BMJ Open

Dynamic prediction of childhood high blood pressure in a population-based birth cohort: a model development study

Marleen Hamoen,[1,2] Yvonne Vergouwe,[1] Alet H Wijga,[3] Martijn W Heymans,[4] Vincent W V Jaddoe,[2] Jos W R Twisk,[4] Hein Raat,[1] Marlou L A de Kroon[1,5]

[1]Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands
[2]Erasmus University Medical Center, The Generation R Study Group, Rotterdam, Netherlands
[3]National Institute for Public Health and the Environment, Center for Nutrition, Prevention and Health Services, Bilthoven, Netherlands
[4]Department of Epidemiology and Biostatistics, Amsterdam Public Health research institute, VU University Medical Center, Amsterdam, Netherlands
[5]Department of Health Sciences, University Medical Center Groningen, Groningen, Netherlands

Correspondence to
Marleen Hamoen;
m.hamoen@erasmusmc.nl

ABSTRACT

Objectives To develop a dynamic prediction model for high blood pressure at the age of 9–10 years that could be applied at any age between birth and the age of 6 years in community-based child healthcare.

Design, setting and participants Data were used from 5359 children in a population-based prospective cohort study in Rotterdam, the Netherlands.

Outcome measure High blood pressure was defined as systolic and/or diastolic blood pressure ≥95th percentile for gender, age and height. Using multivariable pooled logistic regression, the predictive value of characteristics at birth, and of longitudinal information on the body mass index (BMI) of the child until the age of 6 years, was assessed. Internal validation was performed using bootstrapping.

Results 227 children (4.2%) had high blood pressure at the age of 9–10 years. Final predictors were maternal hypertensive disease during pregnancy, maternal educational level, maternal prepregnancy BMI, child ethnicity, birth weight SD score (SDS) and the most recent BMI SDS. After internal validation, the area under the receiver operating characteristic curve ranged from 0.65 (prediction at age 3 years) to 0.73 (prediction at age 5–6 years).

Conclusions This prediction model may help to monitor the risk of developing high blood pressure in childhood which may allow for early targeted primordial prevention of cardiovascular disease.

Strengths and limitations of this study

► This is the first study in which a dynamic prediction model for childhood high blood pressure was developed, using longitudinal information on the child's body mass index.

► The prediction model is based on predictors that are usually recorded or are easy to obtain in community-based child care settings, and may in the future offer an opportunity for targeted primordial prevention of cardiovascular disease.

► The outcome predicted in this study was childhood high blood pressure, and not childhood hypertension, which could only have been diagnosed if blood pressure had been measured on at least three different occasions.

► Before considering implementation of this prediction model, external validation is needed, as well as careful evaluation of possible benefits and harms of targeted preventive strategies.

INTRODUCTION

In recent decades, the prevalence of childhood high blood pressure has increased to 4%–5%, largely driven by the growing prevalence of childhood overweight and obesity.[1] Childhood high blood pressure is usually asymptomatic until complications occur, but may have adverse consequences for later cardiovascular health,[1] partly through early adverse vascular changes and partly through tracking of childhood blood pressure levels into adulthood.[2] Childhood high blood pressure has been associated with atherosclerosis in autopsy studies,[1 3 4] and with unfavourable changes in markers of subclinical atherosclerosis in adulthood, such as increased carotid intima-media thickness and coronary artery calcification, independently of adulthood blood pressure.[5 6] Although blood pressure does not track as strongly from childhood into adulthood as other cardiovascular risk factors such as body fat and cholesterol,[7] longitudinal studies have demonstrated that childhood high blood pressure is moderately predictive of adulthood hypertension,[8 9] a well-known risk factor for cardiovascular disease (CVD).[10]

Several medical societies have recognised the importance of primordial prevention of CVD by preventing high blood pressure from early in the life course onwards,[11 12] for example, by improving nutrition, increasing physical activity and decreasing sedentary behaviour.[1 11 12] Such prevention efforts may be targeted to children at high risk of developing high blood pressure. This would require a tool (eg, based on a prediction model) that accurately identifies these children from the general population. In many developed countries, the general population of children is reached through preventive

child healthcare services, well-child care services or other forms of community-based child healthcare, where there might be a possibility for targeted prevention based on predictive risk assessments.

To apply a prediction model in community-based child healthcare, ideally it should be based on information commonly recorded or easily obtainable in such a setting. Studies on prediction models for future high blood pressure applicable in childhood are scarce,[8 9 13–15] and for most models information is needed that will likely be difficult to obtain in community-based child healthcare settings, such as repeated blood pressure measurements or genetic information.[8 9 14 15] Also, the potential of some of these models for early-life targeted prevention is limited, as they are applicable only later in childhood.[9 13 15] Importantly, generalisation of the performance of these prediction models is uncertain as none of them were validated.

Community-based child healthcare reaches many children from a very early age onwards and usually repeated consultations take place. For example, Preventive Child Health Care in the Netherlands reaches over 90% of all children, with frequent and free preventive health consultations from birth until the age of 18 years.[16] This not only provides the opportunity to start with prevention very early in life, but the follow-up would also allow for updating of predictions based on new information such as more recent measurements of the child's body mass index (BMI). This can be described as *dynamic prediction.*

In this study, as a part of the Prediction Of Child CardioMetabolic Risk project,[17] we aimed to develop a dynamic prediction model feasible to use in community-based child healthcare, from birth to the age of 6 years, to predict high blood pressure at the age of 9–10 years.

METHODS
Participants
Data from Generation R, a population-based prospective birth cohort study, were used. Full details of the study design have been published elsewhere.[18] Pregnant women with an expected delivery date between April 2002 and January 2006, living in Rotterdam, the Netherlands, at the time of birth, were eligible. Most women were enrolled in pregnancy, but enrolment was allowed until birth. During early life, growth data from Preventive Child Health Care centres in the area were retrieved by Generation R.

