Health Services & Outcomes Research Methodology 4: 109–126, 2003

Changes Can be Studied When the Measurement Instrument is Different at Different Time Points

NICOLET C.M. THEUNISSEN* nicolet@ntheunissen.nl; www.ntheunissen.nl
TNO Human Factors, Soesterberg, The Netherlands

JACQUELINE J. MEULMAN meulman@fsw.leidenuniv.nl
Department of Data Theory, Leiden University, The Netherlands

A. LYA DEN-OUDEN l.denOuden@pg.tno.nl
TNO Prevention and Health, Leiden, The Netherlands

HENDRIK M. KOOPMAN h.m.koopman@lumc.nl
Department of Paediatrics, Leiden University Medical Center, The Netherlands

ERIK G.H.W. VERRIPS ghw.verrips@pg.tno.nl
TNO Prevention and Health, Leiden, The Netherlands

S. PAULINE VERLOOVE-VANHORICK sp.verloove@pg.tno.nl
TNO Prevention and Health, Leiden, The Netherlands; Department of Paediatrics, Leiden University Medical Center, The Netherlands

JAN-MAARTEN WIT j.m.wit@lumc.nl
Department of Paediatrics, Leiden University Medical Center, The Netherlands

Received February 21, 2001; Revised April 18, 2003; Accepted April 19, 2003

Abstract. This paper presents a strategy for the analysis of longitudinal data that suffer from differences in measurement instruments over time. For illustration, we used longitudinal Health Status (HS) data from a cohort of preterm born children. The strategy solved three problems: First, commonly known HS domains were not explicitly used. Therefore, expert ratings were used to fit the variables into HS domains. Second, it was unclear which HS variables at 5 years of age matched which variables at 10 years of age. Principal component analyses (PCA) was used to match the data between measurements. Third, the change between points of measurement were estimated by longitudinal PCAs using the matched HS data. The impact of background variables such as gender and birth weight on HS changes was studied. This strategy using an imperfect data set provided valuable information about the development of HS in children born preterm.

Keywords: principal component analysis, CATPCA, longitudinal studies, health status, preterm birth

*Author to whom correspondence should be addressed.

Introduction

A longitudinal design is useful for studying the late effects of perinatal factors such as preterm birth. Unfortunately, this design has some drawbacks including drop outs [10, 21], selection bias and lengthy and expensive commitment. In addition, there is the necessity to use a consecutive series of age specific instruments, which may be difficult to compare between ages [20]. Instruments may also change to reflect scientific advances. For instance, health outcome studies in children until the 80s emphasized the physical functioning approach. More recently, a broader concept of quality of life has been commonly used which also includes psychological and social functioning. New standardized instruments are developed constantly and even a standardized instrument can change over time. Consequently, the measurement instruments may differ at different time points, which makes it difficult to analyze longitudinal changes.

In this paper an analysis strategy is presented that can be used to overcome these obstacles and real longitudinal data are used as an illustration. Before presenting our approach, we first describe the data in some detail, and discus and critique how such data are commonly analyzed.

Illustrative data

The data that we use in this paper is a set of longitudinal data collected from 688 children, preterm born in 1983 and enrolled in the follow-up program 'Project On Preterm and Small for gestational age infants (POPS)'. Details of the design and the results of this project can be found elsewhere [4, 17, 18], but we will give a short description here. Originally, the project was designed to investigate the relation of prenatal and perinatal factors with mortality and morbidity. The survival rate of these children has improved and therefore scientific interest has shifted from mortality and morbidity towards Health Status (HS) in general, described as a combination of physical-, psychological as well as social functioning. Data were collected using trained pediatricians and parents as informants. For the present purposes we will only use the parent reports. Although much HS information obtained from the parents was available, the early part of the POPS project was designed to fit longitudinal analyses using the physician and not the parent as main informant. Furthermore, financial resources varied considerably over time. As a result, the amount and content of HS variables differed between points of measurement, as well as the administration procedure (a combination of mail and interview at age 5, postal questionnaires only at age 10). In the next section we will describe how this kind of data are commonly analyzed and explain the limitations. Next, we will describe a strategy to overcome this limitation.

Common approaches

In the data set we used, the data was not primarily organized to measure HS using the parents as informants. Furthermore, the number and the content of variables varied somewhat between points of measurement. This problem is not unique. For instance, Caputo et al. used age specific instruments to assess development in children with low birth weight [1]. The Cattell Infant Intelligence Scale and the Gesell Developmental Schedules were used at 1 year of age. The Wechsler Intelligence Scale for Children-Revised and the Visual Motor Gestalt Test were used at 7 to 9 years of age. Although from both points of measurement a developmental quotient (DQ) was obtained, the content of these DQs was probably considered to be too age specific. Therefore, analysis was restricted to correlations. In half of the longitudinal studies on preterm born children we inspected, no longitudinal HS analyses [13] were reported, or analyses were restricted to correlations between assessments and changes in HS [1, 6, 7, 9, 12].

Saigal et al. [11] had a similar problem, but used a different approach. They wished to study data already collected in a group of children with extremely low birth weight compared to a control group. These data included a battery of standardized psychometric measures and various clinical data. Therefore, Saigal et al. [11] applied the Multi-attribute HS (MAHS) classification system retrospectively to the data, using particular computer algorithms. However, these computer algorithms were entirely based on expert ratings. This means that experts not only select which data to use, but also determine how to transform the scaling of the data. A new data set would therefore require new expert consensus in order to create new computer algorithms. In the present paper we suggest an alternative strategy to cope with differences between points of measurement.

Questions to be answered

Using our illustrative data set, the following three questions were answered:

1. Which variables are considered HS variables?
2. Which HS variables at 5 years of age match which variables at 10 years of age?
3. How to estimate change when categories are different between points of measurement?

