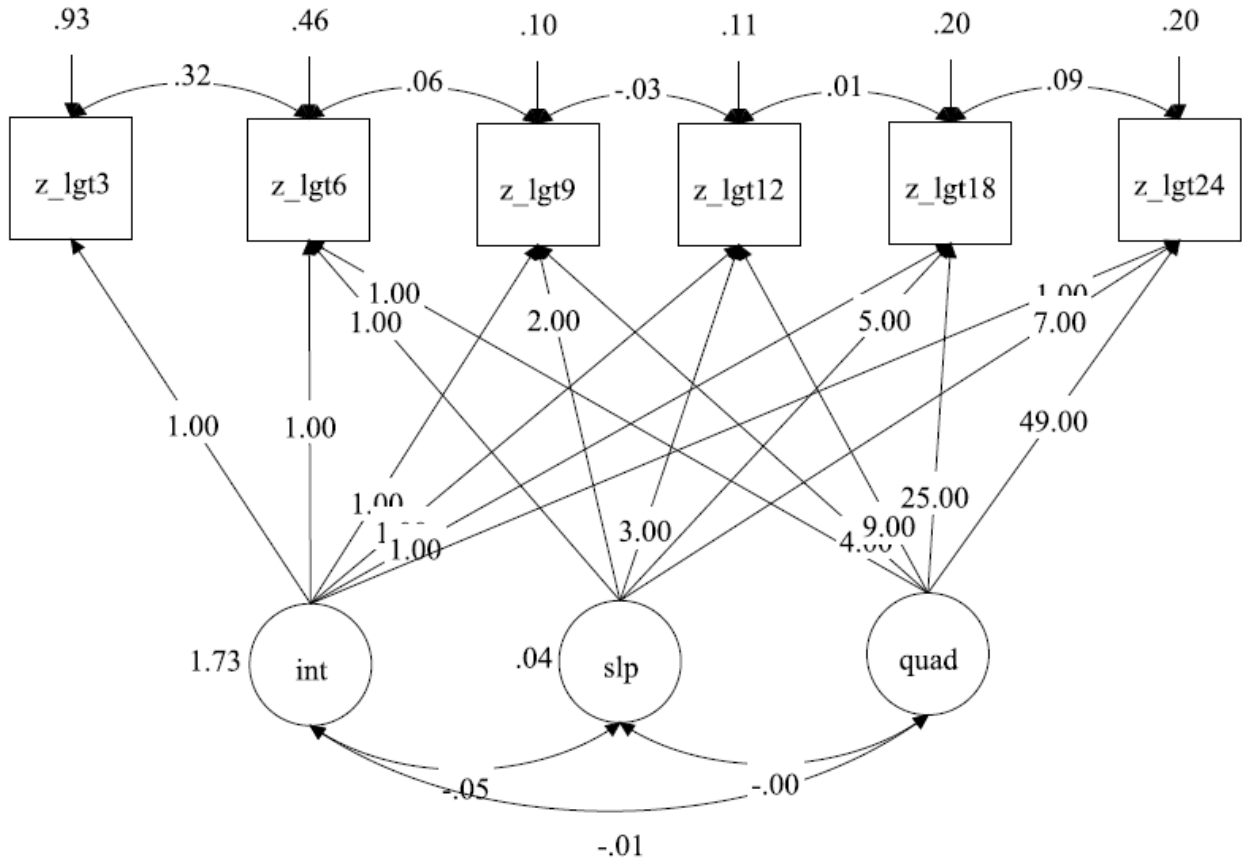
LGT-M2 was built over the LGT-M1 model adding the quadratic latent term. Given its polynomial nature, only standardized lengths means at 9 and 12 means were freed to obtain an optimal fit. The fit obtained was excellent: RMSEA=0.000 (90% CI: 0.000-0.071, p=0.847); CFI=1.000; TLI=1.002 as it can be seen in fig.23 where sample and estimated trajectories are compared.

Figure 23 –Observed and estimated means of length resulting from the LGT-M2 model



Latent intercept, slope and quadratic term were all significant at $p<0.001$. The intercept estimated value at -1.007 was very near to the value of standardized length mean at 3 months that was -1.000, thus the model-implied preterms' length at 3 months was an excellent reproduction of the observed initial value. The estimate of slope at 0.226 indicated a linear component increase in the growth trajectories of 0.226 standardized scores of length every 3 months, and the negative estimate of the quadratic component (-0.020) indicated that the curve increased less steeply as age increases. The three latent factors were not significantly associated, showing that at the individual level there was no relation between starting levels of length, its linear increase and the curvature decrease. Only the intercept had a significant variance, indicating a high variation in individual starting levels of length and no individual differences on slope and quadratic term. The observed values of length had high or very high R^2 values, ranging from 0.650 for length at 3 months to 0.943 for length at 9 months. The LGT-M2 model is represented in Fig. 24.

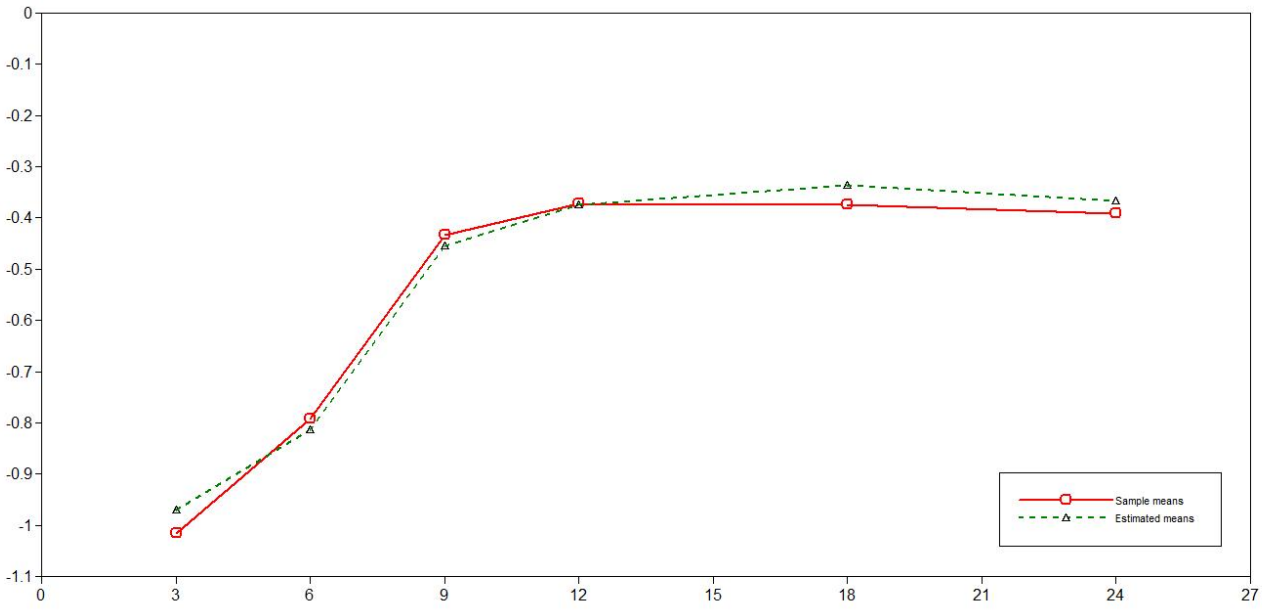
Figure 24 –Diagram of LGT-M2 model



LGT-M3 – Unconditional completely latent trajectory model

Even though LGT-M2 model provided an optimal fit to the data, for the sake of analysis' completeness the unconditional completely latent model was evaluated anyway. It was formulated constraining to 1 the slope at 6 months and letting freely estimated the loadings from time 9 to time 24. To obtain a higher fit it required to free variances of observed values at 3 and at 18 months and the mean of observed values at 24 months. Yet, fit was worse with respect of LGT-M2 and LGT-M1 models, as it is shown by the comparison of sample and estimated means (Fig.25) and by values of RMSEA (0.086, p=0.052), CFI (0.990) and TLI (0.982).

Figure 25 –Observed and estimated means of length resulting from the LGT-M3 model



The main characteristics of the unconditional models on length are summarized in Tab.7. Among these models, the Quadratic LGT-M2 model was chosen as the best fitting because it had better values on each of the fit indexes except for BIC where it was only slightly less fitting than LGT-M1 model. LGT-M2 model was taken as the basis model upon which add covariates to test the effects of clinical and socio-economic characteristics of the preterms on their length growth trajectories.

Tab.7 Main characteristics of Unconditional linear models on repeated measures of length

		LCM UNCONDITIONAL MODELS		
		LGT-M1 Linear	**LGT-M2 Quadratic**	**LGT-M3 Completely latent**
Model fit	RMSEA	0.049	0.000	0.086
	CFI	0.997	1.000	0.990
	TLI	0.994	1.002	0.982
	AIC	3518.612	3512.914	3528.986
	BIC	3592.554	3594.250	3599.230
Means	Intercept	-0.896 (p<0.001)	-1.007 (p<0.001)	-0.970 (p<0.001)
	Slope	0.104 (p<0.001)	0.226 (p<0.001)	0.157 (p=0.003)
	Quadratic term		-0.020 (p<0.001)	
Variances	Intercept	2.011 (p<0.001)	1.730 (p<0.001)	2.215 (p<0.001)
	Slope	0.016 (p<0.001)	0.040 (p=0.270)	0.030 (p=0.144)
	Quadratic term		0.000 (p=0.393)	
Covariances	Intercept-slope	-0.115 (p<0.001)	-0.046 (p=0.601)	-0.175 (p=0.015)
	Intercept-Quadratic term		-0.007 (p=0.414)	
	Slope-Quadratic term		-0.003 (p=0.343)	

Statistics of LGT2 Model are reported in Tab 8.