From the original cohort of 9749 live born children, 8548 children were invited for the follow-up visit at the age of 9–10 years; 5862 of these children attended the research centre, of which 5488 had complete blood pressure measurements. Children aged 8 or 11 years at this visit (n=129) were excluded. As we aimed to develop a prediction model for all children from the general population, no other exclusion criteria were applied. In total, data of 5359 children could be analysed (figure 1).

Outcome measure
Blood pressure was measured in the research centre by well-trained staff at a median age of 9.73 years (range 9.00–10.99 years). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured in supine position, at the right brachial artery, using the validated automatic sphygmomanometer Datascope Accutorr Plus.[19] An appropriately sized cuff was used for each child, with a bladder width of 40% of the arm circumference and a bladder length of more than 80% of the arm circumference. Blood pressure was measured four times with 1 min intervals. For this analysis, mean SBP and DBP values based on the last three measurements were used. Gender-specific, age-specific and height-specific blood pressure percentiles derived from German reference values were used to define the outcome high blood pressure.[20] These percentiles are based on the distribution of blood pressure in a non-overweight population of 12 199 children, that was considered representative of the German population, and also included children with a migrant background (17.1% had a two-sided migrant background, most commonly Turkish or Russian). Height percentiles were comparable to the Dutch population, and blood pressure was measured with the same automated device as in Generation R.[20] We chose to use reference values derived from a non-overweight population, as it is has been recognised that the increasing prevalence of overweight shifts the blood pressure distribution of a population upwards, while this is not a normal or healthy situation.[21] High blood pressure was defined as mean SBP and/or DBP at or above the 95th percentile.

Candidate predictors
Based on previous studies and expert consultations, variables were identified that have been associated with childhood blood pressure, and that are usually recorded or would otherwise be relatively easy to obtain (eg, through self-reports or extracted from medical reports) in community-based child healthcare settings. These are presented in table S1 of online supplementary material 1, with supporting literature. To prevent overfitting of the prediction model, a selection from these potential candidate predictors was made based on (1) expected predictive strength based on the literature, (2) correlations between variables (eg, between maternal and paternal educational level, and maternal smoking during and after pregnancy) and (3) feasibility in community-based child healthcare. This resulted in the following candidate predictors for analysis: maternal prepregnancy BMI, maternal hypertensive disease in pregnancy, maternal educational level, hypertension in biological parents, CVD in family of biological parents, parental smoking, child gender, child ethnicity, gestational age at birth, birth weight SD score (SDS) and repeatedly measured BMI SDS.

Maternal prepregnancy BMI was derived from a questionnaire during pregnancy. The presence of maternal hypertensive disease in pregnancy was based on delivery reports, cross-checked with hospital charts at admission
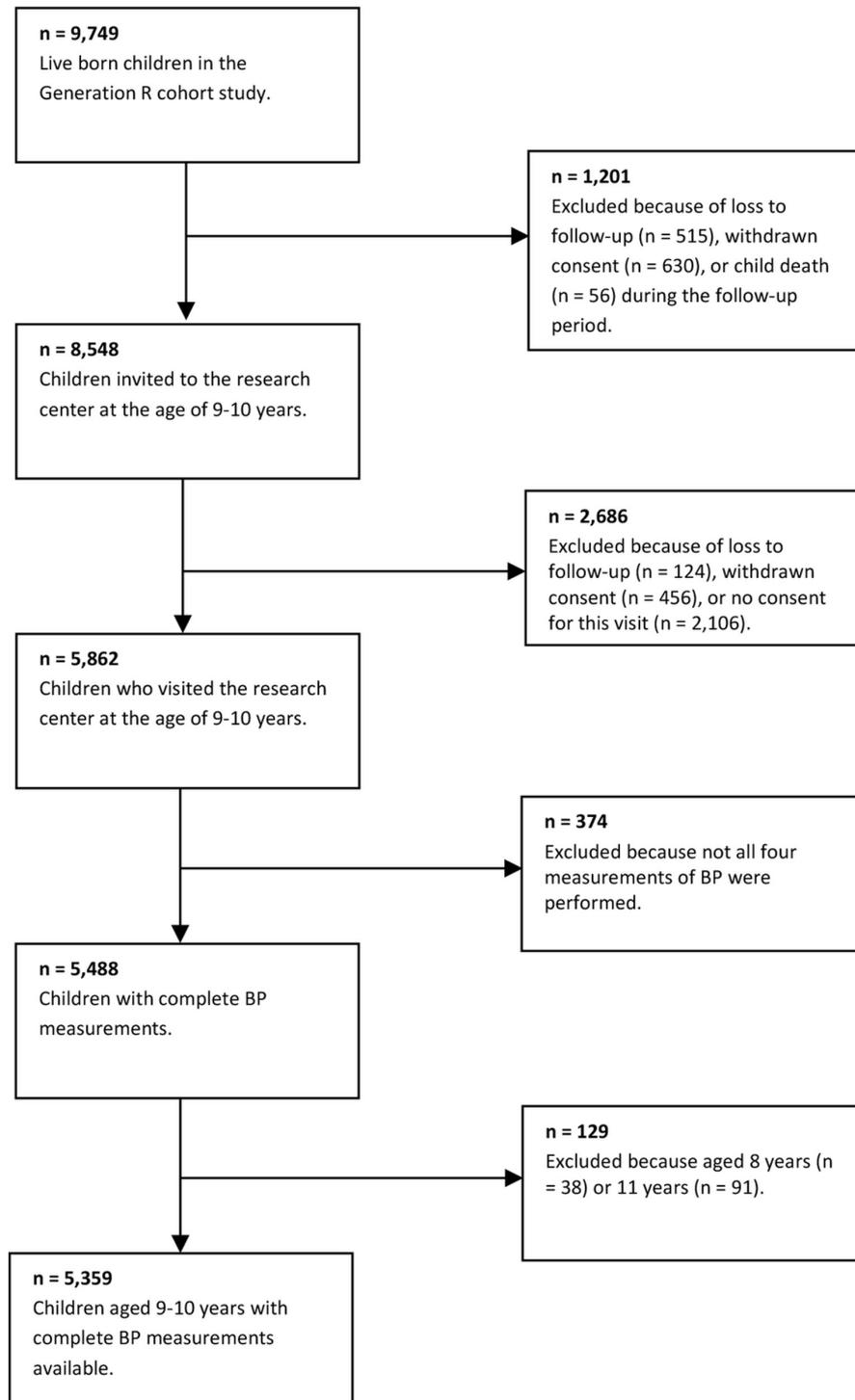
n = 9,749
Live born children in the Generation R cohort study.