1. Which variables are considered HS variables?

In our data set commonly used HS domains were not explicitly mentioned. HS is in the literature divided into subdomains like physical complaints, motor functioning, autonomy, cognitive functioning, emotional functioning and social functioning [19]. Assignment of the variables at age 5 and 10 to HS subdomains depends on the content of the questions and therefore we can not do without expert opinions here.

Strategy of variable selection. First, we confronted five experts with the total list of variables that were assessed by the parents. These experts were two pediatricians (A.O, S.V., see the authors' names) and three developmental psychologists (N.T, H.K, E.V.). They were asked to give their independent opinion whether or not a variable belonged to one of the six subdomains. If this was the case, they wrote the number of the domain behind the variable at the list. This resulted in five independent judgements about each variable. Second, a variable was selected if at least three out of five experts agreed that the variable belonged to a HS sub-domain. If less than three out of five experts agreed, selection was reconsidered by a developmental psychologist and a pediatrician (N.T, A.O.). The final result was again presented to the forum of experts, and they approved of the variable selection made. As a result, a set of 34 variables at 5 years of age and a set of 47 variables at 10 years of age became available (see Appendix).

2. Which HS variables at 5 years of age match which variables at 10 years of age?

At different points in time, different sets of variables were used. Although at first sight the wording of the variables presented in the Appendix seems to match between ages, many variables were not equal in wording in the Dutch original of both question and categories (see for instance the items in physical complaints or motor functioning). Furthermore, the layout and content of the questionnaires was different, which may have influenced the way they were answered. Therefore, we had to compare variables between points of time in order to find a match that could serve in the longitudinal analyses.

Analysis method. To match the selected HS variables we used a generalization of classical principal component analysis (PCA), called CATPCA) [8] (also, see [5, 16] for other applications). CATPCA is a Principal Components Analysis for categorical data or mixed categorical and interval level data. CATPCA is included in SPSS Categories 10.0 [8] (and was formerly known as PRINCALS, which is included in earlier versions of SPSS Categories). CATPCA simultaneously fits the principal components model and finds optimal quantifications (scores) for categorical variables. The ordinal or nominal information in the categorial variables is retained in the optimal quantifications. Replacing the categories of the original variables with the optimal scores, results in optimally transformed variables. Optimal quantification in CATPCA implies that the average proportion of variance accounted for by the transformed variables is as large as possible. CATPCA analyzes the data matrix, not the correlation matrix, obtaining a joint representation of subjects and variables in a number of dimensions (dimensions correspond to principal components).

The results of a CATPCA analysis can be divided into three main parts: quantifications for the categories of the variables, scores for the subjects (with mean of zero and unit variance), and component loadings for the transformed variables. In CATPCA, treatment of missing data can be specified per variable. There are four options: impute the mode of the variable, impute an extra category, passive treatment (exclude missing values on a variable from the analysis), and listwise deletion (exclude objects with missing value on a variable from the analysis). More details can be found in [2], and in the CATPCA Algorithm Description, which can be found at http://www.spss.com/tech/stat/Algorithms.htm). We used passive treatment of missing data in our analyses. This means that in optimizing the quantifications of a variable, only subjects with valid values on the variable are involved, and that only valid values of a variable contribute to the solution.

Data preparation. To match variables we need to exclude the effect of change in time, therefore the matching analyses require a data matrix with subjects in the rows and variables at age 5 and at 10 in the columns. Variables to be included in a CATPCA analysis as ordinal or nominal preferably have well-filled categories. Some variables had categories with very low frequency. These variables were recoded to merge the low-frequency category with another category.

Step A: Grouping variables to control for domains. First, all HS variables were included in a categorical CATPCA analysis to see if the variables would group together by HS sub-domains, regardless of time of measurement. This analysis accounted for a proportion of total variance of 19% (dimension 1 = 13%, dimension 2 = 5%). Additionally, CATPCA reports the eigenvalues (amount of variance accounted for). The eigenvalues are related to Cronbach alpha coefficient [3], which is also reported. We found a Cronbach's alpha of .92 for dimension 1 and an alpha of .78 for dimension 2. The relationships between the variables are displayed graphically in figures 1 and 2. In figure 1, the positions of the variables and their distance to the center of the graph (zero point) are determined by their correlations with the principal components, the so-called component loadings. The distances between the variables (inversely) represent the correlations between them. As can be seen in figure 1, there is a very strong first component (Dimension 1): all variables point in the positive direction. In the second dimension, the analysis partitioned the variables into two clusters: motor (M), autonomy (A) and cognitive (C) variables at the top, emotional (E) and social (S) variables at the bottom, and physical (P) variables in between. The clustering of motor functioning with autonomy, and emotional with social functioning is considered an obvious one. Apparently, the children with motor-autonomy problems experience cognitive problems as well. A more important result of this analysis is that the variables of each HS sub-domain are grouped together, irrespective of time of assessment. Therefore, the variables in these subdomains could be matched at different points in time.

Step B: Likely combinations of variables. Fifteen combinations of variables were inspected that a priori were considered to be good matches according to their meaning (see also the Appendix). For instance, the variable 'cannot concentrate' at age 5 (c1) was almost identical in formulation to a similar variable at age 10 (C1) Their position in relation to the center of the graph is alike and the distance between these variables is small. This is shown in figure 2, a close-up of figure 1. At age 5 another cognitive variable was assessed as being analogous in meaning to the 'cannot concentrate' variable mentioned previously, which is 'problems with concentration' (c2). Figure 2 shows that the distance between these variables at age 5 is small. As we already found a match for c1, this last variable could be removed from further analyses. Other variables, such as 'finding words' (c8), were not repeated which made matching impossible.