Tab 8 Parameter estimates, asymptotic standard errors and p-values of unconditional quadratic curve model for length (LGT-M2 Model), n=298

Parameter	Estimate	Standard Error	p-value
Variances			
$\psi_{\alpha\alpha}$	1.730	0.317	<0.001
$\psi_{\beta1\beta1}$	0.040	0.036	0.270
$\psi_{\beta2\beta2}$	0.000	0.000	0.393
Covariance			
$\psi_{\alpha\beta1}$	-0.046	0.087	0.601
$\psi_{\alpha\beta2}$	-0.007	0.008	0.414
$\psi_{\beta1\beta1}$	-0.003	0.004	0.343
Means			
μ_{α}	-1.007	0.091	<0.001
$\mu_{\beta1}$	0.226	0.028	<0.001
$\mu_{\beta2}$	-0.020	0.003	<0.001
Residual variances			
VAR(ε_0)	0.931	0.148	<0.001
VAR(ε_1)	0.455	0.059	<0.001
VAR(ε_2)	0.098	0.041	0.017
VAR(ε_3)	0.106	0.046	0.022
VAR(ε_4)	0.197	0.060	0.001
VAR(ε_5)	0.197	0.451	0.661
Fit Statistics			
RMSEA	0.000		0.847[a]
CFI	1.000		
TLI	1.002		
AIC	3512.914		
BIC	3594.250		
Number of free parameters (q)	22		
N/q ratio	13.5		

[a] Probability for RMSEA ≤0.05

LGT-M4 –Quadratic curve model with time-invariant covariates

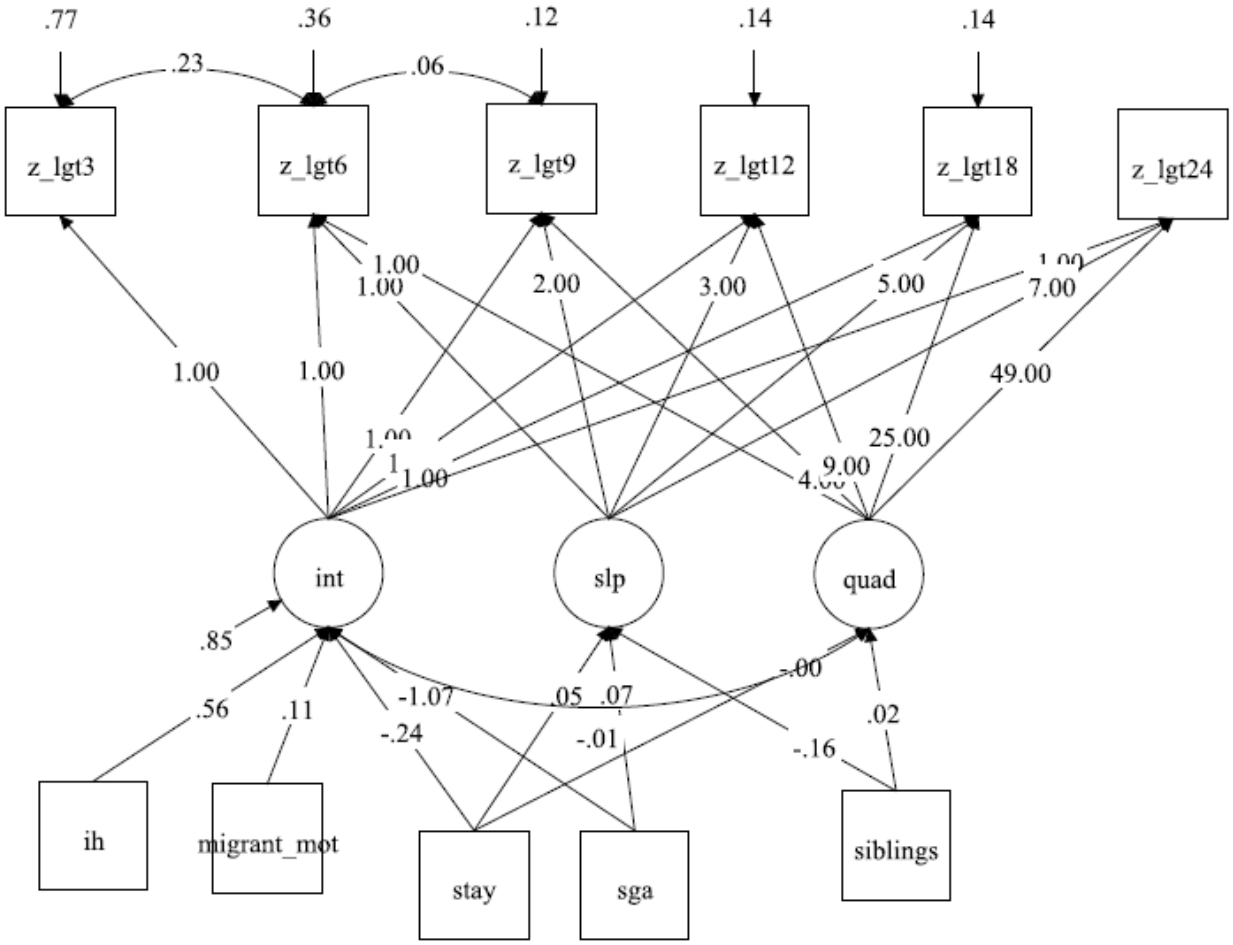
The LGT-M4 model was built upon the LGT-M3 model by adding some time-invariant covariates. In this way, the LGT-M4 model retained the quadratic trajectory that proved to be the best fitting among the unconditional models and added some predictors that were associated with the latent intercept and/or slope. The time-invariant predictors to use in the model included eleven variables described in the Materials and methods Section that on a clinical basis may be associated with the preterms' length trajectory: Gender, GA, SGA, HS, Mother's age at birth, Siblings, Migrant mother's condition, HI, Diet at discharge, Diet at 3 months CA, Sustained human milk

feeding until 3 months CA. All continuous variables (GA, HS, Mother's age and HI) were centered on their medians to obtain an estimated intercept that could be related to logical covariates' values. Length of stay was expressed in months, by dividing the original variable by 30, and HI was divided by 10 to avoid the ill-scaled covariance matrix issue. As a first step, eleven different models with each of the selected predictors as the only time-invariant covariate were evaluated; eight of them proved to be significant on at least one of the three latent variables at $p<0.20$: GA, mother's age, HI, Migrant mother's condition, Siblings, SGA, Diet at discharge, HS. Therefore, LGT-M4 model was built including all these eight predictors and removing one at a time, in decreasing order of p-value, those that were not significant at $p<0.05$ with each of the three latent variables. Starting from the complete model, GA, Mothers's age at birth and Diet at discharge were removed in sequence. At the end of this process, non significant effects at $p<0.05$ from the covariates to the latent variables were removed, in order to obtain a more parsimonious model. Sample size was further reduced to 282 cases due to some missing data on the covariates. The final LGT-M4 Model, where each time-invariant covariate is significant at $p<0.05$ on the intercept and/or the slope, is shown in Fig. 26.

Model fit was practically perfect: RMSEA was 0.000 (90% CI: 0.000-0.000; p=1.000), CFI=1.000 and TLI=1.009. These indexes have values similar than those found in LGT-M3 model, but comparative information-based indexes were sensibly smaller (AIC=3151.572, BIC=3253.476), showing that LGT-M4 model had a much better fit. All the latent variables had significant means at $p<0.001$: the intercept was -0.727, the slope was 0.213 and the quadratic term was -0.022. Therefore, similarly to the unconditional model the estimated growth curve had a positive linear component, indicating an increase of length over time, and a negative quadratic component showing that the curve was progressively reducing its increase rate. The latent intercept and slope were not associated (r=0.410; p=0.233); the slope and the quadratic component were negatively correlated (r=-0.817, $p<0.001$), indicating that preterms with a higher linear increase were those whose curve flattened the most; lastly, the intercept and the quadratic component were negatively correlated with borderline significance (r=-0.688; p=0.082). All the observed variables had very high R^2 values, ranging from 0.704 for

length at 3 months to 0.961 for length at 24 months. The latent variables had moderate to good R^2 values: 0.534 for the intercept, 0.493 for the slope and 0.327 for the quadratic term.

