The Extent to Which Genotype Information May Add to the Prediction of Disturbed Perinatal Adaptation: None, Minor, or Major?

ANDRÁS TRESZL, AMBRUS KAPOSI, JÚLIA HAJDÚ, MIKLÓS SZABÓ, TIVADAR TULASSAY, AND BARNA VÁSÁRHELYI

Research Laboratory of Pediatrics and Nephrology [A.T., A.K., M.S., T.T., B.V.], Hungarian Academy of Sciences, Budapest, Hungary 1083; First Department of Obstetrics and Gynecology [J.H.], Semmelweis University Budapest, Hungary 1085

ABSTRACT: Studies have been performed to describe the significance of genetic polymorphisms in complications associated with disturbed perinatal adaptation. Due to the large number of interacting factors, the results of classic statistical methods are often inconsistent. The random forest technique (RFT) is a robust nonparametric statistical approach that overcomes this problem through the calculation of the importance of each factor. We used RFT to reanalyze the importance of 24 genetic polymorphisms in the classification of preterm infants (birth weight, 680–1460 g, $n = 100$) to affected and unaffected groups according to the presence of acute perinatal complications. The accuracy of classification was between 0.5 and 0.8 for each complication when only birth data were considered. However, when genetic polymorphisms with the highest importance scores (ISs) were included in the analysis, the accuracy of classification according overall morbidity, necrotizing enterocolitis (NEC), acute renal failure (ARF), infant respiratory distress syndrome (IRDS), cardiac failure (CF), and patent ductus arteriosus (PDA) improved from 0.69, 0.60, 0.70, 0.72, 0.68, and 0.57 to 0.77, 0.70, 0.76, 0.77, 0.76, and 0.64, respectively. Our findings suggest that genetic polymorphisms identified by RFT as predictors may improve the risk assessment of preterm infants. RFT is a suitable tool to develop risk factor patterns in this population. (*Pediatr Res* **62:** 610–614, 2007)

Disturbed perinatal adaptation may affect almost all the major organ systems and may lead to a number of perinatal complications such as infant respiratory distress syndrome (IRDS), patent ductus arteriosus (PDA), intraventricular hemorrhage (IVH), acute renal failure (ARF), necrotizing enterocolitis (NEC), and cardiac failure (CF), respectively. Underdeveloped organs, impaired vasoregulation, infection, and inflammation contribute to higher susceptibility. Clinical status of the infants (particularly birth weight and gestational age), therapeutic interventions, and drugs are major determinants. However, not every infant from the same gestational age cohort and under the same conditions develops perinatal complications, and this fact stresses the need for identification of additional factors that have a possible impact on risk. Genetic polymorphisms are possible candidates (for more data, see refs. 1–6). Accumulating literature suggests that the carrier state of specific genetic variants may protect against, whereas others may increase, the risk of some perinatal complications. On the one hand, these findings may help to identify specific elements in disease pathogenesis and therapeutic targets in the future. Moreover, they may be of practical importance as they could help to identify patients at risk. However, what additional information may genotypes provide for the treating physician at the bedside? Can they help identify infants at risk? Available reports do not address these questions, probably as the risk of perinatal complications is influenced by so many interacting factors that classic statistical methods (*e.g.* logistic regression) become unstable and cannot handle simultaneously many genetic variants and other risk factors without requiring prohibitively large sample sizes. Recognizing this drawback, newly developed statistical approaches such as high-dimensional nonparametric methods seem more suitable for the analysis of the impact of genotypes on phenotype. One of these techniques is the random forest technique (RFT) (7). Quite recently, RFT was used to detect interactions between genetic polymorphisms and asthma (8).

Presumably, if genetic polymorphisms do contribute to the later development of perinatal complications, knowledge of a newborn infant's genotype at birth should increase the accuracy of prediction of perinatal complications compared with that obtained solely using clinical data. We tested this hypothesis with RFT on our database containing basic clinical data at birth and genotype data of preterm infants born with a birth weight ≤1500 g.