n = 1,201
Excluded because of loss to follow-up (n = 515), withdrawn consent (n = 630), or child death (n = 56) during the follow-up period.

n = 8,548
Children invited to the research center at the age of 9-10 years.

n = 2,686
Excluded because of loss to follow-up (n = 124), withdrawn consent (n = 456), or no consent for this visit (n = 2,106).

n = 5,862
Children who visited the research center at the age of 9-10 years.

n = 374
Excluded because not all four measurements of BP were performed.

n = 5,488
Children with complete BP measurements.

n = 129
Excluded because aged 8 years (n = 38) or 11 years (n = 91).

n = 5,359
Children aged 9-10 years with complete BP measurements available.

Figure 1 Flow chart of study participants. BP, blood pressure.

and the Netherlands Perinatal Registry. Diseases that were considered as maternal hypertensive disease in pregnancy were pregnancy-induced hypertension, pre-eclampsia and haemolysis elevated liver enzymes and low platelets syndrome (HELLP syndrome). Pre-existing hypertension was not considered as maternal hypertensive disease in pregnancy, unless there was superimposed pre-eclampsia or HELLP syndrome. Maternal educational level (highest level finished) was based on a questionnaire at inclusion, and categorised as none or primary education; secondary

education; and higher education. Hypertension in the biological parents was self-reported in questionnaires at inclusion, and categorised as no parental hypertension or at least one parent with hypertension. Parents who reported not knowing if they had hypertension (3%–5% of received answers) were classified as not having hypertension. Family history of CVD was also self-reported at inclusion, and considered positive if at least one parent had at least one relative (mother, father, sister, brother or child) with hypertension, myocardial infarction before

Table 1 Description of candidate predictors for the population for analysis (n=5359)

Candidate predictors	Descriptive	Available data, n (%)
Parental baseline characteristics		
Maternal prepregnancy BMI (kg/m²)*	22.6 (14.4–50.2)	4038 (75.3)
Maternal hypertensive disease in pregnancy, n (%)	306 (6.4)	4745 (88.5)
Maternal education level		4937 (92.1)
No education/primary education, n (%)	360 (7.3)	
Secondary education, n (%)	2070 (41.9)	
Higher education, n (%)	2507 (50.8)	
Smoking by at least one parent, n (%)	883 (42.4)	2083 (38.9)
Hypertension in at least one biological parent, n (%)	104 (3.7)	2792 (52.1)
CVD in family of at least one biological parent, n (%)	1885 (68.0)	2772 (51.7)
Child characteristics		
Male gender, n (%)	2627 (49.0)	5359 (100.0)
Gestational age at birth in weeks*	40.1 (26.3–43.6)	5327 (99.4)
Birth weight SDS†	−0.07 (1.02)	5311 (99.1)
Child ethnicity		5240 (97.8)
Western, n (%)	3636 (69.4)	
Turkish, n (%)	328 (6.3)	
Moroccan, n (%)	258 (4.9)	
Surinamese, n (%)	366 (7.0)	
Other non-western, n (%)	652 (12.4)	
BMI SDS around 6 months†	0.54 (1.01)	3571 (66.6)
BMI SDS around 1 year†	0.26 (1.03)	3719 (69.4)
BMI SDS around 2 years†	0.11 (1.06)	3491 (65.1)
BMI SDS around 3 years†	0.13 (1.05)	3355 (62.6)
BMI SDS around 4 years†	0.17 (1.12)	2924 (54.6)
BMI SDS at 5–6 years†	0.51 (1.06)	4705 (87.8)
Age in months at BMI measurement around 6 months*	6.2 (5.0–10.0)	3571 (66.6)
Age in months at BMI measurement around 1 year*	11.1 (10.0–13.0)	3719 (69.4)
Age in months at BMI measurement around 2 years*	24.8 (23.0–29.0)	3491 (65.1)
Age in months at BMI measurement around 3 years*	36.7 (34.8–44.0)	3355 (62.6)
Age in months at BMI measurement around 4 years*	45.8 (44.0–52.2)	2924 (54.6)
Age in months at BMI measurement at 5–6 years*	72.1 (60.1–84.0)	4705 (87.8)

*Median (range).
†Mean (SD).
BMI, body mass index; CVD, cardiovascular disease; SDS, SD score.

the age of 65 or stroke. Parental smoking was assessed through questionnaires during pregnancy (asking the partner whether he smoked in the 2 months before pregnancy), and the first 6 months (asking the mother whether she smoked at that time point). Next, parental smoking was categorised as none of the parents smoke or at least one parent smokes. Child ethnicity was based on questionnaires and determined in accordance with Statistics Netherlands according to country of birth of the child's parents: if one parent was born outside the Netherlands, that country was used to determine the child's ethnicity, and if both parents were born outside the Netherlands,

the country of birth of the mother was used to determine the child's ethnicity.[22] Categories were created by considering (1) studies on the association between childhood high blood pressure and ethnicity in the Netherlands, and (2) large ethnic groups in the Netherlands.[22–24] Child ethnicity was categorised as Western (including Dutch, other European, western American, western Asian, Oceanian), Turkish, Moroccan, Surinamese and Other non-western (including Cape Verdean, Dutch Antilles, African, non-western Asian, Indonesian, non-western American). Gestational age at birth in weeks and birth weight were based on delivery reports. Birth weight SDS,

Table 2 Associations between candidate predictors and high blood pressure at the age of 9–10 years

