

RESEARCH

Pre-treatment growth and IGF-I deficiency as main predictors of response to growth hormone therapy in neural models

Urszula Smyczyńska¹, Joanna Smyczyńska², Maciej Hilczer^{2,3}, Renata Stawerska², Ryszard Tadeusiewicz¹ and Andrzej Lewiński^{2,4}

¹Department of Automatics and Biomedical Engineering, AGH University of Science and Technology, Krakow, Poland

²Department of Endocrinology and Metabolic Diseases, Polish Mother's Memorial Hospital – Research Institute, Lodz, Poland

³Department of Paediatric Endocrinology, Medical University of Lodz, Lodz, Poland

⁴Department of Endocrinology and Metabolic Diseases, Medical University of Lodz, Lodz, Poland

Correspondence should be addressed to J Smyczyńska: joanna.smyczynska@o2.pl

Abstract

Mathematical models have been applied in prediction of growth hormone treatment effectiveness in children since the end of 1990s. Usually they were multiple linear regression models; however, there are also examples derived by empirical non-linear methods. Proposed solution consists in application of machine learning technique – artificial neural networks – to analyse this problem. This new methodology, contrary to previous ones, allows detection of both linear and non-linear dependencies without assuming their character *a priori*. The aims of this work included: development of models predicting separately growth during 1st year of treatment and final height as well as identification of important predictors and in-depth analysis of their influence on treatment's effectiveness. The models were derived on the basis of clinical data of 272 patients treated for at least 1 year, 133 of whom have already attained final height. Starting from models containing 17 and 20 potential predictors, respectively for 1st year and final height model, we were able to reduce their number to 9 and 10. Basing on the final models, IGF-I concentration and earlier growth were indicated as belonging to most important predictors of response to GH therapy, while results of GH secretion tests were automatically excluded as insignificant. Moreover, majority of the dependencies were observed to be non-linear, thus using neural networks seems to be reasonable approach despite it being more complex than previously applied methods.

Key Words

- ▶ growth hormone deficiency
- ▶ insulin-like growth factor-I
- ▶ growth hormone therapy
- ▶ prediction models
- ▶ artificial neural networks

Endocrine Connections
(2018) 7, 239–249

Introduction

Growth hormone (GH) therapy is widely approved in children with short stature caused by GH deficiency (GHD). It has also been documented that GH therapy may be beneficial in children with idiopathic short stature (ISS); however, significant variability in its effectiveness in different patients is observed (1, 2). Taking into account the inconsistent and sometimes disappointing effects of treatment, the need to create prediction models of growth response to GH therapy has been defined by Ranke as a new

paradigm for GH treatment in 21st century (3). Moreover, the use of growth prediction models has been proposed as a personalized approach to GH treatment in clinical practice (4). Development of such models should allow to determine more precisely the factors that influence responsiveness to GH, to facilitate realistic expectations concerning the therapy outcome and to identify poor responders, as well as to individualize and optimize treatment modalities (3, 4, 5). Prediction of GH therapy

effectiveness should be a part of strategy for dealing with children with short stature, subjected to GH therapy (5).

In recent years, a number of prediction models have been created for different groups of patients and for different time horizons. Most frequently the models have been derived for prepubertal children with GHD (6, 7, 8, 9, 10, 11), while others were developed either for children in different pubertal stages (12, 13) or separately for prepubertal and pubertal ones (11). In some models, first-year response to treatment has been the only output variable (7, 13), while in others, the predictions have been done for few initial years of therapy (6, 8, 9, 10) or for the attained FH (11, 12). In one model, the total pubertal growth has been predicted (14).

Previously published models have been included as input variables – either only the data available before GH therapy onset or also the information gathered during treatment (6, 10, 11, 13, 14). Besides auxological parameters, GH peak after pharmacological stimulation has been considered as potential predictor of the therapy outcome in all the above-mentioned models (6, 7, 8, 9, 10, 11, 12, 13, 14), in some of them also concentrations of IGF-I (7, 8, 9, 13) and IGFBP-3 (7, 9) have been processed.