PATIENTS AND METHODS

In recent years, we have established a database of preterm neonates with a birth weight of ≤1500 g to investigate the impact of genotypes on disturbed perinatal adaptation. We analyzed the records of those 135 preterm infants who had been genotyped for at least 24 genetic variants (for a brief overview, see Table 1). The records also contain the following parameters: gender (65 girls, 70 boys), birth weight [median (range)]: 1280 g (700–1500) and gestational age 30 wk (24–32), presence of CF ($n = 44$), IRDS ($n = 67$), PDA ($n = 52$), IVH ($n = 44$), NEC ($n = 53$), sepsis ($n = 37$), and ARF ($n = 42$). The diagnosis of these morbidities was set up according to internationally accepted criteria and is described in detail in the references in Table 1. Our findings published earlier along with representative allele prevalence in this perinatal population are presented in Table 1. In this work, infants'

Abbreviations: ARF, Acute renal failure; **CF,** cardiac failure; **IRDS,** infant respiratory distress syndrome **IS,** importance score; **IVH,** intraventricular bleeding; **NEC,** necrotizing enterocolitis; **PDA,** patent ductus arteriosus; **RFT,** random forest technique

Table 1. *Published associations between genotype and disease based on the database used in our analysis*

Gene	Type/site of polymorphism	Putative effect	Prevalence of mutant allele in preterms	Impact of carrier state on neonatal morbidity	Ref.
Angiopoietin	$G^{-35}C$	n.d.	0.07	No impact	Curr Eye Res 2006;31:685
ACE	Insertion/deletion	Increased ACE activity	0.64	Ahock ↓	Acta Paediatr 2006;95:747; Pediatr Nephrol 2001;16:1063
AT1R	$A^{1166}C$	Angiotensin effect ↑	0.25	PDA ↓	Pediatr Res 2003;54:753
E-selectin	$Ser^{128}Arg$	Increased leukocyte-endothelial interaction	0.25	No impact	Acta Paediatr 2006;95:1213
Estrogen receptor α	PvuII restriction site (Pp)	Altered sensitivity to estrogen	0.45	Pp: PDA ↓, NEC ↓; Pp: IVH ↑	J Perinat Med 2005;33:461
HSP 70-2	$A^{1267}G$	HSP-70 ↓	0.40	ARF ↑, IRDS ↑	Pediatr Res 2003;54:452
HSP 70-1	$G^{190}C$	n.d.	0.15	No impact	Pediatr Res 2003;54:452
IL-1β	$C^{3954}T$	IL-1β ↑	0.24	Septic DIC (with IL-10^{-1082}A)	Biol Neonate 2003;83:241
IL-4 receptorα	$A^{1902}G$	IL-4 effect ↑	0.18	NEC ↓	J Pediatr Surg 2003;38:1374
IL-6	$G^{-174}C$	IL-6 ↓	0.31	ARF ↑ (with TNF-α^{-308}A)	Pediatr Nephrol 2002;17:713
IGF type 1 receptor	$G^{+3174}A$	Altered IGF effect	0.49	No impact	Pediatrics 2006;117:591
IFN-γ;	$T^{+874}A$	IFN-γ; ↑	0.46	PDA ↑, shock ↑, IRDS ↑	Arch Dis Child Fetal Neonat Ed 2007;92:F25
IL-10	$G^{-1082}A$	IL-10 ↓	0.51	No impact; potential role in DIC with IL-1β	Pediatr Nephrol 2002;17:713; Biol Neonate 2003;83:241
IL-12	p40 CTCTAA/GC	IL-12 ↓ in heterozygosity	0.42	NEC ↑, pneumonia ↑	Arch Dis Child Fetal Neonat Ed 2007;92:F25
IL-18	$G^{-137}C$, $C^{-607}A$	Altered IL-18 levels	0.28, 0.28	Severe NEC ↑	Eur J Pediatr 2002;161:410
L-selectin	$Pro^{213}Ser$	Altered leukocyte-endothelial interaction	0.20	Chronic lung disease ↑	Acta Paediatr 2006;95:1213
P-selectin	$Thr^{715}Pro$	Soluble P-selectin ↓	0.09	No impact	Acta Paediatr. 2006;95:1213
TLR-4	$Thr^{399}Ile$, $Asp^{299}Gly$	Altered response to bacteria	0.05, 0.03	No impact	J Pediatr Gastroenterol Nutr 2006;42:27
TNF-α	$G^{-308}A$, $G^{-238}A$	TNF-α ↑, n.d.	0.12, 0.04	Need for ventilatory support ↑	Eur Cytokine Netw 2005;16:78; Acta Paediatr 2001;90:1182
VEGF	$C^{-2578}A$, $G^{+405}C$, $T^{-460}C$	VEGF altered (mainly ↑)	0.49, 0.30, 0.49	$^{-2578}$A: NEC ↑, ARF ↓; $^{+405}$G, $^{-460}$C: ROP ↑, vitium ↑	Am Heart J 2006;151:878; Pediatr Res 2005;57:396