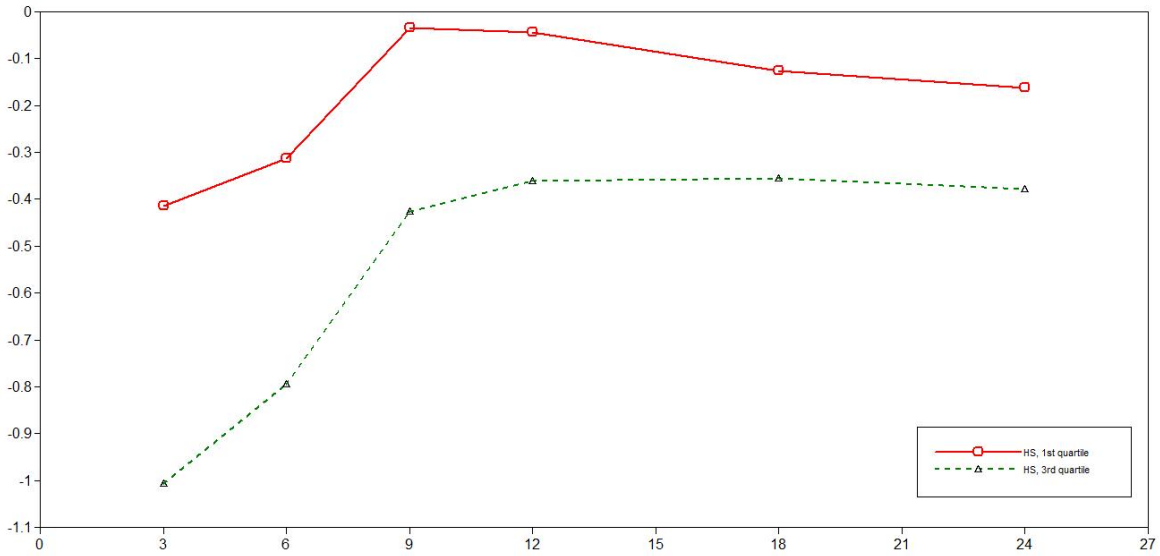
Figure 26 –Diagram of LGT-M4 model



Among the five covariates, Stay (HS) was significantly associated (at $p<0.001$) with each latent variable. It affected negatively the intercept and the quadratic term and influenced positively the slope, indicating that, after adjusting for the other covariates, preterms who had longer stays in the NICU were those who started at a lower length but then experienced a faster growth and in addition had a steeper recovery. In fact, a negative effect of the covariate on the quadratic term multiplied by the negative estimate of the mean quadratic term produces an upward growth curve in its later part. Figure 27 shows that preterms with an HS in the 1st quartile recovered partially their gap in length with respect to preterms who were in the 3rd quartile of HS and that their growth

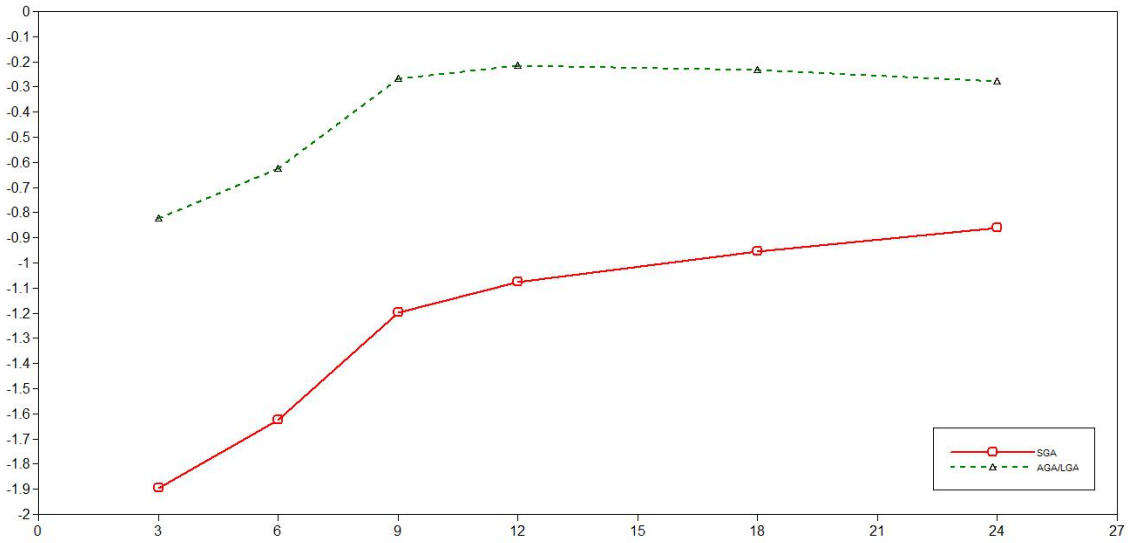
curve was still slightly increasing in the 9-24 months interval, while the estimated trajectory for preterms in the 3rd quartile of HS is decreasing after 9 months of corrected age.

Figure 27 – Comparison of the adjusted estimated length trajectories of preterms in the 1st and 3rd quartiles of NICU stay resulting from the LGT-M4 model



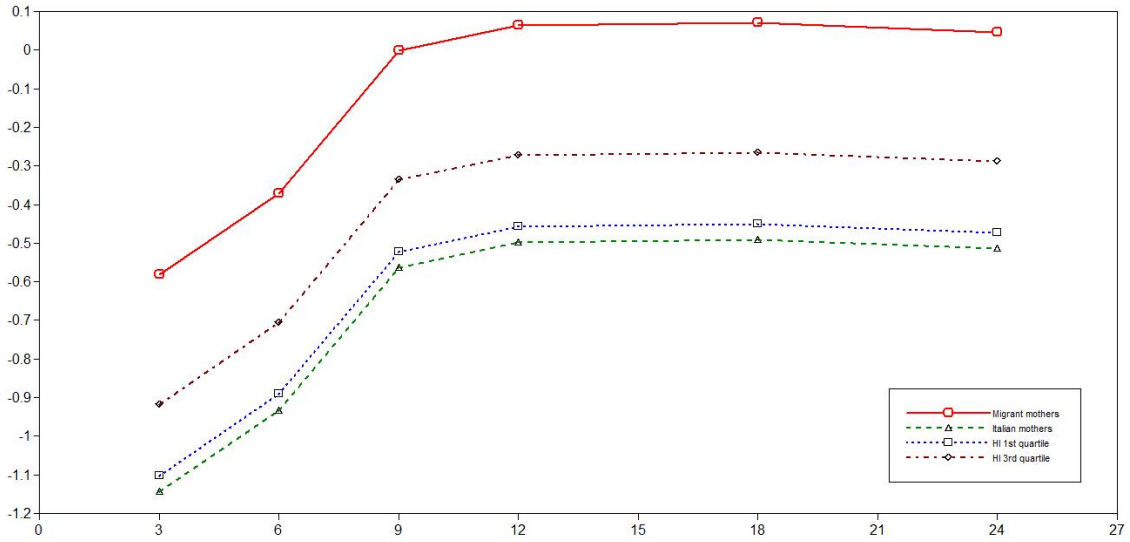
SGA had a similar behaviour (Figure 28): SGA children were much smaller at baseline (on average, their standardized score was 0.996 less than AGA/LGA), but they had a faster growth and an ever increasing curve, because they had no association with the quadratic term of the curve: the increasing rate determined by the slope was then protracted until the 24 months of age.

Figure 28 – Comparison of the adjusted estimated length trajectories of SGA preterms and AGA/LGA preterms resulting from the LGT-M4 model



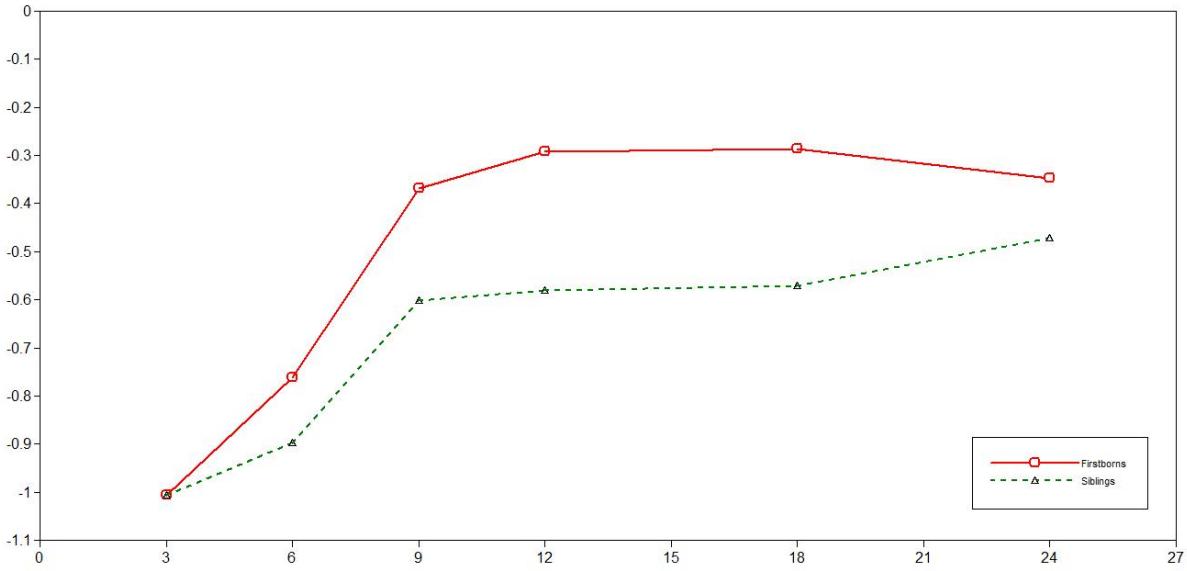
Preterms with a migrant mother and with a higher Hollingshead Index were associated with a higher baseline length; these two variables were not associated with the slope and the quadratic term, therefore their curves are exactly parallel with a shape corresponding to the one of the overall sample and the distance between them was determined by the offset at the starting point (Figure 29). The widest offset was between preterms with Italian or migrant mothers, in favour the latter which have higher estimated length values.

Figure 29 – Comparison of the adjusted estimated length trajectories of preterms with migrant mother, Italian mother, Hollingshead Index at the 1st and 3rd quartiles, resulting from the LGT-M4 model



Preterms with siblings started at the same length as firstborns because the variable was not associated with the estimated mean intercept; afterwards, they showed a slower growth (-0.155, p=0.012) and a later recovery (0.020, p=0.008). In fact, the positive effect of having siblings on the negative quadratic term's mean estimate determined an inversion of the relation with firstborns slope from negative to positive, generating two convergent curves (Figure 30).

Figure 30 – Comparison of the adjusted estimated length trajectories of preterms with sibligns and firstborns, resulting from the LGT-M4 model



Statistics of LGT-M4 Model are reported in Tab 9 and regression coefficients estimates are reported in Tab.10.