Characteristics	Univariable OR (95% CI)	Multivariable OR from final model (95% CI)
Baseline predictors		
Female gender	0.85 (0.63 to 1.15)	–
Child ethnicity		
Western	*Reference category*	*Reference category*
Turkish	4.27 (2.76 to 6.60)	2.63 (1.59 to 4.34)
Moroccan	1.22 (0.58 to 2.59)	0.78 (0.34 to 1.80)
Surinamese	1.12 (0.56 to 2.21)	0.76 (0.37 to 1.56)
Other non-western	2.11 (1.41 to 3.17)	1.67 (1.07 to 2.59)
Gestational age at birth*, 41 weeks vs 39 weeks	0.86 (0.75 to 0.98)	–
Birth weight, +1 SDS vs −1 SDS	0.59 (0.44 to 0.80)	0.60 (0.43 to 0.83)
Maternal prepregnancy BMI*, 25 kg/m^2 vs 21 kg/m^2	1.37 (1.21 to 1.55)	1.16 (1.02 to 1.33)
Maternal hypertensive disease in pregnancy	2.64 (1.67 to 4.15)	2.43 (1.51 to 3.93)
Maternal educational level		
No/primary education	*Reference category*	*Reference category*
Secondary education	0.62 (0.38 to 1.03)	0.86 (0.50 to 1.49)
Higher education	0.25 (0.15 to 0.42)	0.44 (0.23 to 0.81)
Hypertension in at least one biological parent	4.27 (2.08 to 8.75)	–
CVD in first degree relatives of at least one biological parent	1.30 (0.79 to 2.14)	–
Smoking by at least one parent	1.48 (0.91 to 2.41)	–
Time-varying predictors		
BMI SDS +1 vs BMI SDS −1 at age 6 months	1.22 (0.96 to 1.54)	1.28 (1.02 to 1.61)
BMI SDS +1 vs BMI SDS −1 at age 1 year	1.32 (1.05 to 1.66)	1.35 (1.08 to 1.68)
BMI SDS +1 vs BMI SDS −1 at age 2 years	1.54 (1.23 to 1.93)	1.50 (1.20 to 1.86)
BMI SDS +1 vs BMI SDS −1 at age 3 years	1.80 (1.42 to 2.29)	1.66 (1.32 to 2.08)
BMI SDS +1 vs BMI SDS −1 at age 4 years	2.11 (1.63 to 2.73)	1.85 (1.45 to 2.36)
BMI SDS +1 vs BMI SDS −1 at age 5 years	2.47 (1.85 to 3.29)	2.05 (1.56 to 2.69)
BMI SDS +1 vs BMI SDS −1 at age 6 years	2.89 (2.09 to 3.99)	2.28 (1.68 to 3.08)

*OR is presented for a comparison between the outer values of the IQR.
BMI, body mass index; CVD, cardiovascular disease; SDS, SD score.

adjusted for gestational age, was determined according to Niklasson *et al*.[25] BMI was based on protocolled measurements of height and weight at either a Preventive Child Health Care centre (0–4 years) or the Generation R research centre (5–6 years, median 6.0 years). BMI SDS values, adjusted for age and gender, were constructed based on Dutch reference growth curves from 2010.[26]

Statistical analysis

The percentage of missing values for the candidate predictors ranged from 0% to 61%. To be able to use all of the observed information for each candidate predictor, missing values were imputed 30 times with multivariate imputation by chained equations[27]. For three candidate predictors based on complete information from both parents (parental smoking, parental hypertension and CVD in family), the proportions of missing values were high (61%, 48% and 48%, respectively). Therefore, we used the separate variables from each parent in the imputation model; for these variables the proportions of missing values ranged from 14%–45%. The final

Table 3 Discriminative ability of the final dynamic prediction model when applied at different ages between birth and 6 years

Age	Apparent AUC	For optimism corrected AUC
Around 6 months	0.71	0.69
Around 1 year	0.70	0.68
Around 2 years	0.69	0.67
Around 3 years	0.67	0.65
Around 4 years	0.67	0.66
At 5–6 years	0.74	0.73

AUC, area under the ROC curve.

imputation model included all candidate predictors, the outcome and the following auxiliary variables: maternal smoking, paternal smoking, maternal hypertension, paternal hypertension, maternal family history of CVD, paternal family history of CVD, child BMI values at other ages between birth and 6 years, smoking in the home environment, family income, paternal educational level, paternal BMI and maternal BMI at child age 5–6 years.

First, we studied whether the BMI SDS trajectory of a child would predict high blood pressure better than the most recent BMI SDS only. We applied a two-step model to investigate the use of BMI SDS trajectories. In the first step, each child's BMI SDS trajectory was modelled using a random effects model with restricted cubic splines, and in the second step the individual coefficients of each child's trajectory were used as a predictor in a logistic regression model with high blood pressure as the outcome. We saw that, in our study, the trajectory was not of added predictive value when the most recent BMI SDS and birth weight SDS were already included. Therefore, in the subsequent analysis we used only the most recent BMI SDS and not the BMI SDS trajectory.

Next, logistic regression analyses were performed with high blood pressure at age 9–10 years as the outcome, predicted at different ages (6 months, 1 year, 2 years, 3 years, 4 years and 5–6 years). For each age, a backward stepwise selection procedure was performed, using the Akaike Information Criterion for predictor selection. For a variable with one parameter this corresponds to selection at a p value of 0.157.[28] Restricted cubic splines were used to examine non-linearity in the association between birth weight SDS and high blood pressure, because not only low, but also high birth weight might be associated with high blood pressure.[29] An interaction term was considered for birth weight SDS and the most recent BMI SDS value, as it was shown that a combination of high current BMI with low birth weight was predictive of higher SBP.[30] Neither the splines nor the interaction term were of added predictive value in our study, and thus were not used in the final models.