ACE, angiotensin-converting enzyme; AT1R, angiotensin II type 1 receptor; HSP, heat-shock protein; IFN-γ;, interferon-gamma; IGF, insulinlike growth factor; IL, interleukin; n.d., no data are available; TLR, Toll-like receptor; TNF-α, tumor necrosis factor αVEGF, vascular endothelial growth factor.
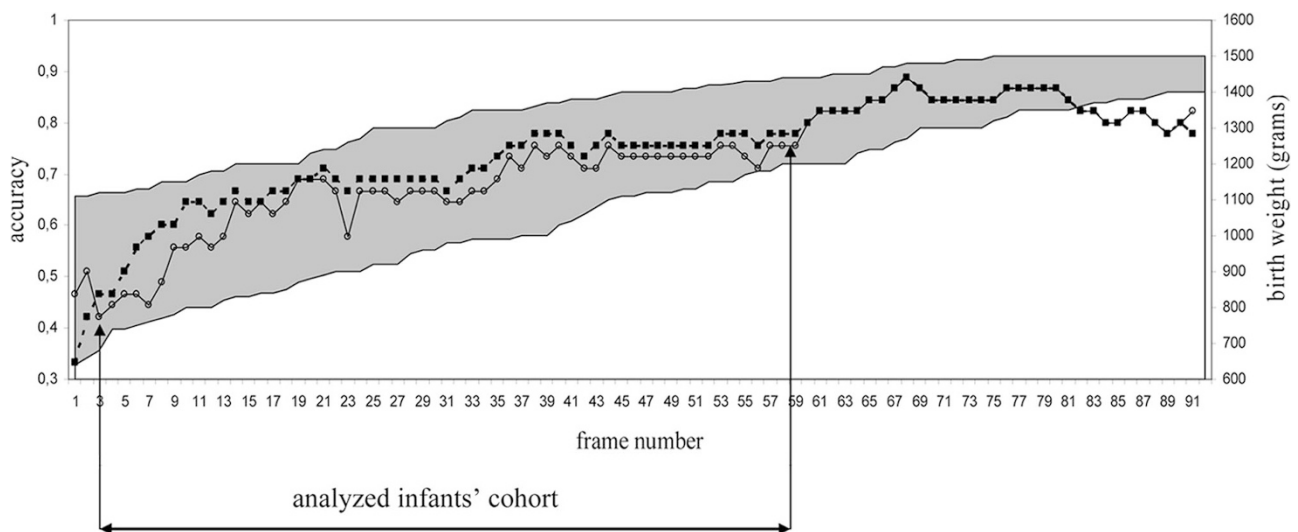

Figure 1. Selection process of that birth weight range where genotypes might add information to patient classification. ○, accuracy values based on birth data; ■, accuracy values based on birth data plus genotype data.

records were analyzed anonymously, and we were blinded to their personal data. Our study was approved by an institutional review board (TUKEB 16/2003).

Statistical analysis. Classification tree–based methods can be very efficient for selecting from large numbers of predictor variables those that best explain a phenotype. The ease of interpretation of classification trees, along

with their flexibility in accommodating large numbers of predictors and ability to handle heterogeneity, has resulted in increasing interest in their application to genetic association and linkage studies.