Tab 9 Parameter estimates, asymptotic standard errors and p-values of quadratic curve model for length with time-invariant covariates (LGT-M4 Model), n=282

Parameter	Estimate	Standard Error	p-value
Variances			
$\psi_{\alpha\alpha}$	0.853	0.143	<0.001
$\psi_{\beta1\beta1}$	0.034	0.018	0.064
$\psi_{\beta2\beta2}$	0.000	0.000	0.108
Covariance			
$\psi_{\alpha\beta1}$	0.070	0.041	0.089
$\psi_{\alpha\beta2}$	-0.014	0.004	0.002
$\psi_{\beta1\beta2}$	-0.003	0.002	0.131
Means			
μ_{α}	-0.727	0.084	<0.001
$\mu_{\beta1}$	0.213	0.029	<0.001
$\mu_{\beta2}$	-0.022	0.003	<0.001
Residual variances			
VAR(ε_0)	0.771	0.105	<0.001
VAR(ε_1)	0.363	0.047	<0.001
VAR(ε_2)	0.124	0.023	<0.001
VAR(ε_3)	0.140	0.022	<0.001
VAR(ε_4)	0.135	0.019	<0.001
VAR(ε_5)	0.050	0.064	0.437
Fit Statistics			
RMSEA	0.000		1.000
CFI	1.000		
TLI	1.009		
AIC	3151.572		
BIC	3253.476		
Number of free parameters (q)	28		
N/q ratio	10.1		

[a] Probability for RMSEA ≤0.05

Tab 10 Coefficient estimates and p-values for latent variables regressed on time-invariant covariates (LGT-M4 model), n=292

Predictor variable	Intercept	Slope	Quad. Term
Length of stay in NICU	-0.237 (p<0.001)	0.048 (p<0.001)	-0.004 (p<0.001)
SGA	-1.071 (p<0.001)	0.070 (p=0.001)	
Migrant mother	0.562 (p<0.001)		
Hollingshead Index	0.114 (p=0.033)		
Siblings		-0.155 (p=0.012)	0.020 (p=0.008)

Multivariate Latent Curve Models based on repeated measures of neurodevelopment and length

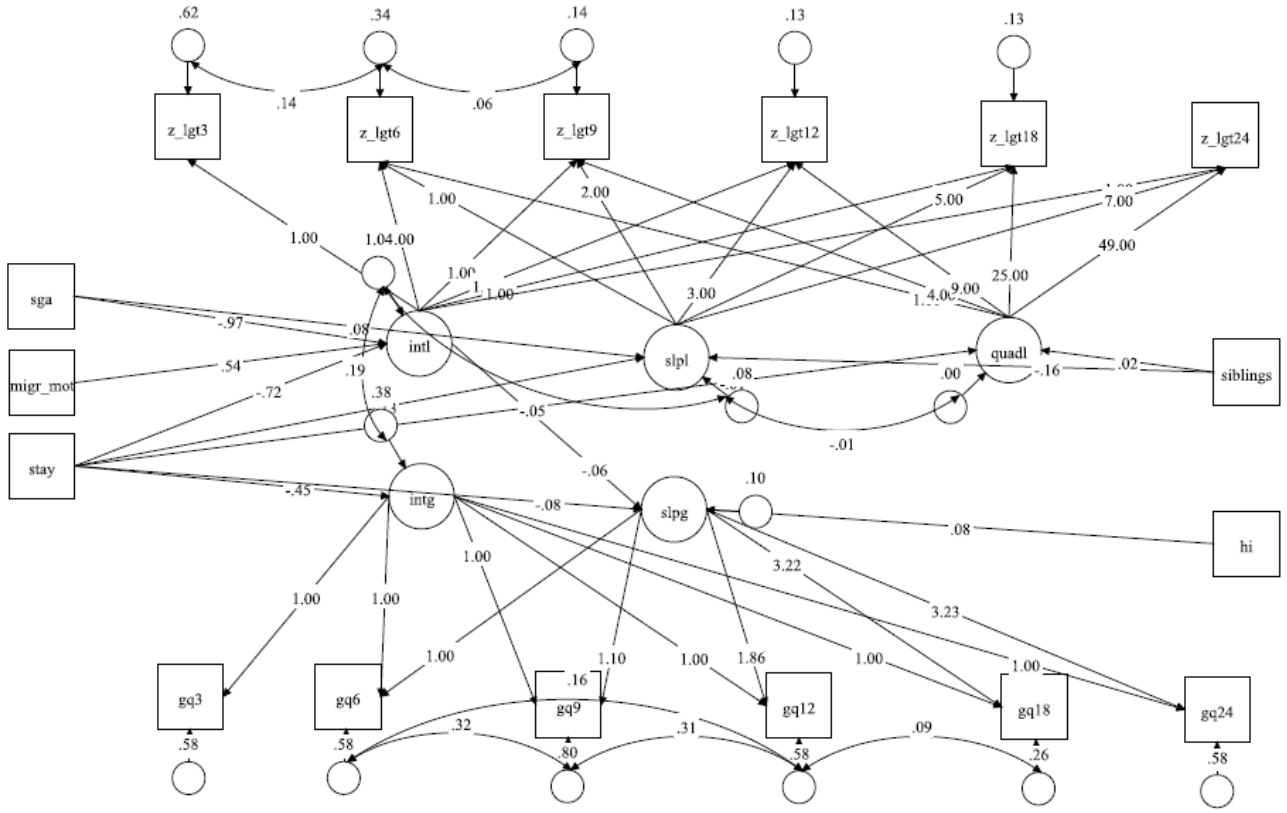
The two best-fitting models obtained for repeated measures of preterms' neurodevelopment and of length were combined into a single multivariate model in order to test the association among the latent variables underlying the two curves. Each of the two latent slopes were regressed on the other curve's intercept, to test the assumption that the model implied mean baseline level of one repeated measure may predict the mean level of change of the other repeated measure, and vice versa. The other latent variables were allowed to covary.

GL-M1 – Multivariate Latent Curve Model with Completely Latent Curve for Neurodevelopment and with quadratic curve for Length

GL-M1 Model was obtained by combining GQ_M4 model and LGT-M4 model. The sample size was 305 because there were 19 infants with at least one missing on the covariates; the number of free parameters was 54, thus leading to a suboptimal N/q ratio of 5.6.

The model fit was excellent: RMSEA=0.015 (90% CI: 0.000-0.035, p<=0.05=1.000), CFI=0.998, TLI=0.997. The covariate indicating a sustained human milk feeding until 3 months of age was removed because it was no more significantly associated with the latent slope of neurodevelopment. The effect from the Hollingshead Index to the intercept of neurodevelopment was removed because it became not significant as well.

Figure 31 –Diagram of GL-M1 model

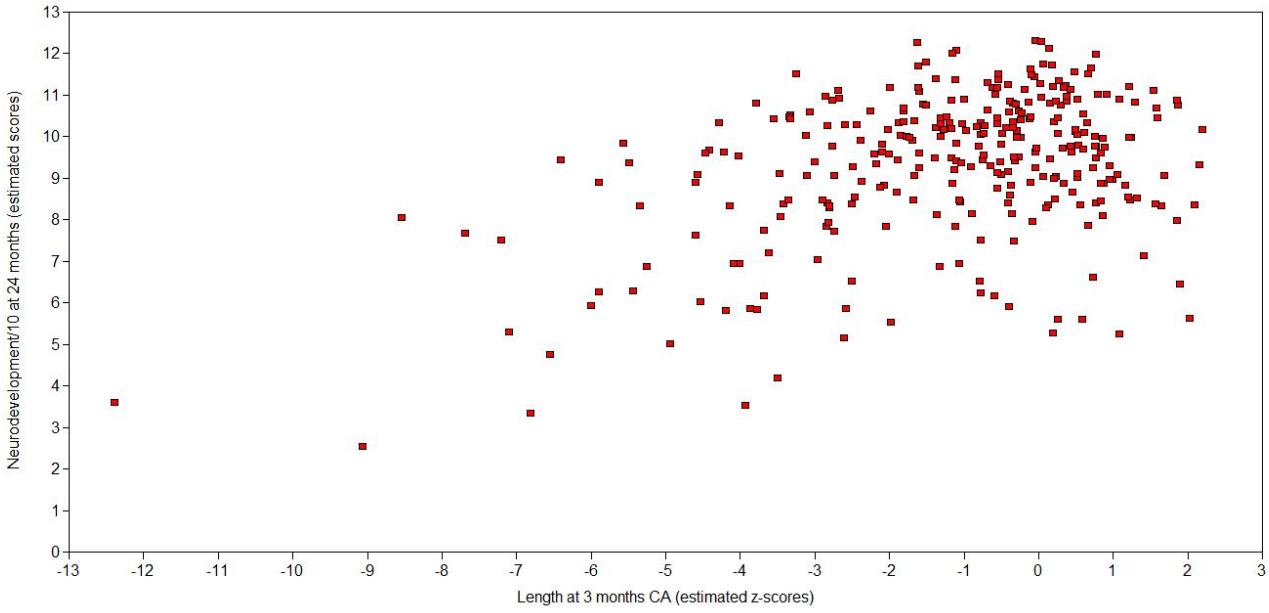


There was a significant negative relation among the slope of neurodevelopment and the intercept of length: $\xi(\beta_N\alpha_L)$=-0.061 (p=0.012). Since the slope of neurodevelopment was negative, this negative relation indicates that preterms with higher initial lengths exhibit a larger decline in neurodevelopment score. The relation between the intercept of neurodevelopment and the slope of length growth proved to be not significant ($\xi(\beta_L\alpha_N)$=-0.013; p=0.829). Significant covariances were found among the two intercepts ($\psi(\alpha_N\alpha_L)$=0.164, p=0.017) showing that higher baseline levels of length are related to higher baseline levels of neurodevelopment; and among the

intercept and the quadratic term of length ($\psi(\alpha_L\gamma_L)$=-0.013, p=0.003), showing that preterms with a higher length baseline level exhibit a lower later recovery.