The most frequently used method of deriving models of GH therapy outcomes has been multiple linear regression (MLR) (6, 7, 10, 11, 12, 13, 14); however, in two models, non-linear technique of empirical curve fitting has been applied (8, 9). The main limitation of MLR models is the necessity to assume linear dependencies between input and output variables, while the real relationships between them can be non-linear and sometimes remain poorly defined or even unknown.

Development of advanced computational statistical methods (machine learning) seems to give the chance to overcome these limitations. Our research group has recently presented for the first time models of prediction of FH of GH-deficient children with the use of artificial neural networks (ANN) (15). Neural networks are very complex computational systems, designed to some extent to resemble neuronal connections in the brain. There are different types of ANN used for modeling of biological processes. In present study, multilayer perceptron (MLP) was selected as a model for prediction of response to GH treatment. This approach was chosen, because it does not require assuming any particular character of dependencies between explained variable (output) and explanatory ones (input). A neural model typically consists of:

- input layer to which explanatory data are delivered,
- hidden layers (1 or 2) that process the data,

- output layer that delivers final result of prediction.

Each of these layers consists of several basic computational units – artificial neurons that process the data in order to produce desired result. In MLP hidden and output neurons calculate the weighted sum of all their inputs and transform it by sigmoid function. Every layer passes results of its calculations to the next until the output is reached and the solution presented to model's user. The process of derivation of neural models, called training, consists of adjusting the weight coefficients of all neurons so that they produce the result that is as close as possible to desired one. Several algorithms exist that can be applied for this purpose. Before application, each model should be tested on a new data set, not used for its derivation.

From the practical point of view, the most important measures of prediction models accuracy are prediction errors (i.e. differences between real and predicted values of the output variable) and the proportion of variability explained and unexplained by the model. Very important is also the knowledge, which of the potential predictors have a real influence on the output value and which are redundant. Next essential issue is to determine the relationships between particular input variables and the output prediction. Finally, the model should enable clinicians to identify non-responders, including both the patients in whom the prediction is poor, as well as those, who do not improve growth despite good prediction.

Subjects and methods

Prediction models of GH therapy effectiveness were derived on the basis of data collected from 272 patients (197 boys, 75 girls) treated with GH due to isolated, non-acquired GHD in 2004–2016 for at least one year (1st year response model), including 133 patients (89 boys, 44 girls) treated up to the attainment of FH (final height model). In both models, most of children were prepubertal at therapy onset; however, 1st-year response model included 59 children who entered puberty before treatment (37 boys, 22 girls), while FH model included 39 ones (24 boys, 15 girls). The exclusion criteria were chronic diseases, multiple pituitary hormone deficiency, genetic syndromes, malnutrition, acquired GHD (brain tumors, cranial irradiation, injuries, etc.). Studies on GH therapy effectiveness were approved by the Committee of Ethics of Scientific Research in Polish Mother's Memorial Hospital – Research Institute in Lodz. Written informed consent has been obtained from each patient and parents

or legal guardians after full explanation of the purpose and nature of all procedures used.

Characteristics of the cohort of patients

The following data concerning the therapy onset or patients' history were collected and pre-processed if necessary:

1. height at therapy onset (H_0), expressed as height SDS for age and gender according to Polish reference data ($H\ SDS_0$) (16);
2. gender (G), transformed to numerical values 0 for boys and 1 for girls;
3. chronological age at therapy onset (CA);
4. height at least 6 months before treatment (H_{BT}) and $H\ SDS_{BT}$, respectively (16);
5. $H\ SDS\ V_0$ – rate of changes of H SDS before treatment
6. body mass (M) and its SDS calculated with respect to height ($M\ SDS$) (16);
7. heights of mother (H_M) and father (H_F), expressed as $H_M\ SDS$ and $H_F\ SDS$, respectively;
8. patient's bone age (BA), assessed according to Greulich-Pyle standards (17), expressed as BA/CA ratio;
9. pubertal stage (PUB), transformed to numerical values: 0 – for prepubertal children, 1 – for pubertal ones;
10. gestational age (GA) and birth weight (BW), expressed as $BW\ SDS$ for GA , according to (18);
11. GH peak after falling asleep (GH_N) – an obligatory screening procedure in Poland, and in two stimulation tests: with clonidine $0.15\ mg/m^2$ orally (GH_C) and with glucagon $30\ \mu g/kg$ i.m., not exceeding $1.0\ mg$ (GH_G); GHD was defined as GH peak below $10.0\ ng/mL$ in all tests; however, some patients with higher GH peak in one test were included to the study, as ones with either neurosecretory dysfunction or with normal GH peak only in nocturnal test (performed before its implementation as an obligatory screening procedure in Poland). This may constitute a bias of the study; however, we decided to include these patients as it was previously reported that assessment of nocturnal GH secretion may not identify all the patients with GHD (1, 19, 20); concentrations of GH were measured with IMMULITE 2000 system;
12. IGF-I and IGFBP-3 serum concentrations, expressed as $IGF-I\ SDS$ and $IGFBP-3\ SDS$ for age and gender, respectively, assuming the log-normal distribution of IGF-I and IGFBP-3 concentrations, according to (21); concentrations of IGF-I and IGFBP-3 were measured with IMMULITE 2000 system;

13. initial GH dose in mg per week per kg of body mass (D); patients were treated with different initial doses of GH from the recommended range, chosen by their physicians, according to their clinical experience.

Additionally, the following data were obtained during or after 1st year of treatment and included as input variables in FH model:

1. patient's height after 1 year of rhGH administration (H_1), expressed as $H_1\ SDS$ (16) and H SDS increase during 1st year of treatment ($H\ SDS\ V_1$), the output variable in 1st year response model, calculated according to the formula:

$$H\ SDS\ V_1 = H_1\ SDS - H_0\ SDS$$

2. IGF-I and IGFBP-3 after at least 3 months of treatment ($IGF-I$ and $IGFBP-3$), expressed as $IGF-I\ SDS$ and $IGFBP-3\ SDS$, respectively (21). Finally, the differences between SDS of concentration before and during treatment were calculated and included in the model ($\Delta IGF\ SDS$ and $\Delta IGFBP-3\ SDS$).

The attainment of FH was confirmed by patients' BA of 16 years for girls and 18 years for boys (22). The output variable in FH model was $FH\ SDS$ (16).

Statistical characteristics of the whole group of patients are collected in Table 1 (data of subgroup who attained FH are presented in brackets). For the purpose of model derivation, whole datasets were randomly divided into 3 subsets:

- training set: 60% of data (163 cases for 1st year and 79 for FH prediction),
- validation set: 20% of data (54 cases for 1st year and 27 for FH prediction),
- testing set: 20% of data (55 cases for 1st year and 27 for FH prediction).

Neural models derivation and analysis

MLP networks for both 1st-year response and FH prediction were trained either by backpropagation (23) or conjugate gradient (24) algorithm. Parameters of networks (number of hidden layers, number of neurons) and training were selected by genetic algorithm (25), that consists of following sequence of operations:

1. Generation of initial population of networks (in our case 200 of them) with parameters set randomly from the range of allowed values.

Table 1 Statistical characteristics of patients' cohort for 1st-year response model and for FH model (in brackets).