Classification trees use all variables and relevant cases when creating nodes in a single tree that represents the outcome of the learning process. To the contrary, RFT creates a large number of trees developed on random samples of input cases. The input data for each tree is based on iterative sampling from the original data. The variables used for constructing splits are a random subsample of the complete set of variables. The necessary calculations are carried out tree by tree as the random forest is constructed. RFT estimates the importance of a variable by looking at how much classification error increases when data for that variable are permuted and all other data are left unchanged and calculates ISs. ISs can be used to prioritize the variables in this model. (If the values of this score from tree to tree are independent, then the SE can be computed. The raw score divided by its SE yields a z score, and, thus, a significance level can be assigned to each importance scores (ISs)).

RFT with the aid of variables classifies patients as to whether they are affected by the complication tested, and the results of classification are compared with actual clinical data. As the result of this process, 2 x 2 confusion matrices (true-negative, false-positive, true-positive, and false-negative groups) are created showing the number of patients classified properly or not. Accuracy values are calculated as the number of true negative + true positive classifications/the number of analyzed patients.

As immaturity is the major determinant of perinatal adaptation, the additional protective/worsening role of individual polymorphisms may be masked to different extent at different developmental status and the large heterogeneity of population may lead to unstable results.

The first step in our analysis was the identification of the gray zone of those infants in whom genotype data may improve the prediction of complications.

We added up the number of complications (*i.e.* CF, NEC, ARF, sepsis, PDA, IVH, IRDS) for each infant and gave them a score (range, 0–7; median value, 3). According to this score, infants were assigned a dummy variable (0 for infants with score ≤3 and 1 for infants with score >3).

Then we ranked the infants according to their birth weight. We created a frame of 45 consecutive infants that was shifted along the data set of all infants ranked according to their birth weights by adding of the next and leaving out the smallest infant of the analyzed cohort. At each position of the frame, we calculated the accuracy of classification when it was based solely on patients' birth data and when it was based on birth data plus all tested genotypes. We created a graph representing the alteration of accuracy values of the prediction of morbidity throughout the whole range of infants (Fig. 1). Obviously, the accuracy values with and without genotype data diverged between 680 and 1430 g; for further analyses, this birth weight range was used.

In this group of infants ($n = 100$), we calculated the IS values of all clinical and genotype data for the classification of each perinatal complication. The clinical and genotype data with the highest IS values (with $p < 0.10$) were used to classify the patients according to the occurrence of severe morbidity, NEC, ARF, IVH, IRDS, CF, PDA, and sepsis; the accuracy of classification was compared with that performed solely based on birth data.

The analyses were performed using Fortran 77 codes (9) and in the *R* statistical environment (10) using the "party" package (11). Program codes are available upon request.

RESULTS

First, as described in detail in the previous section, we searched for that range of birth weight for which genotype data may help classify patients according to the presence or absence of a complication; this range has been established as 680–1430 g and contained 100 infants. The IS values of all clinical and genotype data for the classification of each perinatal complication were calculated; IS values for the classification of patients according to severe morbidity dummy variables are visualized on a heat map figure (Figure S1, supplemental material online). The clinical and genotype data with the highest IS values (with $p < 0.10$) are listed in Table 2.

For each complication (except ARF), the most significant predictors included birth weight and gestational age. Interestingly, some genetic polymorphisms are also major predictors,

Table 2. *List of variables with highest ISs in the classification of patients according to perinatal complications*