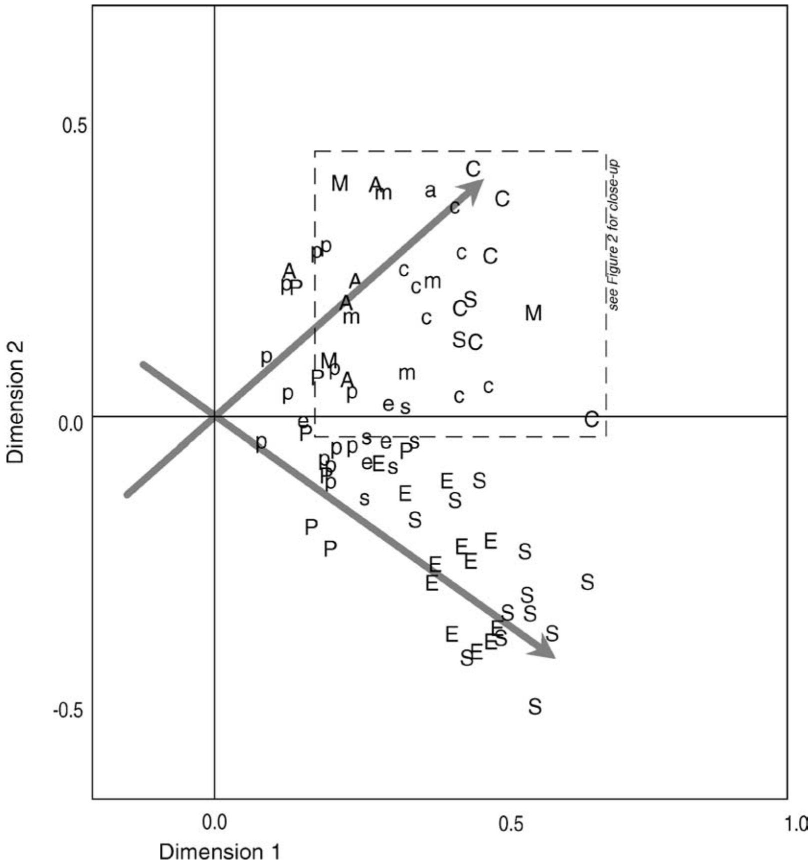

Figure 1. Relationships between HS variables at all different times of measurement, resulting from the quantitative selection. No capitals = items at 5 years of age; CAPITALS = items at 10 years of age; C = cognitive functioning, S = social functioning, P = physical complaints, M = motor functioning, A = autonomy, E = emotional functioning (See text for more clarification).

Step C: New combinations. For some variables, the matching at different time points was debatable and needed further analyses to demonstrate their usefulness. Therefore, a second CATPCA was performed including the 18 combinations mentioned above, excluding variables analogous in meaning to these 18 combinations, and excluding HS aspects that were not repeated in time, but including doubtful cases. The second analysis accounted for a proportion of total variance of 22% (Cronbach's alpha's: dimension 1 = .88, dimension 2 = .68). The same two main clusters were found as in the first analysis, confirming that the selected variables still covered the same HS aspects. In addition to the 15 combinations of variables already established, one new combination was found. *A priori* the two variables did not have the same content, but it could be argued that they belonged together: 'stiff movements' at age 5 and 'walking' at age 10 both represent problems with moving.

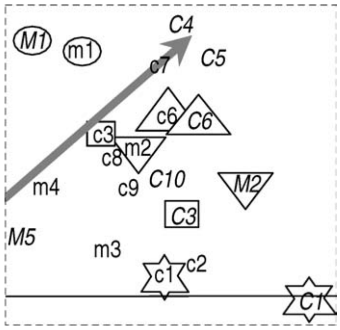

Figure 2. A close up from figure 1, displaying the cognitive functioning and motor functioning items. The numbers in the figure correspond with the numbers in the Appendix. The circles, triangles, stars or squares mark the partner variables.

Step D: Final combinations. A third CATPCA was performed with the 15 old and the one new combination, which accounted for a proportion of total variance of 23% (Cronbach's alpha's: dimension 1 = .89, dimension 2 = .55), with the same two bundles found as in the first analysis. An additional complication was that the categories of two variables to be combined were not comparable, so the categories had to be merged in a rational way. To do this, the recoding was done by using the optimal quantifications given by CATPCA for the separate variables. Table 1 shows how the categories were merged by the authors: the optimal quantifications of both variables were visually compared and used

Table 1. Matching categories by using optimal scaling to obtain a new longitudinal variable: 'moving problems' and 'speech problems'.

Original ordinal categories	Quantification	Original ordinal categories	Quantification	New ordinal categories combined by optimal scaling
Stiff movements (at 5 yr.)		Walking (at 10 yr.)		Moving problems
1 = rarely	−0.48	1 = very good	−0.35	1
2 = occasionally	0.33			2
3 = rather often	1.44			3
		2 = dubious or barely	2.77	4
4 = almost always	3.28			5
Others understand this child (at 5 yr.)		Other people understand what child says (at 10 yr.)		Speech problems
1 = almost always	−0.43	1 = very good	−0.44	1
2 = most of time	0.65			2
		2 = dubious or barely	2.24	3
3 = rarely or not	3.71			4

for new ordinal categories. Final combinations of 16 variables are given in the Appendix (in boxes).

3. How to estimate change when categories are different between points of measurement?

We have finally come to the central objective of this exercise: the longitudinal analyses. Usually in longitudinal analysis on complete datasets, repeated measurement MANOVAs are used. However, because the variables in our data set had different categories at different points of measurement, an alternative strategy had to be followed here as well. The aim of the longitudinal analysis is twofold: to investigate the relationships between HS variables over time, and the relationships between background variables and HS development. These aims were simultaneously addressed by a new CATPCA.