| | Mean | S.D. | Minimum | Maximum |
|-------------------------------|---------------|-------------|---------------|---------------|
| <i>H SDS</i> ₀ | -2.80 (-2.68) | 0.68 (0.50) | -6.37 (-4.00) | -1.59 (-1.61) |
| <i>H SDS V</i> ₀ | -0.22 (-0.23) | 0.23 (0.22) | -0.87 (-0.61) | 0.58 (0.58) |
| CA (years) | 10.9 (12.3) | 2.8 (1.5) | 4.2 (4.3) | 14.6 (14.6) |
| <i>M SDS</i> | 0.26 (0.65) | 1.41 (1.26) | -2.8 (-2.08) | 6.31 (3.61) |
| <i>H_M SDS</i> | -1.00 (-0.94) | 0.94 (0.85) | -4.43 (-3.03) | 2.44 (1.93) |
| <i>H_F SDS</i> | -0.92 (-1.06) | 1.13 (1.12) | -5.09 (-5.09) | 3.12 (3.12) |
| BA/CA | 0.75 (0.80) | 0.13 (0.10) | 0.34 (0.46) | 1.02 (1.02) |
| <i>BW SDS</i> | -0.96 (-0.91) | 1.00 (0.96) | -5.38 (-3.8.) | 1.37 (1.23) |
| GA (weeks) | 39.3 (39.4) | 1.3 (1.3) | 35.0 (35.0) | 43.0 (43.0) |
| <i>GH_N</i> (ng/mL) | 6.1 (6.3) | 3.2 (3.7) | 0.0 (0.0) | 26.9 (26.9) |
| <i>GH_K</i> (ng/mL) | 7.1 (7.4) | 4.0 (4.9) | 0.4 (0.4) | 40.0 (40.0) |
| <i>GH_G</i> (ng/mL) | 5.5 (6.0) | 3.2 (3.7) | 0.1 (0.1) | 21.7 (21.7) |
| <i>IGF-I SDS</i> | -1.91 (-1.94) | 1.1 (1.11) | -5.31 (-5.31) | 0.78 (0.78) |
| <i>IGFBP-3 SDS</i> | -0.36 (-0.41) | 0.95 (1.00) | -4.38 (-2.77) | 3.16 (3.16) |
| <i>D</i> (mg/kg/week) | 0.23 (0.23) | 0.03 (0.03) | 0.1 (0.13) | 0.37 (0.37) |
| <i>H SDS V</i> ₁ | 0.66 (0.57) | 0.35 (0.29) | -0.09 (-0.09) | 1.83 (1.45) |
| Δ <i>IGF SDS</i> | (2.54) | (0.98) | (0.26) | (6.37) |
| Δ <i>IGFBP-3 SDS</i> | (0.87) | (0.93) | (-3.87) | (3.17) |
| <i>FH SDS</i> | (-1.04) | (0.73) | (-3.19) | (0.60) |

2. Training and evaluation of obtained models by calculation of root mean square error (RMSE).
3. Elimination of certain portion of the least accurate networks (in this work 30%).
4. Generation of equal number of new models whose parameters are inherited from pairs of remaining networks (each parameter is assigned randomly to have a value coming from one of the two 'parent' networks) and their training.
5. Repeating steps 3 and 4 several times (here 5 times).

The best networks, in terms of accuracy of prediction, were chosen from final population predicting 1st-year response to GH treatment and FH SDS. Those networks were subjected to input data reduction by input cancellation algorithm (26). It is a model-based approach of redundant explanatory variables removal that is realized in the following way:

1. RMSE is calculated for original dataset.
2. Each input variable is separately replaced in whole dataset with its mean calculated over all cases from this set.
3. RMSE is calculated for all modified dataset.
4. A variable is removed that causes the lowest increase in RMSE when replaced by its mean.
5. Above operations are repeated until only one variable remains in the model.

The assumption behind the algorithm is that changing the value of important predictor influences strongly

the prediction and aggravates its accuracy. At the same time, an insignificant variable is not used by the model in producing the result, therefore, assigning any value (e.g. average over whole set) in its place does not change prediction much. The last step of using this method is selection of number of significant inputs. It is done on the basis of sequence of RMSE values (or coefficient of determination, R^2), calculated after removal of each variable. A rapid decrease in R^2 (or increase in RMSE) means that an important predictor was eliminated and indicates the moment when reduction should be finished. The process of redundant variables elimination is performed automatically and the whole procedure is similar; however – from the technical point of view – much complicated than step-by-step backward elimination of insignificant variables in MLR models.

The networks derived by input cancellation were investigated further in order to identify the role of each input in prediction of response to GH treatment. The task was realized by sensitivity analysis in a similar way to approaches presented in the studies on ecological predictions (27, 28). It consisted in plotting the relationships between input and predicted output (P *HSDS V*₁ for 1st-year response and P *FH SDS* for final height) for whole range of values of each continuous explanatory variable, when all other such inputs assumed the mean value over whole dataset. Binary inputs (gender and pubertal stage) were assigned one of the possible values that are indicated for each figure.