Complication	Predictor	IS	p
Severe morbidity (dummy variable)	Gestational age	10.529	<0.001
	Weight	5.078	<0.001
	IL-18^{-137}C	2.619	0.004
	HSP 71^{190}C	2.596	0.005
	E-selectin^{128}Arg	2.547	0.005
	VEGF^{-460}C	2.448	0.007
	ACE D	2.297	0.011
NEC	IL-12 p40 GC	4.204	<0.001
	Gestational age	3.576	<0.001
	IL-10^{-1082}A	3.284	0.001
	VEGF^{-460}C	3.272	0.001
	Weight	1.814	0.035
	ACE D	1.789	0.037
ARF	IL1-β^{3954}T	1.818	0.035
	IL 4-receptor α^{1902}G	1.629	0.052
	AT1-R^{1166}C	1.618	0.053
	P-selectin^{715}Pro	1.461	0.072
IVH	Gestational age	9.113	<0.001
	ER 2 PvuII p	3.254	0.001
	VEGF^{-460}C	3.182	0.001
	IL-6^{-174}C	2.541	0.006
	AT1-R^{1166}C	2.125	0.017
	VEGF^{+405}G	1.715	0.043
	IL-12 p40 GC	1.466	0.071
	Ang^{-35}C	1.414	0.079
	HSP 71^{190}C	1.300	0.097
CF	Gestational age	6.583	<0.001
	weight	6.539	<0.001
	HSP 71^{190}C	5.835	<0.001
	IFN-gamma^{+874}A	3.748	<0.001
	Ang^{-35}C	3.142	0.001
	ACE D	2.439	0.007
	IL-4 receptor α^{1902}G	1.840	0.033
	ER 2 PvuII p	1.649	0.05
	TLR-4^{399}Ile	1.499	0.067
PDA	Gestational age	7.073	<0.001
	Weight	3.877	<0.001
	VEGF^{+405}G	3.785	<0.001
	ACE D	3.205	0.001
	Ang^{-35}C	2.570	0.005
	IL-12 p40 GC	1.564	0.059
Sepsis	Weight	9.847	<0.001
	Gestational age	6.273	<0.001
	E-selectin^{128}Arg	4.737	<0.001
	TLR-4^{399}Ile	4.717	<0.001
	Ang^{-35}C	3.671	<0.001
	TLR-4^{299}Gly	3.643	<0.001
	L-selectin^{213}Ser	2.622	0.004
	HSP 71^{190}C	1.962	0.025
	IL-18^{-607}A	1.779	0.038
	IL-6^{-174}C	1.543	0.061
IRDS	Gestational age	14.338	<0.001
	Weight	9.732	<0.001
	E-selectin^{128}Arg	3.878	<0.001
	TLR-4^{399}Ile	3.647	<0.001
	IL-12 p40 GC	3.637	<0.001
	IL-18^{-137}C	3.198	0.001
	TNF-α^{-238}A	2.229	0.013
	IL-18^{-607}A	2.153	0.016
	VEGF^{+405}G	1.545	0.061
	P-selectin^{715}Pro	1.289	0.099

Table 3. *Accuracy of classification of patients with the use of solely birth data or birth data plus variables with the highest ISs*

	Classification according to birth data			Classification according to variables with the highest ISs		
	Unaffected	Affected	Accuracy	Unaffected	Affected	Accuracy
Morbidity dummy variable						
Negative	64	28	0.69	64	20	0.77
Positive	3	5		3	13	
NEC						
Negative	55	33	0.60	53	21	0.70
Positive	7	5		9	17	
ARF						
Negative	70	30	0.70	68	22	0.76
Positive	0	0		2	8	
IVH						
Negative	65	30	0.69	56	27	0.63
Positive	1	4		10	7	
IRDS						
Negative	30	11	0.72	34	10	0.77
Positive	17	42		13	43	
CF						
Negative	68	27	0.68	70	24	0.76
Positive	2	3		0	6	
PDA						
Negative	52	36	0.57	49	26	0.64
Positive	7	5		10	15	
Sepsis						
Negative	75	25	0.75	71	20	0.76
Positive	0	0		4	5	

and in some complications, their IS approached or even surpassed that of birth data (such as IL-12 p40 GC for NEC). For ARF, the ISs were low in general and did not include birth weight and gestational age, just some genetic polymorphisms.

Based according to calculated ISs, we created optimized classification patterns that included each IS with a *p* value <0.10. At this step, we calculated the accuracy of classification and related to that based solely on birth data (Table 3).

The accuracy of classification according to overall morbidity, NEC, ARF, IRDS, CF, and PDA improved by 0.06–0.10, when the genotype data with the highest IS values were also incorporated to the analysis.