Data preparation. To study change over time, the data were reorganized as follows: The columns of the data matrix contained the 16 selected variable combinations, and the rows the observations on the 688 children at time point 5 and 10 years of age, placed below each other. Thus, the data matrix contains 1376 rows, each subject appearing two times. The CATPCA analysis will give one set of optimal quantifications and one set of component loadings for the variables, but two sets of scores for each subject, one for each time point, displaying the development over time with respect to the variables. In addition to this data matrix, categorical interaction variables were created combining the information about time point and each of the background variables gender and birth weight. Because there were two time points and gender has 2 categories, and birth weight 3 categories, the time $\times$ gender variable has 4 categories, and the time $\times$ birth weight variable 6 categories.

Step 1: Performing the longitudinal analyses. The longitudinal CATPCA analysis was based on the original 16 HS variables. To investigate the relation between the time $\times$ background variables and the structure found using the 16 HS variables, the time $\times$ background variables should be treated as supplementary variables, that is, they should not be included in the analysis, but fitted into the recovered structure between the 16 HS variables. CATPCA has an option to treat variables as supplementary. The 16 HS variables were given an ordinal optimal scaling level, while the time $\times$ background variables were treated as nominal variables because their categories are unordered. The proportion of variance accounted for by the longitudinal CATPCA was 31%. Dimension 1 accounts for 22% (alpha = .76) and dimension 2 for 9% (alpha = .36).

Step 2: Inspecting ordinality of HS variables. At the vertical y-axis in figure 3, category quantifications are displayed for ordinal variables having more than two categories. The zero score represents the mean score of the total sample over two time points. The horizontal *x*-axis gives the original category numbers. As can be seen in the figure, the categories of the presented variables are ordinal and in four out of six cases almost at interval level. In all HS variables the lowest category represents the absence of a

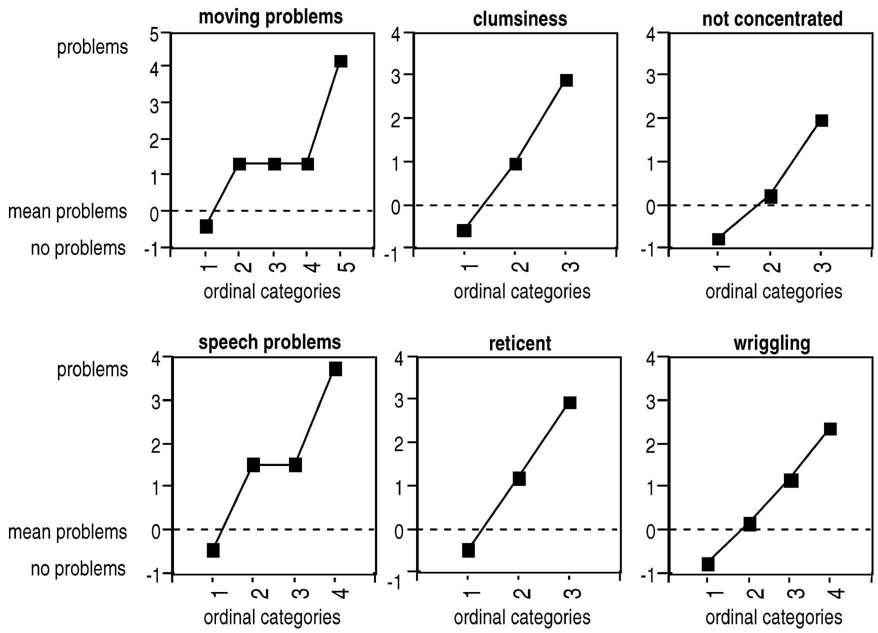

Figure 3. Category optimal scaling scores of variables with more than two categories, resulting from the longitudinal analysis.

problem (see Appendix for the category contents). Only the lowest ordinal categories, representing no problems, have quantification scores below zero. The majority of the other categories, representing HS problems, obtained scores higher than zero. It can be concluded that a score of zero represents minor problems, which in turn corresponds with the fact that the study group on the whole had minor HS problems on average.

Step 3: Interpreting the first dimension. Using the CATPCA results, the HS variables can be depicted graphically as vectors (arrows) in figure 4. Like in figure 3, the zero-point in figure 4 represents the mean component score of the sample. All variables have a positive correlation with the first dimension (horizontal x-axis) resulting in a Cronbach's alpha of .76. Since the coding of these variables, implies that high category scores mean more HS problems, the underlying factor of Dimension 1 is HS problems. 'Lung problems' does not fit very well (short vector) and seems to be not important as an HS indicator. The lower score ends of the vectors (at the left-hand side of figure 4) indicate the absence of problems such as 'not tense' and 'no speech problems'. Not shown in this figure is that seventeen percent of the subjects have their component scores in the area spanned by the no-problem ends of the vectors at age 5 but not at age 10. This indicates no HS problems yet at 5, but HS problems by 10 years of age. Sixteen percent had no HS problems at age 10 although they had HS problems at age 5. Thirty-five percent of the