As we aimed for a dynamic prediction model that could be applied at any age from birth until the age of 6 years, we first checked whether at each age the same baseline predictors remained in the model after backward selection. Furthermore, we checked whether the size of the baseline regression coefficients was similar at each age, and whether the coefficient for BMI SDS would increase with increasing age. As these conditions were satisfied, we developed a dynamic prediction model by including the selected baseline predictors and an interaction between BMI SDS and age. By doing so, the predictive value of BMI SDS was allowed to vary with the child's age at measurement, while associations of the other predictors were kept constant. This approach is referred to in the literature as *dynamic logistic regression* or *pooled logistic regression*,[31] and reduces the need for age-specific models. To take into account that each child contributes to the model with multiple measurements of BMI SDS (ranging from 0 to 14 measurements), robust standard errors were calculated by fitting the model using generalised estimating equations (GEE) with the *independence* correlation structure. GEE is usually used to deal with repeatedly measured outcomes, but can also be used to adjust standard errors for repeatedly measured predictors or exposures, such as BMI SDS and age in our study.[31 32] For this purpose, an independent working correlation matrix must be specified.[32] From the GEE analysis, we obtained the final model with coefficients, and the estimates for the apparent (ie, determined before internal validation) discrimination at different ages.

As currently there is no R package available to internally validate prediction models based on GEE, we performed internal validation procedures on the logistic regression models at each age using bootstrapping. 250 random sets of data were generated, with the same size as the original dataset, drawn with replacement from the original data. These datasets were used to estimate, for each age, the optimism in the quality of the prediction model. Compared with other internal validation techniques such as cross-validation, bootstrapping is better able to capture model uncertainty caused by variable selection methods such as backward selection.[33 34]

As performance measures we assessed, at different ages, the apparent area under the receiver operating characteristic curve (AUC) and the AUC adjusted for the optimism calculated with the bootstrapping procedure. The AUC is a measure of discrimination of a prediction model, and reflects the ability of the model to correctly assign a higher risk to individuals who have the outcome compared with those who do not.[34] Next, we assessed the calibration slopes calculated in the bootstrap procedure which represent, at different ages, the ability of the model to estimate the level of risk accurately. It can range from 0 to 1, where 1 means that the model is perfectly calibrated.[34] The mean of these calibration slopes was used to shrink the regression coefficients estimated with the GEE approach. As a final step, the intercept was adjusted to re-establish the calibration-in-the-large, so that the mean of the predicted risks was again in line with the mean of the observed risks.[34] The resulting final model was applied in an Excel risk calculator. All analyses were performed in R, version 3.3.2.

Patient involvement

Outcome definition and selection of candidate predictors were presented and discussed in a meeting with a group of stakeholders involved with our research project, including child healthcare professionals and a parent representative from a Dutch parent organisation.

RESULTS
Model development

Table 1 shows the characteristics of the 5359 included children, of which 150 (2.8%) were twins and 323 (6.1%)

Prediction model for childhood high blood pressure

Predictors		Values
Maternal pre-pregnancy BMI	in kg/m²	26
Maternal educational level		
No/primary education	yes = 1 / no = 0	1
Secondary education	yes = 1 / no = 0	0
Higher education	yes = 1 / no = 0	0
Maternal hypertensive disease in pregnancy	yes = 1 / no = 0	0
Birth weight SDS	SDS	-0.5
Child ethnicity		
Western	yes = 1 / no = 0	0
Turkish	yes = 1 / no = 0	1
Moroccan	yes = 1 / no = 0	0
Surinamese	yes = 1 / no = 0	0
Other non-western	yes = 1 / no = 0	0
Age in months	in months	60
BMI SDS	SDS	2.0

Predicted probability of high blood pressure — 18%

Figure 2 Screenshot of spreadsheet with risk calculations based on the dynamic prediction model. The prediction is calculated using the regression equation of the final prediction model, after shrinkage of the coefficients and updating of the intercept. *Linear predictor (LP) for high BP at age 9–10 years*=−3.805+0.818*maternal hypertensive disease in pregnancy−−0.135*maternal educational level: secondary education−0.763*maternal educational level: higher education+0.035*maternal prepregnancy BMI+0.888*Turkish ethnicity−0.226*Moroccan ethnicity−0.256*Surinamese ethnicity+0.470*other non-western ethnicity−0.237*birth weight SDS−0.005*age in months at BMI measurement+0.089*BMI SDS+0.004*age in months at BMI measurement*BMI SDS. *Probability of high BP at age 9–10 years*=1/(1 + e^{-LP}). The Excel risk calculator is available on request to the corresponding author. BMI, body mass index; BP, blood pressure; SDS, SD score.

were born preterm. In total, 227 children (4.2%) had high blood pressure at the age of 9–10 years. Predictors selected in the logistic regression models were: maternal hypertensive disease in pregnancy, maternal educational level, maternal prepregnancy BMI, child ethnicity, birth weight SDS and child BMI SDS at the specific age. These models are presented in table S2 of online supplementary material. The predictive value of the child's BMI SDS increased with increasing age. Table 2 shows the ORs for the univariable and multivariable associations between the candidate predictors and high blood pressure based on GEE. The strongest predictors were child age in combination with BMI SDS, maternal educational level and maternal prepregnancy BMI. Apparent AUCs

for the model ranged from 0.67 to 0.74 at different ages (table 3).

Internal validation and final model
After correcting for optimism, AUCs at different ages ranged from 0.65 to 0.73 (table 3). Calibration slopes ranged from 0.91 to 0.94; a mean shrinkage factor of 0.92 was applied to the coefficients of the GEE model. The intercept had to be adjusted with −0.23. Figure 2 shows the Excel risk calculator. The predicted probabilities in our dataset ranged from 0.2% to 53.4% (median 3.0%). Within children, predicted probabilities can vary at different ages given their BMI SDS development (figure 3 and table 4). Table 5 shows the prevalence of high blood
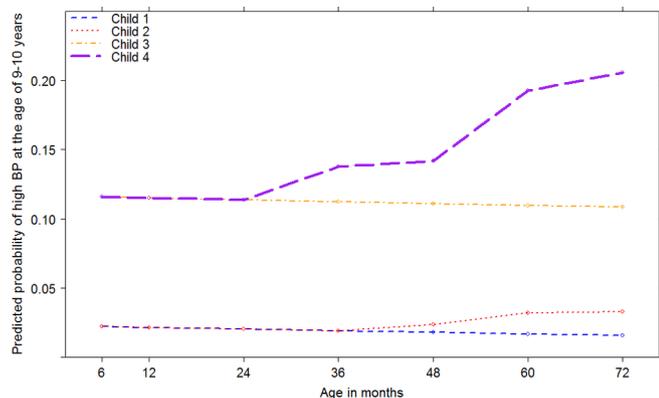
Figure 3 Predicted probabilities of high blood pressure at the age of 9–10 years, predicted from the age of 6 months to 6 years, comparing four children with different baseline characteristics and different BMI SDS development (as shown in table 4). BMI, body mass index; BP, blood pressure; SDS, SD score.

pressure across four categories of predicted risk at the ages 1, 3 and 5–6 years. For example, in children with a predicted risk of more than 15% at the age of 5–6 years, the measured prevalence of high blood pressure at the age of 9–10 years was 24.8%.