DISCUSSION

During past years, several associations between genetic variants and perinatal complications [including CF (12), sepsis (4), PDA (5), ARF (3), NEC (1), and IVH (6)] have been reported. These studies are generally of case-control design, and, therefore, efforts have been done to enroll populations as homogeneous as possible. However, it is likely that preterm infants' investigated populations are still heterogeneous in terms of gestational age, applied therapies, *etc*. Therefore, logistic regression analysis is generally used to adjust the genotype-phenotype association to risk factors. The downside of this approach is that, as a result of this process, several subgroups with very low patient numbers are created and this leads to the instability of this model. This is a probable explanation for contradictory results of genetic studies performed in preterm infants with sepsis (4,13,14), disturbances of perinatal cardiorespiratory adaptation (12,15), or chronic lung disorder (16–19). Recognizing this limitation, we used

RFT, a new statistical approach, to assess whether genotypes may help to identify infants at risk. RFT is a type of high-dimensional nonparametric predictive model that consists of a collection of classification or regression trees. Thus, RFT appears to be particularly well suited to address a primary problem posed by large-scale association studies (20).

The results of our analyses are in accordance with everyday practice: the major determinant of perinatal complications is prematurity reflected by low birth weight and gestational age. Indeed, this is in line with the general clinical experience as proficient neonatologists are usually able to classify the preterm infant in the delivery room whether he or she is at a higher risk of perinatal complications. However, can the accuracy of classification be improved when some information about genetic variants is also known?

Before addressing this question, we assumed that the impact of genotypes on perinatal complications is influenced by developmental status. Infants with low gestational age are probably more susceptible to disturbances of perinatal adaptation independently of their genetic inheritance; therefore, genetic variants play fewer roles in the complications in a very immature population. In contrast, the importance of inherited susceptibility is higher in more mature infants, but the incidence of perinatal morbidities among near-term infants is low; therefore, a large number of patients is required to investigate the role of genotypes in perinatal complications in these populations. To identify that gray zone preterm population in whom prematurity itself does not fully mask the effect of genetic variants, but the incidence of morbidity is still high enough to be investigated, we screened the importance of 24 genetic variants in 135 preterm infants.

In agreement with our hypothesis, the generated curves representing the accuracy of classifications of infants according to the presence of several morbidities with and without genotype data diverged somewhat in a subset of neonates, namely in those whose birth weight ranged between 680 and 1430 g. We analyzed this population further.

We demonstrated that knowing the carrier state of some genetic polymorphisms in this subgroup of infants might help to predict the risk of some perinatal complications. The identified genes are partly identical to those reported previously in other experimental settings and case-control studies (such as IL-12 p40 GC for NEC; ER 2 *Pvu*II p and ACE D for CF; for references, see Table 1).

It is worth mentioning that RFT revealed other polymorphisms not detected by the standard statistical approaches, which may indicate other directions for research and therapy. Although our results might give rise to several hypotheses, we should stress that our analysis was of descriptive nature and focused on finding possible predictive markers of perinatal complications.

Moreover, there are some limitations of our analysis. Although RFT is a robust method, and especially suitable for studies with low patient numbers (21), the number of patients analyzed is still small. Our population is possibly biased by the over- and underrepresentation of some complications. Furthermore, we considered just 24 genetic variants, and it is reasonable to postulate that the expansion of the analysis to other genotypes may reveal the importance of additional genetic variants as well. Finally, we omitted important clinical conditions such as intrauterine growth retardation, postnatal therapeutic interventions, and drugs from our analysis and focused exclusively on the additional predictive value of genotypes to that of birth data. Considering these factors in a larger number of patients would specify, however, further subgroups with special genetic susceptibility to perinatal complications. We should also stress that the accuracy we obtained with or without genetic patterns is still far less than 1.00, indicating that there are still major unrecorded perinatal, postnatal, and, possibly, genetic factors that contribute to the risk of perinatal morbidities.

To sum up our results, what are the implications of our analysis for the interested clinicians? First, we think that our method offers a new statistical way to test the possible role of genetic polymorphisms in disease prediction of preterm infants. Second, we identified a subgroup in that the association between genotype and complications is worthy of investigation. Third, our holistic approach designates those genetic variants that may contribute to the most important perinatal complications in preterm infants.