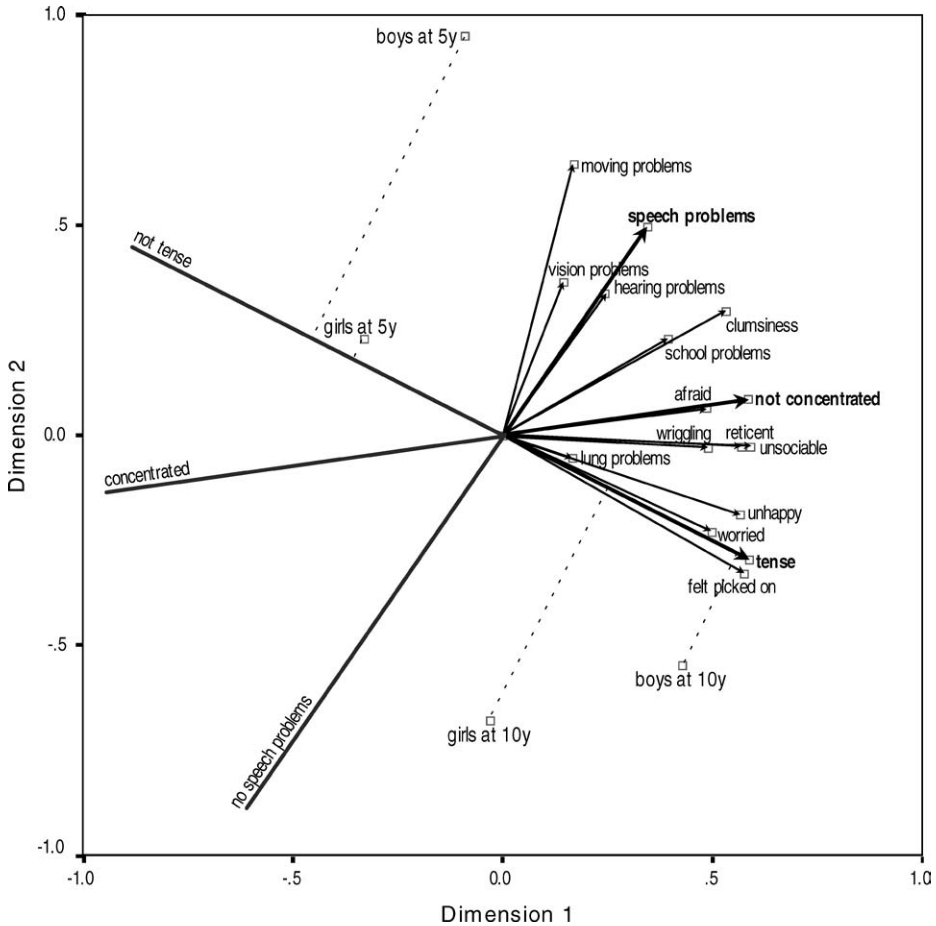

Figure 4. Relationships between HS variables in combination with gender displayed over time.

subjects had HS problems at both ages and 32% of the sample had no HS problems at both ages.

Step 4: Interpreting the second dimension. To investigate the nature of the problems at 5 or 10 years of age, the second dimension was inspected. The second dimension (vertical axis) separates the HS variables: It distinguishes between physical and motor problems in the upper part of the Figure, and emotional and social problems in the lower part. Cognitive problems are positioned in between. Smaller angles between the vectors represent higher correlations between the variables. The bundle of 'speech problems,' 'moving problems,' 'vision problems' and 'hearing problems' has hardly any correlation with the bundle 'tense,' 'unhappy,' 'worried' and 'felt-picked-on.' The first bundle can be interpreted as problems in basic functioning and the last as negative moods. Between these bundles is a bundle close to the horizontal axis with 'school problems,' 'not concentrated,' 'unsociable,' 'wriggling,' 'reticent,' 'afraid,' and 'clumsiness.' Together they indicate concentration problems with both cognitive and behavioral components.

Step 5: A closer look at the bundles. Usually, Cronbach's alpha between variables is computed on the basis of the original variables. In our case, the variables were optimally transformed. CATPCA produces transformed variables, which are approximated by the sum over dimensions of the component scores weighted by the component loadings in each dimension (The approximated scores for a variable are obtained by multiplying the subject (component) scores by the variable's component loadings, for both dimensions, and then summing over dimensions). Looking at the representation of the variables in two dimensions, Cronbach's alpha for the three variable bundles were computed on the estimated scores. Basic functioning problems had a Cronbach's alpha of .97, concentration problems had an alpha of .99, and negative moods had an alpha of .99 as well.

Step 6: Selecting key variables. Often, bundles of variables are added up to form scales. As an alternative, we may select from each bundle of variables a key variable, to represent the other variables in the bundle for further investigation. Selection was not based on the highest loading, but on the representativeness for the other variables. This representativeness was based both on content and on a central position of the key variable in the bundle. However, choosing another variable from the bundle as the key variable will give highly comparable results. From the basic functioning bundle we selected 'speech problems,' from the negative moods bundle we selected 'tense,' and from the concentration bundle we selected 'not concentrated.'

Step 7: Relationships between background variables and HS development. In figure 4 the relationships between the time $\times$ gender variable and the HS variables are graphically presented by displaying the category quantifications of the time $\times$ gender variable in the plot of the component loadings of the HS variables. The interpretation of the relationships is obtained by projecting the category points onto the vectors representing the HS variables. For example, the dotted lines in figure 4 point to the position of boys and girls at age 5 and age 10 with respect to 'tense' (the negative moods bundle). We see a large time effect; both boys and girls score low on 'tense' at age 5 compared to age 10, and boys at age 10 score higher than girls. As another example, projection of the category points onto 'speech problems' (the basic functioning bundle), shows a decrease of problems from age 5 to age 10. The relations between gender, time, and the key variables are more clearly displayed in the figure 5(a)–(c). Moreover, the relations of the other two background variables with the three key variables are displayed as well. The plots in figure 5 were derived from figure 4 in the following way: projection of the category quantifications of a time $\times$ background variable on a HS variable gives projected scores for the time $\times$ background variables. These projected scores are displayed in figure 5 (such plots are optional output in CATPCA). In addition we have connected the projected scores for time point 1 and time point 2 to show the development over time more clearly. Differences between groups in arrow length shows group effects, total length shows change over time