Results

From the population of 200 networks for 1st year prediction, produced in final iteration of genetic algorithm, a network with 2 hidden units – MLP 17:17-2-1:1 (number of input variables: #n of input neurons – #n of hidden units in consecutive layers – #n of output neurons: #n of outputs) was selected as the most accurate one. Similar procedure developing a model for final height indicated MLP 20:20-3-2-1:1 as the best model with all potential predictors. MLP 17:17-2-1:1 was trained by backpropagation, while MLP 20:20-3-2-1:1 by conjugate gradients algorithm. Both of those networks are presented schematically in Fig. 1; input variables abbreviations are explained in the 'Characteristics of the cohort of patients' section where they are indicated by bold font.

Those two models were subjected to input reduction by input cancellation algorithm. Results of this operation are presented in Fig. 2 as values of coefficient of determination (R^2) after removal of each input. Basing on the plot for 1st year (Fig. 2A), it was observed that 7 variables can be eliminated with only slight decrease in testing R^2 ; therefore, it was decided that final model should contain 9 inputs (MLP 9:9-2-1:1). In the model for final height, 10 explaining variables could be removed without significant change to R^2 (Fig. 2B), so final model was MLP 10:10-3-2-1:1. Variables that remained in both models are illustrated in Fig. 2C, where blue ellipse indicates 1st-year model and pink one FH model. Both reduced networks are presented in Fig. 3. Summary of errors and coefficients of determination calculated for both full and selected reduced models is presented in Table 2.

After input reduction, a sensitivity analysis was performed for reduced models MLP 9:9-2-1:1 and MLP 10:10-3-2-1:1. Input–output dependencies in 1st-year model are presented in Fig. 4; all of them were done for prepubertal children, separately for girls (red curves) and boys (blue curves). The observed relationships are non-linear and slightly different for boys and girls.

Analogical analysis for FH model (MLP 10:10-3-2-1:1) is shown in Fig. 5. In this case, gender and pubertal stage were not included in the model as redundant variables; therefore, the results are valid for both boys and girls either prepubertal or pubertal.

Finally, we were interested in accuracy of prediction for particular patients. It was analyzed with the use of Bland–Altman plots (Fig. 6). For 95% of patients, prediction error did not exceed 0.5 s.d. in the model of 1st-year response to GH therapy and 1.0 s.d. in the model for FH; the mean error was about 0.27 s.d. and 0.50 s.d., respectively (for details see Table 2).

Discussion

In the present study, ANN model of 1st-year response to GH therapy explained 45% of its variability, with RMSE about 0.27 s.d./year that corresponds to about 1 cm/year. Other models of 1st-year response to treatment were created only for prepubertal children and explained from 33% (10) to 70% (7) of its variability, with the error of 1.46 cm (6) to 2.5 cm (7) in linear models and 0.23 s.d. in non-linear model (8). We did not manage