DISCUSSION
We developed a dynamic model to predict, for children from birth until the age of 6 years in the general population, their risk of high blood pressure at the age of 9–10 years, based on information that is relatively easy to obtain. The dynamic nature of the prediction model allows for incorporating new information on BMI SDS that becomes available as a child gets older, so that the predicted risk can be updated.

After internal validation, the discriminative ability of the prediction model was moderate, and highest at the age of 5–6 years (AUC 0.73) which can be explained by the higher predictive value of BMI SDS at an age closer to the age at outcome assessment. Although the overall discriminative performance was not excellent or good, the prediction model did allow for identification of a

group of children at a considerably higher risk than the overall study population to have high blood pressure at the age of 9–10 years. The prediction model might therefore prove helpful to community-based child healthcare professionals, because it would allow them to objectively select children for targeted prevention. On the other hand, before considering implementation of this prediction model, first external validation studies are needed, in order to study the generalisability of the prediction model and to see what adaptations to specific populations might be necessary to improve the performance of the model.

In line with previous studies, we found the following predictors for childhood high blood pressure: maternal hypertensive disease in pregnancy,[35 36] maternal educational level,[37 38] maternal prepregnancy BMI,[39] child ethnicity,[23 24] birth weight,[29 30] and BMI SDS.[9 13–15] Candidate predictors that did not improve the model were gestational age at birth, child gender, hypertension in the biological parents, CVD in the family of the biological parents and parental smoking, even though in previous studies these were associated with blood pressure in the child.[14 40–43] One reason these predictors proved unimportant in our study could be that they were correlated with other predictors we included, for example, parental hypertension with maternal BMI. For parental hypertension, another explanation might be that parents were still relatively young and therefore the prevalence was low. Further, the high proportions of missing values for parental smoking, parental hypertension and family history of CVD could have decreased the power to detect associations between these candidate predictors and childhood high blood pressure. A possible reason for the association between gender and high blood pressure reported in a previous study is that in that study the outcome was assessed in adolescence,[41] where effects of puberty or other adolescence-related variables may play a role in differentiating the risk of high blood pressure.

As blood pressure was measured in Generation R participants at the age of 5–6 years, we have considered adding SBP and/or DBP at this age to the model to improve its discriminative ability, but decided not to, because in most countries routine measurement of blood pressure

Table 4 Characteristics of the four children with different baseline characteristics and different BMI SDS development

Child	Maternal prepregnancy BMI (kg/m²)	Maternal educational level	Hypertensive disease in pregnancy	Birth weight SDS	Child ethnicity	BMI SDS age 6 months	BMI SDS age 1 year	BMI SDS age 2 years	BMI SDS age 3 years	BMI SDS age 4 years	BMI SDS age 5 years	BMI SDS age 6 years
1	22.6	Higher education	No	−0.07	Western	0	0	0	0	0	0	0
2	22.6	Higher education	No	−0.07	Western	0	0	0	0	+1	+2	+2
3	22.6	No/primary education	No	−0.07	Turkish	+1	+1	+1	+1	+1	+1	+1
4	22.6	No/primary education	No	−0.07	Turkish	+1	+1	+1	+2	+2	+3	+3

BMI, body mass index; SDS, SD score.

Table 5 Prevalence of high blood pressure at the age of 9–10 years across four risk categories at the age of 1, 3 and 5–6 years

Risk category	Age 1 year n=3719*		Age 3 years n=3355*		Age 5–6 years n=4705*	
	Number of children, n (%)†	High BP, n (%)‡	Number of children, n (%)†	High BP, n (%)‡	Number of children, n (%)†	High BP, n (%)‡
0%–4.9%	2845 (76.5)	73 (2.6)	2723 (81.2)	79 (2.9)	3624 (77.0)	76 (2.1)
5%–9.9%	682 (18.3)	46 (6.7)	484 (14.4)	35 (7.2)	752 (16.0)	51 (6.8)
10%–14.9%	151 (4.1)	20 (13.2)	112 (3.3)	7 (6.3)	188 (4.0)	25 (13.3)
15% or more	41 (1.1)	9 (22.0)	36 (1.1)	12 (33.3)	141 (3.0)	35 (24.8)

*Children with observed BMI SDS values around that specific age.
†Proportion of all children.
‡Proportion of children with high BP within risk category.
BMI, body mass index; BP, blood pressure; SDS, SD score.

has not been incorporated into community-based child healthcare.[44 45] Even though American and European medical societies recommend routine measurement of blood pressure for children from the age of 3 years,[21 46] the debate on its usefulness is still ongoing.[46 47] If blood pressure measurement would become a standard procedure in community-based child healthcare, updating the prediction model with information on current blood pressure should be considered.

This study has several strengths and limitations. First of all, we used data from a large prospective cohort study, allowing us to consider many possible candidate predictors. Based on literature and expert opinion, we believe that we included all the most relevant candidate predictors, while also taking into account that the model should be applicable in most community-based child healthcare settings. Second, we could study the change in the predictive value of childhood BMI SDS, and by using pooled logistic regression analysis we were able to incorporate this information into one dynamic prediction model that can be applied at different ages, increasing the ease of its use in practice.