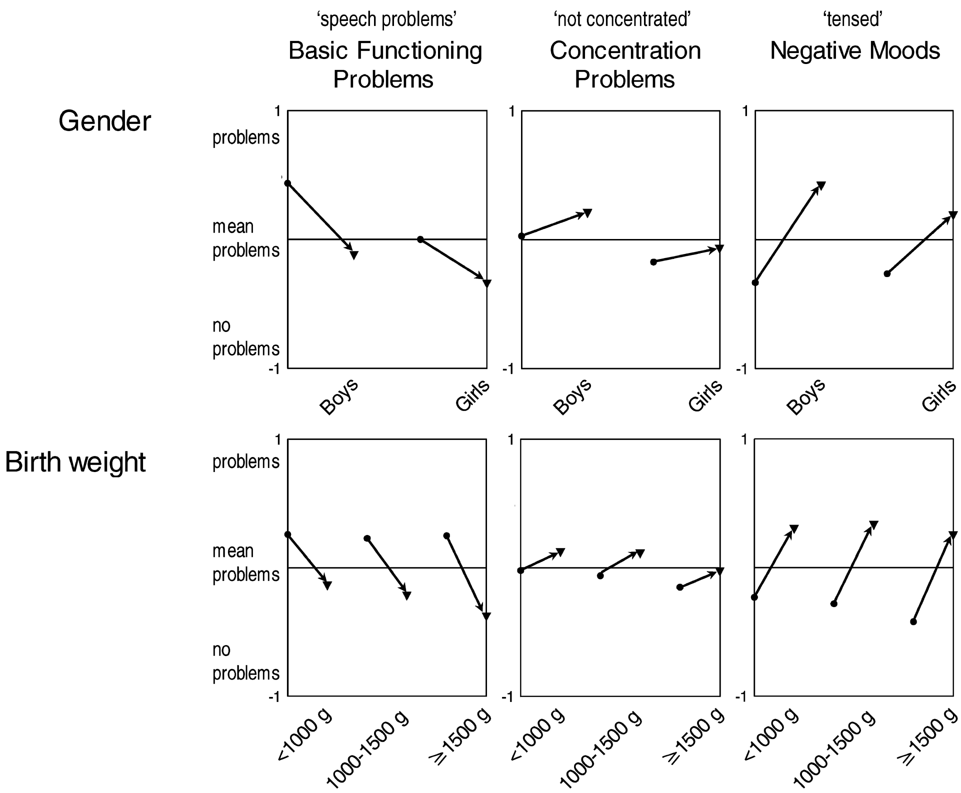

Figure 5. Background variables at 5 and 10 years of age with respect to basic functioning problems, concentration problems and negative moods. 5a = Gender, 5b = Birth weight, ● = 5 years of age, ▼ = 10 years of age.

and parallel lines indicate no interaction between group and time. Differences with respect to the optimal scaled key variables were tested with Repeated Measures ANOVA, testing the sources of variation between groups and within groups between age 5 and 10 (see Table 2).

Final results. Overall a large time effect is shown. The amount of basic functioning problems tends to decrease in time, negative moods increase and concentration problems increase slightly in all background variable groups. Boys have more problems in basic functioning and more concentration problems than girls (see figure 5(a)). At 5 years of age, boys have less negative moods than girls. At 10 years of age, boys have slightly more negative moods than girls. The group differences between boys and girls as well as time differences between 5 and 10 years of age are significant. In addition, gender × time interactions are significant in basic functioning problems and negative moods. It can be concluded that boys have more problems than girls at both points of measurement, although the nature of the problems are the same for both sexes. Differences between birth weight groups are

Table 2. Repeated measures analysis of variance between birth-factors and age of measurement with respect to the three key variables.

		Basic functioning problems	Concentration problems	Negative moods
		'speech problems' F	'not concentrated' F	'tensed' F
Gender	between subjects (groups)[a]	88.2***	40.1***	4.6*
	within subjects (age)	450.0***	37.7***	608.4***
	interaction	25.1***	2.8	37.8***
Birth weight (g)	between subjects (groups)	1.9	3.4*	3.1*
	within subjects (age)	338.1***	28.3***	406.3***
	interaction	6.3**	0.4	1.5

[a]Differences between measurements at 5 and 10 years of age; $^{*}p \leq .05$, $^{**}p \leq .01$, $^{***}p \leq .001$.

significant for negative mood and concentration problems only (see Table 2). Lower birth weight relates to more HS problems (see figure 5(b)).

Validity of the strategy

Expert ratings were used to fit variables from a large pool into HS domains, which supported content validity of the initially selected variables. During the analysis performed to match variables over time, the variables of each HS sub-domain grouped together, irrespective of time of assessment (see figure 1). In turn, this proved the construct validity of the expert ratings.

Although it appropriate to analyze ordinal variables with optimal scaling techniques, it could be questioned whether the variables that were matched at age 5 and age 10 would have grouped together when not quantified by the CATPCA. Therefore, Spearman's rho correlations were computed using the 16 final pairs of original variables at age 5 and age 10. Convergent validity was confirmed completely: all original variables at 5 years of age correlated significantly ($p < .01$) with their partner variables at 10 years of age (results not shown). Discriminant validity was confirmed for the most part: 13 out of 16 of the variables at 5 years of age correlated better with their partner variable at 10 years of age than with other variables (results not shown). From the three original variables that failed discriminant validity, two correlated higher with another variable within the same problem area (concentration problems). Thus, the common approach of using rank-correlations would basically have given the same relationships between variables, and construct validity would have been supported.

The main problem areas we found and the order in which they appeared—basic functioning, concentration and negative moods—are considered to be clinically and psychologically relevant, and are supported by findings in the literature (for instance, [14]). More pediatric details about the HS changes of the POPS cohort are described elsewhere [15]. Groups at the highest risk were children born preterm with handicaps, with a low birth weight corrected for gestational age, sex and parity, and boys. It is of interest that the groups which appear at risk according to the present method, are known risk groups this can be interpreted as a criterion related validation of the study's method.

Conclusion and discussion

We presented a to reconstruct changes in HS in spite of the initial differences in measurement instruments at different time points. As an illustration, a challenging longitudinal data set was used. To start with, it was unclear which variables were considered HS variables. Furthermore, it was not always clear which HS variables at 5 years of age matched which variables at 10 years of age. Finally, change between points of measurement had to be estimated in spite of differences between categories. By using the present reconstruction technique, it was unnecessary to restrict analyses to correlations between assessments and we could study changes in HS. Furthermore, expert ratings were used to select but not to recode the variables. Interpretation of which original score represents which problem level is entirely empirically based. Consequently, it was possible to stay close to the original data. As a bonus, instead of studying the changes in HS subdomains one by one as previously with children born preterm [1, 9, 13, 14], the present method allows for new combinations of domains. Consequently, it is possible to pinpoint the problems that are relevant for the development of preterm born children.