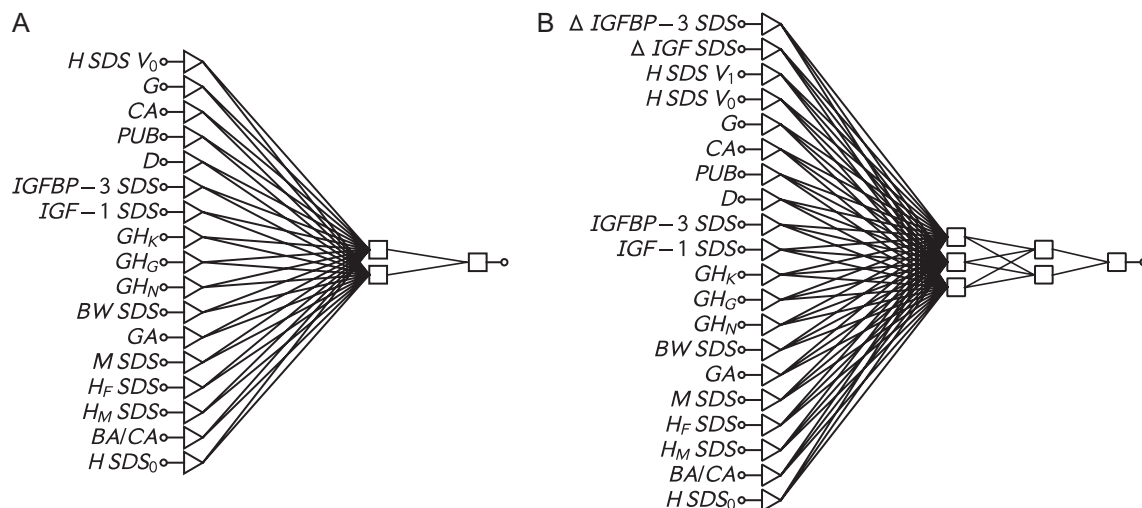
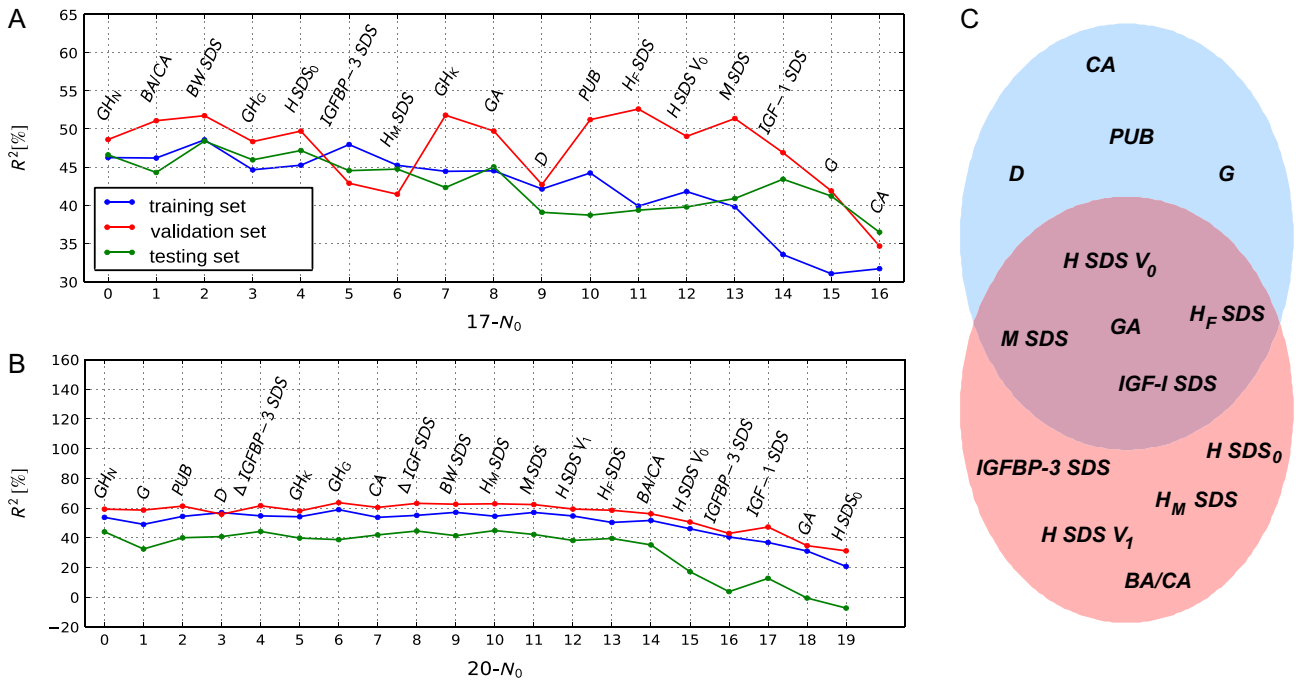


Figure 1 Schematic representations of models for 1st-year response to GH treatment (A) and final height (B) with all available variables.



to find any published model of 1st-year response to GH therapy for pubertal children or for children with different stages of puberty. Unfortunately, in some of the published models, the data on either explained variability or prediction error are missing. Moreover, despite the fact that the common model has been created for boys and girls, some gender-related differences in

the relationships among the analyzed variables could be observed.