A possible weakness of the cohort is the loss to follow-up: about 40% of the children in the original cohort did not attend the visit at the age of 9–10 years. In general, children remaining in the study more often had a Western ethnic background, and their mothers were older and better educated.[18] Since all ethnicities and educational levels were still well-represented in the follow-up, we think it unlikely that this loss to follow-up appreciably biased the associations found. An important limitation of our study is that for some candidate predictors based on information from both parents (parental smoking, parental hypertension and CVD in the family), the proportions of missing values were very high. Even though we included all available information from each individual parent in the imputation model, the missing values could have reduced the power to detect these variables as predictors for high blood pressure in our model. Therefore, we cannot exclude that information about these variables in reality might be of added predictive value (and hence

increase the performance of the model). This should be investigated in external validation studies with more complete data on these variables. Another point that should be noted is that we performed the internal validation of the GEE model indirectly, that is, through bootstrapping the logistic regression models at each age, as there is not yet a software package available that is able to perform this directly on the GEE model. The results were stable over the different ages and standard errors are correct in the analyses for one time point. Therefore, the estimated optimism may be considered as realistic, although external validation is again recommended.

It should be noted that we could only measure the outcome as *high blood pressure* and not as *hypertension*. To diagnose childhood hypertension, blood pressure needs to be measured on three separate occasions,[21] which was not the case in this study. Most of the children with high blood pressure in our cohort would not be diagnosed with hypertension if they had been followed up on two more occasions. Based on previous studies, we estimate that about 1 in 4 to 5 children with high blood pressure in our study would be diagnosed as hypertensive.[41 48] On the other hand, the high blood pressure prevalence in our study might be slightly underestimated, because blood pressure was measured in a supine position which tends to give lower blood pressure values than measurement in a sitting position, as in the study for the reference values.[20] This might have filtered out some of the children without true hypertension. Setting aside these limitations, several studies have shown that high blood pressure in childhood measured on only one occasion is associated with an increased risk of hypertension in later life.[8 9] Therefore, extra attention to these children could still be warranted, although we must be aware that this has not yet been studied for the more recent reference values for high blood pressure based on non-overweight populations.

If external validity can be confirmed, we would propose that, based on this prediction model, only minimally intensive (and not invasive or harmful) strategies should be offered to high-risk children,

considering that (1) the discriminative performance is only moderate, (2) it concerns prediction of *high blood pressure* and not *hypertension* and (3) it has not yet been studied whether targeted interventions in this population would be effective. As mentioned previously, the prediction model might be helpful to guide community-based child healthcare professionals in better distributing their time and efforts, by identifying children that need relatively more attention to prevention of CVD, for example in the form of tailored lifestyle and nutritional advice,[11 12] measurement of the child's current blood pressure, and monitoring of blood pressure during follow-up. In overweight or obese children, a higher predicted risk could help to underline the importance of improving weight status. Depending on the strategies to be offered, higher or lower cut-offs to define the high risk group might be used, and the use of multiple cut-offs and differentiated strategies might also be considered. In very young children (eg, <4 years of age) with a high predicted risk, it might be a strategy to wait for the result of the next risk assessment before starting with targeted prevention. Before implementation, the possible benefits and harms of the preferred strategies should be discussed. It would also be important to investigate how parents and health professionals could experience the use of such a prediction model, including the acceptability and effectiveness of risk communication. Lastly, if the model would be implemented in the future, the effects of applying the model in combination with targeted prevention on the occurrence of high blood pressure should be investigated in a randomised or cluster randomised trial.

In summary, we developed a dynamic prediction model to predict the development of childhood high blood pressure based on information that is usually recorded or is easy to obtain in community-based child healthcare practice. This can be seen as a first step towards applying childhood prediction models for future high blood pressure in order to offer targeted primordial prevention of CVD.

REFERENCES

1. Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and AdolescentsNational Heart, Lung, and Blood Institute. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics* 2011;128:S213–56.
2. Chen X, Wang Y. Tracking of blood pressure from childhood to adulthood: a systematic review and meta-regression analysis. *Circulation* 2008;117:3171–80.
3. Berenson GS, Srinivasan SR, Bao W, et al. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. The bogalusa heart study. *N Engl J Med* 1998;338:1650–6.
4. Strong JP, Malcom GT, McMahan CA, et al. Prevalence and extent of atherosclerosis in adolescents and young adults: implications for prevention from the pathobiological determinants of atherosclerosis in youth study. *JAMA* 1999;281:727–35.
5. Hartiala O, Magnussen CG, Kajander S, et al. Adolescence risk factors are predictive of coronary artery calcification at middle age: the cardiovascular risk in young Finns study. *J Am Coll Cardiol* 2012;60:1364–70.
6. Raitakari OT, Juonala M, Kähönen M, et al. Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the cardiovascular risk in young finns study. *JAMA* 2003;290:2277–83.
7. Twisk JW, Kemper HC, van Mechelen W, et al. Tracking of risk factors for coronary heart disease over a 14-year period: a comparison between lifestyle and biologic risk factors with data from the amsterdam growth and health study. *Am J Epidemiol* 1997;145:888–98.
8. Juhola J, Magnussen CG, Viikari JS, et al. Tracking of serum lipid levels, blood pressure, and body mass index from childhood to adulthood: the cardiovascular risk in young finns study. *J Pediatr* 2011;159:584–90.
9. Ferrer M, Fernández-Britto JE, Bacallao J, et al. Development of hypertension in a cohort of Cuban adolescents. *MEDICC Rev* 2015;17:41–7.
10. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388:1659–724.
11. Weintraub WS, Daniels SR, Burke LE, et al. Value of primordial and primary prevention for cardiovascular disease: a policy statement from the American Heart Association. *Circulation* 2011;124:967–90.
12. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: the sixth joint task force of the european society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of 10 societies and by invited experts) developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2016;37:2315–81.
13. Li L, Pinot de Moira A, Power C. Predicting cardiovascular disease risk factors in midadulthood from childhood body mass index: utility of different cutoffs for childhood body mass index. *Am J Clin Nutr* 2011;93:1204–11.
14. Juhola J, Oikonen M, Magnussen CG, et al. Childhood physical, environmental, and genetic predictors of adult hypertension: the cardiovascular risk in young Finns study. *Circulation* 2012;126:402–9.