The longitudinal results of the present study may in fact be an artifact of the reconstruction method. At age 5 the parents completed the questionnaire together with the pediatrician 'if necessary'. At age 10 the pediatrician was not present. The parents may have been biased at age 5 by the stress put on basic functioning problems by pediatricians, since clinicians emphasize other HS problems than parents do. Nevertheless, the initial selection of HS variables, the relationship between partner variables, the main problem areas we found, and the order in which they appeared—basic functioning, concentration and negative moods—are considered valid. In the future, prospective research could be used to judge the validity of the findings.

The described strategy is complex some details had to be omitted. Researchers considering this strategy can obtain more details from the authors.

In conclusion, longitudinal studies have to cope with the consequences of using age specific instruments and with changes in scientific ideas during the research period. In this paper we described a strategy to overcome the problem of using different measurement instruments at different time points. At the same time, we stayed as close to the original data as possible. The 'common knowledge' results we found in preterm born children gave confidence in the method used. Although prospective research will be needed to judge the validity of the findings, the method presented allows an imperfect data set to provide valuable information about changes in health status.

Appendix

The qualitative selection result of health status variables at 5 and 10 years of age (translated from the Dutch original) is shown below.

5 years of age	10 years of age	combined label
Cognitive functioning		
c1: My child cannot concentrate, cannot focus on something for a long time *1=not applicable, 2= somewhat, 3=clearly/often*	**C1:** Cannot concentrate, cannot focus on something for a long time *1=not at all, 2= somewhat/ sometimes, 3=clearly/often*	not concentrated
c2: problems with concentration		
c3: My child gets low grades at school *1=no, 2= somewhat or clearly/often*	**C3:** Low grades at school *1=not at all, 2= somewhat/ sometimes or clearly/often*	school problems
	C4: school results: language or Dutch **C5:** school results: math or rythmetics	
c6: Children are often better understood by their parents than by other people. Others understand my child *1=almost always 2=most of time, 4=rarely or not*	**C6:** How well does other people understand what your child says *1= very good, 3=dubious or barely*	speech problems
c7: talking **c8:** finding words **c9:** intelligibility/audibility	**C10:** how well does the child understand others?	
Motor functioning		
m1: The movements of my child make a stiff and wooden impression *1=rarely, 2=occasionally, 3=rather often, 5=almost always*	**M1:** How well does your child walk *1= very good, 4=dubious or barely*	moving problems
m2: My child is clumsy and butterfingered *1=rarely, 2=occasionally, 3=rather often or almost always*	**M2:** Clumsy or bad coordination *1=not at all, 2= somewhat/ sometimes, 3=clearly/often*	clumsyness
m3: bumping into something **m4:** trips over doorsteps	**M5:** slow movements, or lack of energy	
Physical complaints		
Seeing without glasses *1=good, 2=dubious or bad*	How well does your child see (if necessary with glasses or contact lenses) *1= very good, 2=dubious or barely*	vision problems
squinting owns glasses		
Hearing without hearing aid *1=good, 2=dubious or bad*	How well does your child hear *1= very good, 2=dubious or barely*	hearing problems
ear inflammation		
Asthmatic attacks *1=never, 2=sometimes(<3/y.) often(3-10/y.) or very often (>10/y)*	Asthma *1=not at all, 2= somewhat/ sometimes or clearly/often*	lung problems
throat inflammation tightness of the chest short of breath wheezy breathing coughing bronchitis pseudocroup attacks	stomach ache, without medical explanation headache without medical explanation nauseous without medical explanation exhausted	

To be continued at the next page

Continuation of the Appendix

5 years of age	10 years of age	combined label
Autonomy		
	sports achievement compared with peers	
problems with laces and buttons	medical reasons for not joining sports or gymnastics at school	
	how well is the child performing hobbies in comparison with peers	
	wet one's pants	
	wet one's bed	
Social functioning		
My child feels that others pick on him/her *1=not applicable, 2= somewhat or clearly/often*	Feels that others pick on him/her *1=not at all, 2= somewhat/ sometimes or clearly/often*	felt picked on
	be ragged a lot	
My child can't get on well with other children *1=not applicable, 2= somewhat or clearly/often*	Can't get on well with other children *1=not at all, 2= somewhat/ sometimes or clearly/often*	unsociable
	disrupts other children's games	
	activities with friends	
	getting along with other children	
	others don't like him/her	
	how many really good friends?	
My child is reticent, others do not know what is on the child's mind *1=not applicable, 2= some what, 3=clearly/often*	Reticent, others do not know what is on the child's mind *1=not at all, 2= somewhat/ sometimes, 3=clearly/often*	reticent
My child is wriggling constantly *1=rarely, 2=occasionally, 3=rather often, 4=almost always*	Makes often restless movements with hands or feet or wiggles in chair *1=not at all, 2= somewhat/ sometimes, 3=rather often 4= very often*	wriggling
fiddling all the time	does not respond	
doesn't do what it has to do	disobedient at home	
	behavior towards parents	
	objects or arguments a lot	
Emotional functioning		
My child is unhappy, sad, depressed *1=not applicable, 2= somewhat or clearly/often*	Unhappy, sad, depressed *1=not at all, 2= somewhat/ sometimes or clearly/often*	unhappy
My child is worried *1=not applicable, 2= somewhat or clearly/often*	Is worried *1=not at all, 2= somewhat/ sometimes or clearly/often*	worried
My child is fearful or too afraid *1=not applicable, 2= some what or clearly/often*	Is fearful or too afraid *1=not at all, 2= somewhat/ sometimes or clearly/often*	afraid
	shy	
My child is nervous, tense *1=not applicable, 2= some what or clearly/often*	Nervous or tensed *1=not at all, 2= somewhat/ sometimes or clearly/often*	tense
	sulks and pouts	
	fights a lot	
	destroys its own belongings	
	cruel or mean to others	
	physically attacks others	
	outbursts of anger	
	quickly jealous	

Acknowledgments

We would like to thank all the previous participants from the Collaborative Project on Preterm and Small for Gestational Age (POPS) Infant in the Netherlands. We would also like to express our deepest appreciation to Drs. A.J. van der Kooij and Prof. L.J. Hubert for sharing their knowledge. We are indebted to the parents for their participation.