The best ANN model of FH SDS presented in the current study also explained 45% of its variability for testing group, with RMSE not exceeding 0.50s.d. (that corresponds to about 3-cm difference between predicted and attained FH). For comparison, our previous model

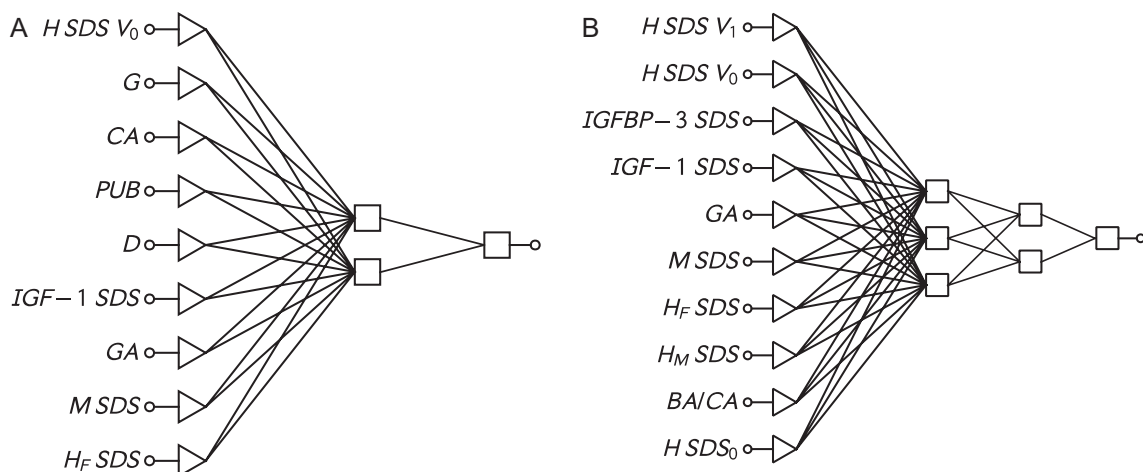


Figure 3
 Schematic representations of reduced models for 1st-year response to rhGH treatment (A) and final height (B).

Table 2 RMSE and R^2 for full and reduced models predicting 1st-year response to GH treatment and FH SDS.

| Prediction horizon | Network architecture | RMSE | | | R^2 (%) | | |
|--------------------|----------------------|-------|-------|-------|-----------|------|------|
| | | TR | V | TS | TR | V | TS |
| 1st year | MLP 17:17-2-1:1 | 0.258 | 0.255 | 0.267 | 42.8 | 47.3 | 48.7 |
| | MLP 9:9-2-1:1 | 0.254 | 0.249 | 0.277 | 44.5 | 49.7 | 45.0 |
| FH | MLP 20:20-3-2-1:1 | 0.489 | 0.498 | 0.493 | 53.7 | 59.3 | 44.0 |
| | MLP 10:10-3-2-1:1 | 0.485 | 0.476 | 0.498 | 54.5 | 62.9 | 44.9 |

MLP number (n#) of inputs: n# of input neurons, n# of neurons in consecutive hidden layers, n# of output neurons: n# of outputs. TR, training set; TS, testing set; V, validation set.

that included only pre-treatment data explained 42% variability of FH SDS with RMSE 0.63s.d. Other models for FH prediction were derived either for only prepubertal children or separately for prepubertal and pubertal ones, which makes the direct comparison of our results with data from other studies impossible. However, the prediction error in our FH model was similar or even lower than that in published MLR models (11, 29). In our previous study (15), slightly better accuracy together with higher rate of

explained variability of FH have been observed for ANN model vs MLR model, derived on the same data.