REFERENCES

1. Treszl A, Tulassay T, Vásárhelyi B 2006 Genetic basis for necrotizing enterocolitis—risk factors and their relations to genetic polymorphisms. Front Biosci 11:570–580
2. Parton LA, Strassberg SS, Qian D, Galvin-Parton PA, Cristea IA 2006 The genetic basis for bronchopulmonary dysplasia. Front Biosci 11:1854–1860
3. Vásárhelyi B, Tóth-Heyn P, Treszl A, Tulassay T 2005 Genetic polymorphisms and risk for acute renal failure in preterm neonates. Pediatr Nephrol 20:132–135
4. Baier RJ, Loggins J, Yanamandra K 2006 IL-10, IL-6 and CD14 polymorphisms and sepsis outcome in ventilated very low birth weight infants. BMC Med 4:10–
5. Treszl A, Szabó M, Dunai G, Nobilis A, Kocsis I, Machay T, Tulassay T, Vásárhelyi B 2003 Angiotensin II type 1 receptor A1166C polymorphism and prophylactic indomethacin treatment induced ductus arteriosus closure in very low birth weight neonates. Pediatr Res 54:753–755
6. Baier RJ 2006 Genetics of perinatal brain injury in the preterm infant. Front Biosci 11:1371–1387
7. Breiman L 2001 Random forests. Mach Learn 45:5–32
8. Bureau A, Dupuis J, Falls K, Lunetta KL, Hayward B, Keith TP, Van Eerdewegh P 2005 Identifying SNPs predictive of phenotype using random forests. Genet Epidemiol 28:171–182
9. Breiman L, Cutler A 2004 Random forests. Available at: http://www.stat.berkeley.edu/~breiman/RandomForests/cc_software.htm (accessed June 20, 2007)
10. R Development Core Team 2006 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org (accessed June 20, 2007)
11. Hothorn T, Hornik K, Zeileis A 2006 Party: a laboratory for recursive part(y)itioning (R package version 0.9-0) 2006 Available at: http://CRAN.R-project.org/ (accessed June 20, 2007)
12. Nobilis A, Szabó M, Kocsis I, Sulyok E, Tulassay T, Vásárhelyi B 2006 Angiotensin-converting enzyme DD genotype is preventive against circulatory failure in very low-birth-weight neonates. Acta Paediatr 95:747–750
13. Treszl A, Kocsis I, Szathmari M, Schuler A, Heninger E, Tulassay T, Vásárhelyi B 2003 Genetic variants of TNF-α, IL-1beta, IL-4 receptor α-chain, IL-6 and IL-10 genes are not risk factors for sepsis in low-birth-weight infants. Biol Neonate 83:241–245
14. Ahrens P, Kattner E, Kohler B, Hartel C, Seidenberg J, Segerer H, Moller J, Gopel W Genetic Factors in Neonatology Study Group 2004 Mutations of genes involved in the innate immune system as predictors of sepsis in very low birth weight infants. Pediatr Res 55:652–656
15. Harding D, Dhamrait S, Marlow N, Whitelaw A, Gupta S, Humphries S, Montgomery H 2003 Angiotensin-converting enzyme DD genotype is associated with worse perinatal cardiorespiratory adaptation in preterm infants. J Pediatr 143:746–749
16. Bokodi G, Derzbach L, Vásárhelyi B 2006 Re: Deletion allele of angiotensin-converting enzyme. J Pediatr 149:579–580
17. Kazzi SN, Quasney MW 2005 Deletion allele of angiotensin-converting enzyme is associated with increased risk and severity of bronchopulmonary dysplasia. J Pediatr 147:818–822
18. Adcock K, Hedberg C, Loggins J, Kruger TE, Baier RJ 2003 The TNF-α −308, MCP-1-2518 and TGF-beta1 +915 polymorphisms are not associated with the development of chronic lung disease in very low birth weight infants. Genes Immun 4:420–426
19. Bokodi G, Treszl A, Derzbach L, Balogh Á, Vásárhelyi B 2005 The association of the carrier state of the tumor necrosis factor-α (TNFα) (-308)A allele with the duration of oxygen supplementation in preterm neonates. Eur Cytokine Netw 16:78–80
20. Diaz-Uriarte R, Alvarez de Andres S 2006 Gene selection and classification of microarray data using random forest. BMC Bioinform 7:3
21. Heidema AG, Boer JM, Nagelkerke N, Mariman EC, van der A DL, Feskens EJ 2006 The challenge for genetic epidemiologists: how to analyze large numbers of SNPs in relation to complex diseases. BMC Genet 7:23