All these effects combined, the scatterplot among the estimated values of length at 3 months corrected age and the estimated neurodevelopment scores at 24 months (Fig.32) indicates a mild positive relation (higher initial length values corresponding to higher neurodevelopment).

Figure 32 –Scatterplot among estimated z-scores of Length at 3 months CA and estimated Neurodevelopment scores at 24 months CA (GL-M1 model)



In addition to the loss of significance of Sustained human milk, other changes were found in the effects of covariates on the latent growth factors. Length of stay had significant effects on every latent factor, even on the slope of neurodevelopment (where in GQ-M4 Model it had p=0.158). As to neurodevelopment, preterms with a longer NICU stay had lower baseline values ($\xi_1(\alpha_N)$=-0.445, p<0.001) and a less declining slope ($\xi_1(\beta_N)$=-0.087, p=0.005); as to length, the effect of a longer NICU stay was to decrease the baseline level ($\xi_1(\alpha_L)$=-0.710, p<0.001), to raise the growth slope ($\xi_1(\beta_L)$=0.138, p<0.001) and to weaken growth acceleration ($\xi_1(\gamma_L)$=-0.011, p<0.001). However, the total effect on neurodevelopment slope resulted negative non significant ($T\xi_1(\beta_N)$=-

0.043, p=0.084), because it was obtained by summing the counterbalanced negative direct effect and the positive indirect effect.

Hollingshead Index affected significantly the slope of neurodevelopment ($\xi_2(\beta_N)$=0.078, p<0.001) and the baseline length level ($\xi_2(\alpha_L)$=0.120, p=0.025), indicating that preterms from more educated and working are expected to have higher initial levels of length and a smaller decline in neurodevelopment.

Migrant condition of mothers was significantly related directly only with the baseline length level ($\xi_3(\alpha_L)$=0.576, p<0.001), because the direct association with the slope of neurodevelopment fell to borderline significance ($\xi_2(\beta_N)$=-0.106, p=0.075); however, the total effect on the slope of neurodevelopment (corresponding to the sum of the direct effect and of the indirect effect connecting the migrant mothers' variable to the slope of neurodevelopment through the initial level of length) was significant ($T\xi_2(\beta_N)$=-0.141, p=0.016). Preterms from a migrant mother are then expected to have a faster decline in neurodevelopment mainly because of their higher initial levels of length and the subsequent negative effect of initial length in neurodevelopment.

SGA condition was significantly associated with two latent factors of length, specifically with lower initials levels ($\xi_4(\alpha_L)$=-1.042, p<0.001) and a faster growth ($\xi_4(\beta_L)$=0.068, p=0.001). As a consequence, the total effect from SGA to the slope of neurodevelopment was significant ($T\xi_4(\beta_N)$=0.064, p=0.009): as the relation was negative, SGA preterms are expected to have a slower decline in neurodevelopment.

The presence of siblings was significantly associated with length growth, showing a slower velocity ($\xi_5(\beta_L)$=-0.159, p<0.010) and a higher recovery ($\xi_1(\gamma_L)$=0.020, p=0.007).

Observed variables of length had higher proportions of explained variance (ranging from 0.705 to 0.958) when compared to the observation of neurodevelopment, that ranged from 0.499 to 0.883; this is consistent with the previous result of a better fit of the length model with respect to the neurodevelopment model. Similarly, the latent factors of length had higher values of explained variance though in both cases the latent intercept showed the highest values.

Tab 11 Parameter estimates, asymptotic standard errors and p-values of multivariate latent curve model for neurodevelopment and length with time-invariant covariates (GL-M1 model), n=305

Parameter	Estimate	Standard Error	p-value
Variances			
$\psi(\alpha_N)$	0.406	0.088	<0.001
$\psi(\alpha_L)$	0.860	0.144	<0.001
$\psi(\beta_N)$	0.098	0.024	<0.001
$\psi(\beta_L)$	0.037	0.019	0.047
$\psi(\gamma_L)$	0.001	0.000	0.088
Covariances			
$\psi(\alpha_N\alpha_L)$	0.164	0.068	0.017
$\psi(\alpha_N\gamma_L)$	-0.001	0.003	0.777
$\psi(\beta_N\beta_L)$	0.018	0.011	0.081
$\psi(\beta_N\gamma_L)$	-0.001	0.001	0.351
$\psi(\alpha_N\beta_N)$	-0.013	0.027	0.617
$\psi(\alpha_L\beta_L)$	-0.068	0.039	0.082
$\psi(\alpha_L\gamma_L)$	-0.013	0.005	0.003
$\psi(\beta_L\gamma_L)$	-0.004	0.002	0.101
Means			
$\mu(\alpha_N)$	11.408	0.049	<0.001
$\mu(\alpha_L)$	-0.734	0.084	<0.001
$\mu(\beta_N)$	-0.582	0.066	<0.001
$\mu(\beta_L)$	-0.368	0.707	0.603
$\mu(\gamma_L)$	-0.022	0.003	<0.001
Residual variances			
$VAR(\varepsilon_{0N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{1N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{2N})$	0.797	0.080	<0.001
$VAR(\varepsilon_{3N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{4N})$	0.251	0.058	<0.001
$VAR(\varepsilon_{5N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{0L})$	0.756	0.105	<0.001
$VAR(\varepsilon_{1L})$	0.355	0.045	<0.001
$VAR(\varepsilon_{2L})$	0.128	0.023	<0.001
$VAR(\varepsilon_{3L})$	0.139	0.022	<0.001
$VAR(\varepsilon_{4L})$	0.133	0.019	<0.001
$VAR(\varepsilon_{5L})$	0.054	0.062	0.387
Fit Statistics			
RMSEA	0.015		1.000
CFI	0.998		
TLI	0.997		
AIC	7655.861		
BIC	7856.758		
Number of free parameters (q)	54		
N/q ratio	5.6		
Estimated power	0.93		

[a] Probability for RMSEA ≤0.05

Tab 12 Coefficient estimates and p-values for latent variables regressed on time-invariant covariates (GL-M1 model), n=305

	Neurodevelopment			Length		
Predictor variable	**Intercept**	**Slope (direct effect)**	**Slope (total effect)**	**Intercept**	**Slope**	**Quad. Term**
Baseline mean of Length	-	-0.061 (p=0.012)		-	-	-
Baseline mean of Neurodevelopment	-	-		-	-0.013 (p=0.829)	-
Length of stay in NICU	-0.445 (p<0.001)	-0.087 (p=0.005)	-0.043 (p=0.084)	-0.710 (p<0.001)	0.138 (p<0.001)	-0.011 (p<0.001)
SGA	-	-	0.064 (p=0.009)	-1.042 (p<0.001)	0.068 (p<0.001)	-
Migrant mother	-	-0.106 (p=0.075)	-0.141 (p=0.016)	0.576 (p<0.001)	-	-
Hollingshead Index	-	0.078 (p=0.001)	0.070 (p=0.001)	0.120 (p=0.025)	-	-
Siblings	-	-		-	-0.159 (p=0.010)	0.020 (p=0.007)

GL-M2 – Multivariate Latent Curve Model with Completely Latent Curves

Joining the two best-fitting univariate models to obtain GL-M1 model produced a multivariate model that was not homogeneous with respect to the curves' parameterization, since the curve of length was obtained with a quadratic model and the curve of neurodevelopment was obtained with a completely latent curve. Even though GL-M1 model fit very well to data, it is useful to consider an additional multivariate model where the curves parameterization is the same for both measures. The completely latent curve parameterization was used to evaluate GL-M2 model because it allowed to compare the curves' shapes through the loadings estimated values.

GL-M2 model fit was was good: RMSEA=0.038 (90% CI: 0.023-0.051; p=0.938), CFI was 0.987, TLI was 0.984. However, these indexes and AIC and BIC information-based indexes were higher than those found in GL-M1, showing a worse fit. Two of the significant relations found in GL-M1 became non significant: the effect from HI to the slope of length reduced to 0.120 and its estimated p-value was 0.062; the effect of having siblings on the slope of length fell to -0.082, with an estimated p-value of 0.056. The effect of length baseline level on the slope of neurodevelopment remained negative and significant, while the influence from initial neurodevelopment to the length slope remained not significant, but grew in magnitude. The correlation among length's intercept and slope became significant (r=-0.482, p<0.001). The total effects from the covariates to the slope of neurodevelopment were confirmed in magnitude and significance.

The series of estimated loadings from the slopes to the observed values depict two quite different shapes of the growth curves. The ending point is substantially the same: λ_{24N}=3.282 and λ_{24L}=3.280, but while length shape had an early steep increase (λ_{9L}=2.284) that subsequentely decelerated until the end of the study, neurodevelopment grew very slowly until 9 months (λ_{9N}=1.073) and then had a very steep increase from 12 to 18 months, where its growth rate was higher than length's growth rate (λ_{18L}=3.162; λ_{18N}=3.269).

This different parameterization of the model confirmed the relation among the initial level of length and the slope of neurodevelopment and the strong influence of NICU stay on both trajectories. Two of the other predictors, namely HI and siblings, showed a reduced and non significant effect on the baseline level and the slope of length respectively. The two trajectories revealed different patterns of growth that ended in a similar gain upon the baseline levels: length has most of its growth in the 3-9 months interval while neurodevelopment changed mainly in the 9-18 month interval.