References

1. Caputo, D.V., Goldstein, K.M., and Taub, H.B., "Neonatal compromise and later psychological development: A 10-year longitudinal study," in *Preterm birth and psychological development* (S.L. Friedman and M. Sigman, eds.) Academic Press, New York, etc., 353–385, 1981.
2. Gifi, A., *Nonlinear multivariate analysis*, Wiley, Chichester, 1990.
3. Heiser, W.J. and Meulman, J.J., "Homogeneity analysis: Exploring the distribution of variables and their nonlinear relationships," in *Correspondence analysis in the social sciences: Recent developments and applications* (M. Greenacre and J. Blasius, eds.), Academic Press, Harcourt Brace & Company, London, San Diego, New York, Boston, Sydney, Tokyo, Toronto, 179–209, 1994.
4. Hille, E.T.M., den-Ouden, L., Bauer, L., Van-den-Oudenrijn, C., Brand, R., and Verloove-Vanhorick, S.P., "School performance at nine years of age in very premature and very low birth weight infants: Perinatal risk factors and predictors at five years of age," *J. Pediatr* 125, 426–434, 1994.
5. Kerkhof, A.J.F.M., van-der-Wal, J., and Hengeveld M.W., "Typology of persons who attempted suicide with predictive value for repetition: A prospective cohort study," in *Current issues of suicidology* (H.J. Möller, A. Schmidtke, and R. Welz, eds.), Springer-Verlag, Berlin, 194–203, 1988.
6. Largo, R.H., Graf, S., Kundu, S., Hunziker, U., and Molinari, L., "Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants," *Dev. Med. Child Neurol.* 32, 30–45, 1990.
7. Menyuk, P., Chesnick, M., Liebergott, J.W., and Korngold, B., "Predicting reading problems in at-risk children," *Journal of Speech and Hearing Research* 34, 893–903, 1991.
8. Meulman, J.J., Heiser, W.J., *SPSS Inc., SPSS Categories*, SPSS Inc., Chicago, 1999.
9. Portnoy, S., Callias, M., Wolke, D., and Gamsu, H., "Five-year follow-up study of extremely low-birthweight infants," *Dev. Med. Child. Neurol*. 30, 590–598, 1988.
10. Raboud, J.M., Singer, J., Thorne, A., Schechter, M.T., and Shafran, S.D., "Estimating the effect of treatment on quality of life in the presence of missing data due to drop-out and death," *Qual. Life Res.* 7, 487–494, 1998.
11. Saigal, S., Rosenbaum, P., Stoskopf, B., et al., "Comprehensive assessment of the health status of extremely low birth weight children at eight years of age: Comparison with a reference group," *J. Pediatr.* 125, 411–417, 1994.
12. Saigal, S., Szatmari, P., and Rosenbaum, P.L., "Can learning disabilities in children who were extremely low birth weight be identified at school entry?" *J. Dev. Behav. Pediatr.* 13, 356–362, 1992.
13. Saylor, C.F., Casto, G., and Huntington, L., "Predictors of developmental outcomes for medically fragile early intervention participants," *J. Pediatr. Psychol.* 21, 869–887, 1996.
14. Silva, P.A., McGee, R., and Williams, S., "A longitudinal study of the intelligence and behavior of preterm and small for gestational age children," *J. Dev. Behav. Pediatr.* 5, 1–5, 1984.
15. Theunissen, N.C.M., den-Ouden, A.L., Meulman, J.J., Koopman, H.M., Verloove-Vanhorick,, S.P., and Wit, J.M., "Health status development in a cohort of preterm children," *J. Pediatr.* 137, 534–539, 2000.
16. van-der-Ham, T., Meulman, J.J., van-Strien, D.C., and van-Engeland, H., "Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective," *Brit. J. Psychiat.* 170, 363–368, 1997.
17. Veen, S., Ens-Dokkum, M.H., Schreuder, A.M., Verloove-Vanhorick, S.P., Brand, R., and Ruys, J.H., "Impairments, disabilities, and handicaps of very preterm and very-low-birthweight infants at five years of age," *Lancet* 338, 33–36, 1991.
18. Verloove-Vanhorick, S.P., Verwey, R.A., Ebeling, M.C.A., Brand, R., and Ruys, J.H., "Mortality in very preterm and very low birth weight infants according to place of birth and level of care: Results of a national

collaborative survey of preterm and very low birth weight Infants in the Netherlands," *Pediatr*. 81, 404–411, 1988.

19. Vogels, T., Verrips, G.H., Verloove-Vanhorick, S.P., et al., "Measuring health-related quality of life in children: the development of the TACQOL—Parent Form," *Qual. Life Res.* 7, 457–465, 1998.
20. Wadsworth, M.E., Rodgers, B., and Wadsworth, M.E.J., "Longterm follow-up studies; a critical overview," *Rev. Epidemiol. Sante. Publique* 37, 533–547, 1989.
21. Zwinderman, A.H., "Statistical analysis of longitudinal quality of life data with missing measurements," *Qual. Life Res.* 1, 219–224, 1992.