15. Vos LE, Oren A, Bots ML, *et al*. Does a routinely measured blood pressure in young adolescence accurately predict hypertension and total cardiovascular risk in young adulthood? *J Hypertens* 2003;21:2027–34.
16. Siderius EJ, Carmiggelt B, Rijn CS, *et al*. Preventive child health care within the framework of the dutch health care system. *J Pediatr* 2016;177S:S138–S141.
17. de Kroon ML, Wijga A, Vergouwe Y, *et al*. Prediction of preadolescent overweight and poor cardiometabolic outcome in children up to 6 years of age: research protocol. *JMIR Res Protoc* 2016;5:e85.
18. Kooijman MN, Kruithof CJ, van Duijn CM, *et al*. The Generation R Study: design and cohort update 2017. *Eur J Epidemiol* 2016;31:1243–64.
19. Wong SN, Tz Sung RY, Leung LC. Validation of three oscillometric blood pressure devices against auscultatory mercury sphygmomanometer in children. *Blood Press Monit* 2006;11:281–91.
20. Neuhauser HK, Thamm M, Ellert U, *et al*. Blood pressure percentiles by age and height from nonoverweight children and adolescents in Germany. *Pediatrics* 2011;127:e978–e988.
21. Flynn JT, Kaelber DC, Baker-Smith CM, *et al*. Clinical practice guideline for screening and management of high blood pressure in children and adolescents. *Pediatrics* 2017;140:e20171904.
22. Statistics Netherlands. *Jaarrapport Integratie 2014 [Annual Report Integration 2014]*. Statistics Netherlands: Den Haag/Heerlen, 2014.
23. de Hoog ML, van Eijsden M, Stronks K, *et al*. Ethnic differences in cardiometabolic risk profile at age 5-6 years: the ABCD study. *PLoS One* 2012;7:e43667.
24. Gishti O, Gaillard R, Felix JF, *et al*. Early origins of ethnic disparities in cardiovascular risk factors. *Prev Med* 2015;76:84–91.
25. Niklasson A, Ericson A, Fryer JG, *et al*. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). *Acta Paediatr Scand* 1991;80:756–62.
26. Schönbeck Y, Talma H, van Dommelen P, *et al*. Increase in prevalence of overweight in Dutch children and adolescents: a comparison of nationwide growth studies in 1980, 1997 and 2009. *PLoS One* 2011;6:e27608.
27. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377–99.
28. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. *Journal of the Royal Statistical Society: Series C* 1999;48:313–29.
29. Zhang Y, Li H, Liu SJ, *et al*. The associations of high birth weight with blood pressure and hypertension in later life: a systematic review and meta-analysis. *Hypertens Res* 2013;36:725–35.
30. Uiterwaal CS, Anthony S, Launer LJ, *et al*. Birth weight, growth, and blood pressure: an annual follow-up study of children aged 5 through 21 years. *Hypertension* 1997;30:267–71.
31. Hansen BE, Buster EH, Steyerberg EW, *et al*. Prediction of the response to peg-interferon-alfa in patients with HBeAg positive chronic hepatitis B using decline of HBV DNA during treatment. *J Med Virol* 2010;82:1135–42.
32. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11:561–70.
33. Steyerberg EW. *Clinical prediction models*. New York: Springer-Verlag, 2009.
34. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925–31.
35. Fraser A, Nelson SM, Macdonald-Wallis C, *et al*. Hypertensive disorders of pregnancy and cardiometabolic health in adolescent offspring. *Hypertension* 2013;62:614–20.
36. Alsnes IV, Vatten LJ, Fraser A, *et al*. Hypertension in pregnancy and offspring cardiovascular risk in young adulthood: prospective and sibling studies in the HUNT study (Nord-Trøndelag Health Study) in Norway. *Hypertension* 2017;69:591–8.
37. Shi Y, de Groh M, Morrison H. Increasing blood pressure and its associated factors in Canadian children and adolescents from the canadian health measures survey. *BMC Public Health* 2012;12:388.
38. Duncan MJ, Vale S, Santos MP, *et al*. The association between cardiovascular disease risk and parental educational level in portuguese children. *Int J Environ Res Public Health* 2012;9:4311–20.
39. Lawlor DA, Najman JM, Sterne J, *et al*. Associations of parental, birth, and early life characteristics with systolic blood pressure at 5 years of age: findings from the Mater-University study of pregnancy and its outcomes. *Circulation* 2004;110:2417–23.
40. Bao W, Srinivasan SR, Wattigney WA, *et al*. The relation of parental cardiovascular disease to risk factors in children and young adults. The bogalusa heart study. *Circulation* 1995;91:365–71.
41. Sorof JM, Lai D, Turner J, *et al*. Overweight, ethnicity, and the prevalence of hypertension in school-aged children. *Pediatrics* 2004;113:475–82.
42. Steen E, Bonamy AK, Norman M, *et al*. Preterm birth may be a larger risk factor for increased blood pressure than intrauterine growth restriction. *Acta Paediatr* 2015;104:1098–103.
43. Simonetti GD, Schwertz R, Klett M, *et al*. Determinants of blood pressure in preschool children: the role of parental smoking. *Circulation* 2011;123:292–8.
44. Baltag V, Pachyna A, Hall J. Global overview of school health services: data from 102 countries. *Health Behav Policy Rev* 2015;2:268–83.
45. Wolfe I, McKee M. *European child health services and systems*. Lessons without borders: Open University Press, 2013.
46. Lurbe E, Agabiti-Rosei E, Cruickshank JK, *et al*. 2016 European society of hypertension guidelines for the management of high blood pressure in children and adolescents. *J Hypertens* 2016;34:1887–920.
47. Chiolero A, Bovet P, Paradis G. Screening for elevated blood pressure in children and adolescents: a critical appraisal. *JAMA Pediatr* 2013;167:266–73.
48. McNiece KL, Poffenbarger TS, Turner JL, *et al*. Prevalence of hypertension and pre-hypertension among adolescents. *J Pediatr* 2007;150:640–4.