Tab 13 Parameter estimates, asymptotic standard errors and p-values of multivariate latent curve model for neurodevelopment and length with time-invariant covariates (GL-M2 model), n=305

Parameter	Estimate	Standard Error	p-value
Variances			
$\psi(\alpha_N)$	0.405	0.088	<0.001
$\psi(\alpha_L)$	1.384	0.152	<0.001
$\psi(\beta_N)$	0.098	0.024	<0.001
$\psi(\beta_L)$	0.088	0.024	<0.001
Covariances			
$\psi(\alpha_N\alpha_L)$	0.212	0.073	0.004
$\psi(\beta_N\beta_L)$	0.015	0.007	0.034
$\psi(\alpha_N\beta_N)$	-0.013	0.027	0.631
$\psi(\alpha_L\beta_L)$	-0.168	0.041	<0.001
Means			
$\mu(\alpha_N)$	11.409	0.049	<0.001
$\mu(\alpha_L)$	-0.758	0.087	<0.001
$\mu(\beta_N)$	-0.572	0.064	<0.001
$\mu(\beta_L)$	1.012	0.557	0.069
Residual variances			
$VAR(\varepsilon_{0N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{1N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{2N})$	0.799	0.080	<0.001
$VAR(\varepsilon_{3N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{4N})$	0.252	0.059	<0.001
$VAR(\varepsilon_{5N})$	0.577	0.041	<0.001
$VAR(\varepsilon_{0L})$	0.220	0.021	<0.001
$VAR(\varepsilon_{1L})$	0.220	0.021	<0.001
$VAR(\varepsilon_{2L})$	0.220	0.021	<0.001
$VAR(\varepsilon_{3L})$	0.220	0.021	<0.001
$VAR(\varepsilon_{4L})$	0.086	0.019	0.563
$VAR(\varepsilon_{5L})$	0.220	0.021	<0.001
Fit Statistics			
RMSEA	0.038		0.938
CFI	0.987		
TLI	0.984		
AIC	7690.290		
BIC	7868.865		
Number of free parameters (q)	48		
N/q ratio	6.4		
Estimated power	0.90		

[a] Probability for RMSEA ≤ 0.05

Tab 14 Estimated loadings Coefficient estimates (GL-M2 model), n=305

Loadings	Neurodevelopment	Length
λ_0 (3 months)	0.000	0.000
λ_1 (6 months)	1.000	1.000
λ_2 (9 months)	1.073	2.284
λ_3 (12 months)	1.906	2.594
λ_4 (18 months)	3.269	3.162
λ_5 (24 months)	3.282	3.280

Tab 15 Coefficient estimates and p-values for latent variables regressed on time-invariant covariates (GL-M2 model), n=305

Predictors	Neurodevelopment			Length	
	Intercept	Slope (direct effect)	Slope (total effect)	Intercept	Slope
Baseline mean of Length		-0.050 (p=0.005)			-
Baseline mean of Neurodevelopment		-			-0.070 (p=0.153)
Length of stay in NICU	-0.445 (p<0.001)	-0.079 (p=0.005)	-0.043 (p=0.086)	-0.729 (p<0.001)	0.107 (p=0.001)
SGA	-	-	0.054 (p=0.005)	-1.078 (p<0.001)	0.133 (p=0.015)
Migrant mother	-	-0.108 (p=0.066)	-0.137 (p=0.019)	0.586 (p<0.001)	-
Hollingshead Index	-	0.078 (p<0.001)	0.073 (p=0.001)	0.102 (p=0.062)	-
Siblings		-			-0.082 (p=0.056)

DISCUSSION

In this study the trajectories of preterms' neurodevelopment and length from 3 to 24 months of corrected age have been examined and several latent curve models have been fitted to the data in order to examine the underlying pattern of change of these two outcomes, the interrelation among the two trajectories and the effect exerted on them by some covariates of clinical and socio-economic relevance.

Neurodevelopment at 24 months is considered as fairly indicative of infants' development at later ages, therefore it is important to acquire the best available knowledge about the path leading to this step of development. The association among growth and neurodevelopment has already been extensively analyzed, but most of the published studies focus on growth in the NICU, while sensitive periods for neurodevelopment span from time before and after term to the end of the first year of life.[3] Furthermore, when longitudinal studies were carried on, very often only the endpoint information was used, thus missing information on what happened in the intermediate periods. As Belfort et al[3] suggest, from a clinical viewpoint time matters because in the preterms' first two years of life different types of nutrition are administered from different sets of health care providers. This study attempted to overcome those limitations by describing the patterns of neurodevelopment and of growth in the post-discharge period taking into account all available data gathered at nearby follow-up visits, as well as clinical and socio-economic data baseline information.

A conditional completely latent curve model (GQ-M4 model, tabb. 5 and 6) was the more adequate to represent the trajectory of preterms neurodevelopment. This type of model is a non-linear growth model with time-invariant covariates associated with the latent initial mean value and the latent mean slope. The observed trajectory of standardized Griffiths scores was stepwise declining, with two periods of steeper decrease (from 3 to 6 months and from 9 to 18 months) interposed on two periods of stability from 6 to 9 months and from 18 to 24 months. The neurodevelopment stepwise pattern was faithfully reproduced by GQ-M4 model and a mean slope of -5.47 was estimated, a figure corresponding to the average loss in Griffiths' GQ score experienced from

3 to 6 months of age.[3] By weighing this average rate of change for the loadings that identified the trajectory, the mean decrease estimate in GQ scores at the end of the 3-24 months interval was -17.96. However, it should be noted that the initial level of the curve could be overestimated, since neurodevelopment assessment at 3 months may depend on benevolent parents' reports. Moreover, GQ scores become progressively more accurate at later ages because as infants grow a greater number and more demanding tasks are administered and for this reason a delay in development that was unobservable in earlier measures may be diagnosed later. As a consequence, the average negative slope that was observed may look steeper than the true slope and it should not be considered as an indication of an increasing preterns' neurodevelopment delay, because it could be determined by the progressive improvement of scores' accuracy.

Preterms showed significant individual variability both in initial values and in decline rate and four variables were found to be associated with the trajectory. Among these, the duration of NICU hospitalization and sustained human milk feeding were clinical factors, while socio-economic status and the presence of a migrant mother were socio-economic factors. It should be remembered that biometric variables were not considered at this stage of analysis, since they were entered as a parallel growth process in the third type of LCM model.

Covariates had an influence on neurodevelopment that was coherent with their nature: NICU stay, which is a proxy of infants' perinatal morbidity and as such brings informations about preterms' initial conditions, affected only the baseline 3 months' neurodevelopment level. On the contrary the two socio-economic factors and sustained milk feeding are highly dependent from infants' familiar environment and therefore had an influence only on the average slope.

The positive influence on neurodevelopment slope exerted by sustained milk feeding was found in several other observational studies[61–63] and randomized trials[64]; it was explained by the presence of nutrients in breast milk that may benefit the developing brain or alternatively by the physical and social interactions between the mother

[3] The actual estimate of slope (-0.547) was multiplied by 10 to return to the original GQ score metric.

and the infant inherent in breastfeeding that may benefit both mothers, in terms of psycho-physical behaviour, and infants, receiving better care that may anticipate their neurological development processes.

The positive association among socio-economic status (as measured by Hollingshead Index) and the slope of neurodevelopment indicates that preterms from disadvantaged families had a higher risk of achieving a lower neurodevelopment. A previous cohort study[65] that used socioeconomic status as an exposure variable found similar results and suggested that deprived families may experience feelings of shame, inferiority and powerlessness that, combined with financial hardships, may increase the risk of illness and depression, thus diverting mothers from child care and breastfeeding.

The negative influence on neurodevelopment found for migrant mothers is not attributable to worse socioeconomic conditions of migrant families, since its effect was adjusted for the Hollingshead Index. It is then likely that this net effect might relate to linguistical impairments, that can provide their negative influence by means of greater difficulties in understanding clinicians' prescriptions and some of the Griffiths Scale tests. In fact, standardized tools such as the Griffiths and Bayley scales include verbal skills based on the language of the country where the infants are evaluated, which for infants born from migrant women does not correspond to their mother tongue. A cultural bias of the personal-social subscale of the Griffiths scale was already highlighted by Luiz et al.[66] when comparing the scores obtained by Black and White children in South Africa.

The trajectory of preterms' length found in this study shows a steep increase from 3 to 12 months, followed by a flat evolution in the second year of life. This finding is entirely adherent to the results of previous studies.[8]

Among the factors influencing length's trajectory, neonatal morbidities and SGA caused preterms to be smaller at 3 months, but those preterms had a higher increase during follow-up, although they remained smaller al 24 months. The presence of neonatal morbidities was already recognized as one of the major growth restriction factors, since a difficult neonatal course may affect physiologic and regulatory systems, leading to poorer feeding and subsequently to a restricted growth.[8,67] A higher increase of SGA preterms compared to AGA preterms was already observed,[68] although this catch-up did not end in the achievement of a normal standardized

height at school age. An important factor for length restriction is genetic potential;[22,68] unfortunately, for this study there are no biometric measures of preterms' parents available, although the relationship among the initial level of length and migrant mothers may partially reflect the existence of different genetic features. The positive association among length and a higher socio-economic status is difficult to interpret, since it seems questionable to associate a higher stature to a higher status. In the same way, the long-term positive effect of having siblings must be investigated more deeply, since it may be due to a combination of effects relating to the family's emotional bonds.

The last finding of the study is the relationship between the trajectories of length growth and of neurodevelopment. It was evaluated by multivariate latent curve models GL-M1 (tabb. 11 and 12) and GL-M2 (tabb. 13-15). The association between the estimated mean baseline levels of length and the estimated mean growth rate of neurodevelopment was negative, suggesting that for preterms who begin follow-up at a higher length a larger decline in standardized scores of neurodevelopment should be expected. On the other hand, the initial levels of the two outcomes were positively correlated, indicating that preterms taller at 3 months had also higher neurodevelopment at 3 months. Therefore this combined result can be explained as a regression towards the mean: specifically, the neurodevelopment level at 24 months may be reached either from a more declining higher initial growth level or from a progressive recovery from lower initial growth levels, while a higher initial preterm's length was not predictive of a better final neurodevelopment.

Among the covariates influencing the two trajectories, NICU stay had a significant direct association with each of the four latent growth parameters: examining these relations separately, preterms with a longer stay, as a consequence of a more severe morbidity, are expected to start follow-up at lower neurodevelopment and height levels, to grow more in height and to have a more decreasing neurodevelopment curve. However, for the mediating role exerted by initial levels of growth on the slope of neurodevelopment, these effects were actually opposed and resulted in a non significant total effect. When looking at the preterm growth process in a comprehensive way, that is taking into account the interrelations among height and neurodevelopment growth,

preterm perinatal morbidity then seemed not so influent as it was by examining the two processes separately. On the contrary, the other covariates proved to be significantly associated with neurodevelopment growth: SGA for its strong relation with initial height level and socio-economic status for its direct relation with the neurodevelopment slope. Preterms born from a migrant mother were found to be associated with a more declining neurodevelopment slope in the univariate model, but in the multivariate model this direct effect is no more significant. However, through the relation with length at 3 months, once again seen as a mediator variable, a significant total effect on the neurodevelopment slope was found, suggesting that preterms with migrant mothers experienced a larger neurodevelopment decline mainly because they were taller at the beginning of follow-up and because taller preterms had a more declining neurodevelopment curve. In fact, preterms from migrant mothers were significantly taller at 3 months both on raw values (59.1 vs 58.0 cm; t-test=-2.22, p=0.027) and on standardized values (-0.61 vs -1.13; t-test=-2.29, p=0.023).

Some of the results found in this study may be controversial and some discrepancies between results obtained in univariate and multivariate models may seem confusing. However, one of the strength of the study is actually the methodology that allows to evaluate jointly the neurodevelopment and growth trajectories. Using Latent Curve multivariate models, magnitude and significance of the relationships among the two trajectories are not estimated from the observed values but from their latent initial values and slopes. These values are actually the results of a synthesis among interindividual and intraindividual variations over time and as such reflect the overall covariance pattern among the two phenomena.

A probably controversial result that was found concerns the non significant influence of preterms' perinatal morbidity, expressed by length of NICU stay, on the neurodevelopment rate of change. This may especially be surprising given that at the univariate analyses this predictor resulted the more influential among the four submitted in the LCM models. Such finding was obtained only thanks to a multivariate analysis that evaluated simultaneously the growth and neurodevelopment curves, but certainly deserves to be more extensively analyzed.

An interesting insight was provided regarding the effect exerted on neurodevelopment by the presence of a preterms' migrant mother. While this study confirms (but with borderline significance) the existence of a direct relation among this exposure and this outcome, the interesting new finding (once again revealed only by the multivariate model) is that there may actually be a mediating effect of the initial level of growth. Therefore, being born from migrant mothers may be related not only to socio-cultural (and specifically linguistic) issues, but also to genetical factors. A possible explanation of preterms with migrant mothers' higher neurodevelopment decline may be then that they, being taller at the beginning of follow-up, somehow anticipated the development process and their regression to the mean neurodevelopment level was faster than for preterms with Italian mothers. Unfortunately, lack of information about preterms' parents biometrical characteristics is precluding further considerations about this hypothesis. A longer follow-up, continued until school age and adolescence, would be highly useful to corroborate this finding.

Another limitation of the study is the possible bias associated to the first measures of neurodevelopment, where preterms' scores might be overestimated by parents' answers. While it is very likely that this bias is randomly distributed among preterms, nonetheless it may affect neurodevelopment trajectory estimates and their relationships with the length trajectory. Prolongation of follow-up in order to have additional later measures of growth and neurodevelopment would allow to evaluate models shifted forward in time, then testing whether the 3 months visit may actually be biased.

While it may be questionable to utilize length trajectory as representative of preterms' growth, its influence (both direct and as a mediator) was so relevant that the importance of multivariable modeling of the neurodevelopment trajectory can not be neglected. Therefore, further studies exploring the relationships among trajectories of neurodevelopment and other biometric measures of preterms' growth, such as weight and cranial circumference, may give further insights on the neurodevelopment process.

REFERENCES

1. Lucas A, Morley R, Cole TJ, et al. Early diet in preterm babies and developmental status at 18 months. *Lancet*. 1990;335(8704):1477–1481.

2. Ehrenkranz R a, Dusick AM, Vohr BR, Wright LL, Wrage L a, Poole WK. Growth in the neonatal intensive care unit influences neurodevelopmental and growth outcomes of extremely low birth weight infants. *Pediatrics*. 2006;117(4):1253–61. doi:10.1542/peds.2005-1368.

3. Belfort MB, Rifas-Shiman SL, Sullivan T, et al. Infant growth before and after term: effects on neurodevelopment in preterm infants. *Pediatrics*. 2011;128(4):e899–906. doi:10.1542/peds.2011-0282.

4. Prader A. Catch-up growth. *Postgrad Med J*. 1977;18(February):646–661. Available at: http://europepmc.org/abstract/MED/724586. Accessed February 6, 2014.

5. Tanner JM. Regulation of growth in size in mammals. *Nature*. 1963;199:845–.

6. De Wit CC, Sas TCJ, Wit JM, Cutfield WS. Patterns of catch-up growth. *J Pediatr*. 2013;162(2):415–20. doi:10.1016/j.jpeds.2012.10.014.

7. Boersma B, Wit JM. Catch-up growth. *Endocr Rev*. 1997;18(5):646–661.

8. Sullivan MC, McGrath MM, Hawes K, Lester BM. Growth trajectories of preterm infants: birth to 12 years. *J Pediatr Health Care*. 2008;22(2):83–93. doi:10.1016/j.pedhc.2007.02.008.

9. Griffiths. *Griffiths Mental Development Scales – GMDS.*; 1970.

10. Giannì ML, Picciolini O, Vegni C, Gardon L, Fumagalli M, Mosca F. Twelve-month neurofunctional assessment and cognitive performance at 36 months of age in extremely low birth weight infants. *Pediatrics*. 2007;120(5):1012–9. doi:10.1542/peds.2006-3364.

11. Sansavini A, Savini S, Guarini A, Broccoli S, Alessandroni R, Faldella G. The effect of gestational age on developmental outcomes: a longitudinal study in the first 2 years of life. *Child Care Health Dev*. 2011;37(1):26–36. doi:10.1111/j.1365-2214.2010.01143.x.

12. Johnson S, Marlow N. Developmental screen or developmental testing? *Early Hum Dev*. 2006;82:173–183. doi:10.1016/j.earlhumdev.2006.01.008.

13. Savini S, Sansavini A, Guarini A, Alessandroni R, Faldella G. Il follow-up del nato pretermine: esperienze, strumenti e nuove problematiche. In: Sansavini A, Faldella G, eds. *Lo sviluppo dei bambini nati pretermine*. Milano, Italy: FrancoAngeli; 2013:384.

14. Lefebvre F, Glorieux J, St-Laurent-Gagnon T. Neonatal survival and disability rate at age 18 months for infants born between 23 and 28 weeks of gestation. *Am J …*. 1996:833–838. Available at: http://www.sciencedirect.com/science/article/pii/S0002937896703095. Accessed February 6, 2014.

15. Claas M, Bruinse H. Two-year neurodevelopmental outcome of preterm born children≤ 750 g at birth. *Arch Dis …*. 2011. doi:10.1136/adc.2009.174433.

16. Hindmarsh G, O'Callaghan M. Gender differences in cognitive abilities at 2 years in ELBW infants. *Early Hum …*. 2000;60:115–122. Available at: http://www.sciencedirect.com/science/article/pii/S0378378200001055. Accessed February 6, 2014.

17. Barnett A, Guzzetta A. Can the Griffiths scales predict neuromotor and perceptual-motor impairment in term infants with neonatal encephalopathy? *Arch Dis …*. 2004:637–644. doi:10.1136/adc.2002.019349.

18. Claas MJ, de Vries LS, Bruinse HW, et al. Neurodevelopmental outcome over time of preterm born children ≤750 g at birth. *Early Hum Dev*. 2011;87:183–191. doi:10.1136/adc.2009.174433.

19. Gutbrod T, Wolke D, Soehne B, Ohrt B, Riegel K. Effects of gestation and birth weight on the growth and development of very low birthweight small for gestational age infants : a matched group comparison. *Arch Dis Childhood-Fetal Neonatal …*. 2000;82:208–215.

20. Rijken M, Wit JM, Le S, Veen S. The effect of perinatal risk factors on growth in very preterm infants at 2 years of age : The Leiden Follow-Up Project on Prematurity. 2007:527–534. doi:10.1016/j.earlhumdev.2006.10.002.

21. Leppänen M, Lapinleimu H, Lind A, et al. Antenatal and Postnatal Growth and 5-Year Cognitive Outcome in Very Preterm Infants. *Pediatrics*. 2013. doi:10.1542/peds.2013-1187.

22. Brandt I, Sticker EJ, Gausche R, Lentze MJ. Catch-up growth of supine length/height of very low birth weight, small for gestational age preterm infants to adulthood. *J Pediatr*. 2005;147(5):662–8. doi:10.1016/j.jpeds.2005.06.034.

23. Mercier CE, Dunn MS, Ferrelli KR, Howard DB, Soll RF. Neurodevelopmental outcome of extremely low birth weight infants from the Vermont Oxford network: 1998-2003. *Neonatology*. 2010;97(4):329–38. doi:10.1159/000260136.

24. Casey PH. Growth of low birth weight preterm children. *Semin Perinatol*. 2008;32(1):20–7. doi:10.1053/j.semperi.2007.12.004.

25. Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of Very Low Birth Weight Infants to Age 20 Years. *Pediatrics*. 2003;112:e30–e38. doi:10.1542/peds.112.1.e30.

26. Guo SS, Roche AF, Chumlea WC, Casey PH, Moore WM. Growth in weight, recumbent length, and head circumference for preterm low-birthweight infants during the first three years of life using gestation-adjusted ages. *Early Hum Dev*. 1997;47:305–325. doi:10.1016/S0378-3782(96)01793-8.

27. Morris BH, Smith KE, Swank PR, Denson SE, Landry SH. Patterns of physical and neurologic development in preterm children. *J Perinatol*. 2002;22:31–36. doi:10.1038/sj.jp.7210590.

28. Bertino E, Di Nicola P, Varalda A, Occhi L, Giuliani F, Coscia A. Neonatal growth charts. *J Matern Neonatal Med*. 2012;25:67–69. doi:10.3109/14767058.2012.664889.

29. Westby Wold SH, Sommerfelt K, Reigstad H, et al. Neonatal mortality and morbidity in extremely preterm small for gestational age infants: a population based study. *Arch Dis Child Fetal Neonatal Ed*. 2009;94(5):F363–7. doi:10.1136/adc.2009.157800.

30. Latal-Hajnal B, Siebenthal K von. Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome. *J …*. 2003;3476(03):163–170. Available at: http://www.sciencedirect.com/science/article/pii/S0022347603002439. Accessed March 3, 2014.

31. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. *J Pediatr*. 1978;92:529–534. doi:10.1016/S0022-3476(78)80282-0.

32. Prematurity IC for the C of R of. The international classification of retinopathy of prematurity revisited. *Arch Ophthalmol*. 123(7):991–999. Available at: http://dx.doi.org/10.1001/archopht.123.7.991.

33. Potijk M, Kerstjens J, Bos A, Reijneveld S, de Winter A. Developmental Delay in Moderately Preterm-Born Children with Low Socioeconomic Status: Risks Multiply. *J Pediatr*. 2013;163(5):1289–1295. doi:10.1016/j.jpeds.2013.07.001.

34. Bodeau-Livinec F, Zeitlin J, Blondel B, et al. Do very preterm twins and singletons differ in their neurodevelopment at 5 years of age? *Arch Dis Child Fetal Neonatal Ed*. 2013;98(6):F480–7. doi:10.1136/archdischild-2013-303737.

35. Jedrychowski W, Perera F, Jankowski J, et al. Effect of exclusive breastfeeding on the development of children's cognitive function in the Krakow prospective birth cohort study. *Eur J Pediatr*. 2012;171(1):151–8. doi:10.1007/s00431-011-1507-5.

36. Koutra K, Chatzi L, Roumeliotaki T, et al. Socio-demographic determinants of infant neurodevelopment at 18 months of age: Mother-Child Cohort (Rhea Study) in Crete, Greece. *Infant Behav Dev*. 2012;35(1):48–59. doi:10.1016/j.infbeh.2011.09.005.

37. Auger N, Park AL, Gamache P, Pampalon R, Daniel M. Weighing the contributions of material and social area deprivation to preterm birth. *Soc Sci Med*. 2012;75:1032–1037. doi:10.1016/j.socscimed.2012.04.033.

38. Hollingshead A. Four factor index of social status. *Yale J Sociol*. 1975;8:21–52. Available at: http://elsinore.cis.yale.edu/sociology/yjs/yjs_fall_2011.pdf#page=21.

39. Preacher K, Wichman A, MacCallum R, Briggs N. *Latent growth curve modeling*. Thousand Oaks, CA: Sage; 2008:96.

40. Hancock GR, Choi J. A Vernacular for Linear Latent Growth Models. *Struct Equ Model A Multidiscip J*. 2006;13:352–377. doi:10.1207/s15328007sem1303_2.

41. Curran PJ, Hussong AM. Structural Equation Modeling of Repeated Measures Data : Latent Curve Analysis. In: *Modeling Intraindividual Variability With Repeated Measures Data Methods and Applications*.; 2002:59–86.

42. McArdle JJ. A structural modeling experiment with multiple growth functions. In: *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences.*; 1989:71–117.

43. Hancock GR, Lawrence FR. Using latent growth models to evaluate longitudinal change. In: Hancock GR, Mueller RO, eds. *Structural equation modeling: A second course*. Greenwich, CT: Information Age; 2006:171–196.

44. Kline RB. *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press; 2011:427. doi:10.1038/156278a0.

45. Browne MW. Asymptotically distribution-free methods for the analysis of covariance structures. *Br J Math Stat Psychol*. 1984;37 (Pt 1):62–83. doi:10.1111/j.2044-8317.1984.tb00789.x.

46. Mardia KV. Measures of multivariate skewness and kurtosis with applications. *Biometrika*. 1970;57(3):519–530. doi:10.1093/biomet/57.3.519.

47. Mardia K. Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhya Indian J os Stat Ser B*. 1974;36:115–128. doi:10.2307/25051892.

48. Yuan K-H, Bentler PM. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociol Methodol*. 2000;30(1):165–200. doi:10.1111/0081-1750.00078.

49. Muthén B, Asparouhov T. *Using Mplus Monte Carlo Simulations In Practice : A Note On Non-Normal Missing Data In Latent Variable Models*.; 2002:7.

50. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107:238–246. doi:10.1037/0033-2909.107.2.238.

51. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model A Multidiscip J*. 1999;6(1):1–55. doi:10.1080/10705519909540118.

52. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973;38:1–10. doi:10.1007/BF02291170.

53. Steiger JH, Lind JC. Statistically based tests for the number of common factors. *Annu Meet Psychom Soc*. 1980:–.

54. Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. *Sociol Methods Res*. 1992;21(2):230–258. doi:10.1177/0049124192021002005.

55. Bentler PM, Chou CP. Practical issues in structural modeling. *Sociol Methods Res*. 1987;16:78–117. doi:10.1177/0049124187016001004.

56. Bentler PM. *EQS: Structural Equations Program Manual*. Encino, CA; 1995.

57. Hoogland JJ, Boomsma A. Robustness Studies in Covariance Structure Modeling: An Overview and a Meta-Analysis. *Sociol Methods Res*. 1998;26:329–367. doi:10.1177/0049124198026003003.

58. Thoemmes F, MacKinnon D, MR R. Power analysis for complex mediational designs using Monte Carlo methods. *Struct Equ Model*. 2010;17(3):510–534. doi:10.1080/10705511.2010.489379.POWER.

59. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods*. 1996;1(2):130–149. doi:10.1037/1082-989X.1.2.130.

60. Hancock GR, Freeman MJ. Power and Sample Size for the Root Mean Square Error of Approximation Test of not Close Fit in Structural Equation Modeling. *Educ Psychol Meas*. 2001;61(5):741–758. doi:10.1177/00131640121971491.

61. Belfort MB, Rifas-Shiman SL, Kleinman KP, et al. Infant feeding and childhood cognition at ages 3 and 7 years: Effects of breastfeeding duration and exclusivity. *JAMA Pediatr*. 2013;167(9):836–44. doi:10.1001/jamapediatrics.2013.455.

62. Brion M-JA, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*. 2011;40(3):670–80. doi:10.1093/ije/dyr020.

63. Quigley MA, Hockley C, Carson C, Kelly Y, Renfrew MJ, Sacker A. Breastfeeding is associated with improved child cognitive development: a population-based cohort study. *J Pediatr*. 2012;160(1):25–32. doi:10.1016/j.jpeds.2011.06.035.

64. Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Arch Gen Psychiatry*. 2008;65(5):578–84. doi:10.1001/archpsyc.65.5.578.

65. Flacking R, Wallin L, Ewald U. Perinatal and socioeconomic determinants of breastfeeding duration in very preterm infants. *Acta Paediatr*. 2007;96(8):1126–30. doi:10.1111/j.1651-2227.2007.00386.x.

66. Luiz DM, Foxcroft CD, Stewart R. The construct validity of the Griffiths Scales of Mental Development. *Child Care Health Dev*. 2001;27(1):73–83. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11136343.

67. Sung IK, Vohr B, Oh W. Growth and neurodevelopmental outcome of very low birth weight infants with intrauterine growth retardation: comparison with control subjects matched by birth weight and gestational age. *J Pediatr*. 1993;123:618–624.

68. Knops NBB, Sneeuw KC a, Brand R, et al. Catch-up growth up to ten years of age in children born very preterm or with very low birth weight. *BMC Pediatr*. 2005;5:26. doi:10.1186/1471-2431-5-26.

69. Bollen KA, Curran PJ. *Latent Curve Models. A Structural Equation Perspective*. New York, NY: Wiley; 2006:285 + xii.

70. Wang J, Wang X. *Structural Equation Modeling: Applications Using Mplus*. Chichester, West Sussex UK: Wiley; 2012:478.