The best MLP model for 1st-year response contained 9 input variables selected out of 17 potential ones, as the most significant input variables there were selected: patients' age at therapy onset, gender and pre-treatment IGF-I SDS, while the model eliminated all the input data concerning direct assessment of GH secretion. The best MLP model for FH prediction eliminated 10 out of 20

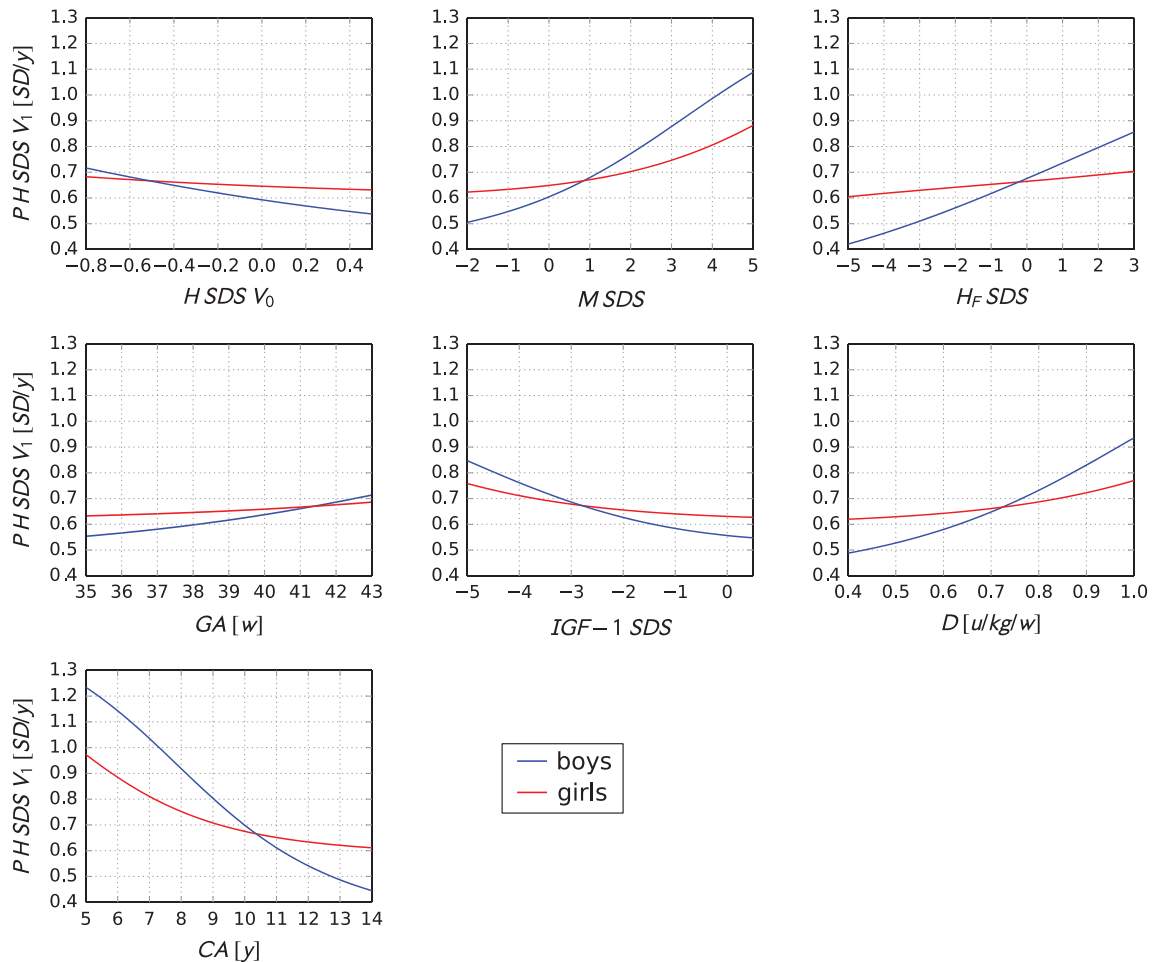


Figure 4 Sensitivity analysis for 1st-year model – MLP 9:9-2-1:1.

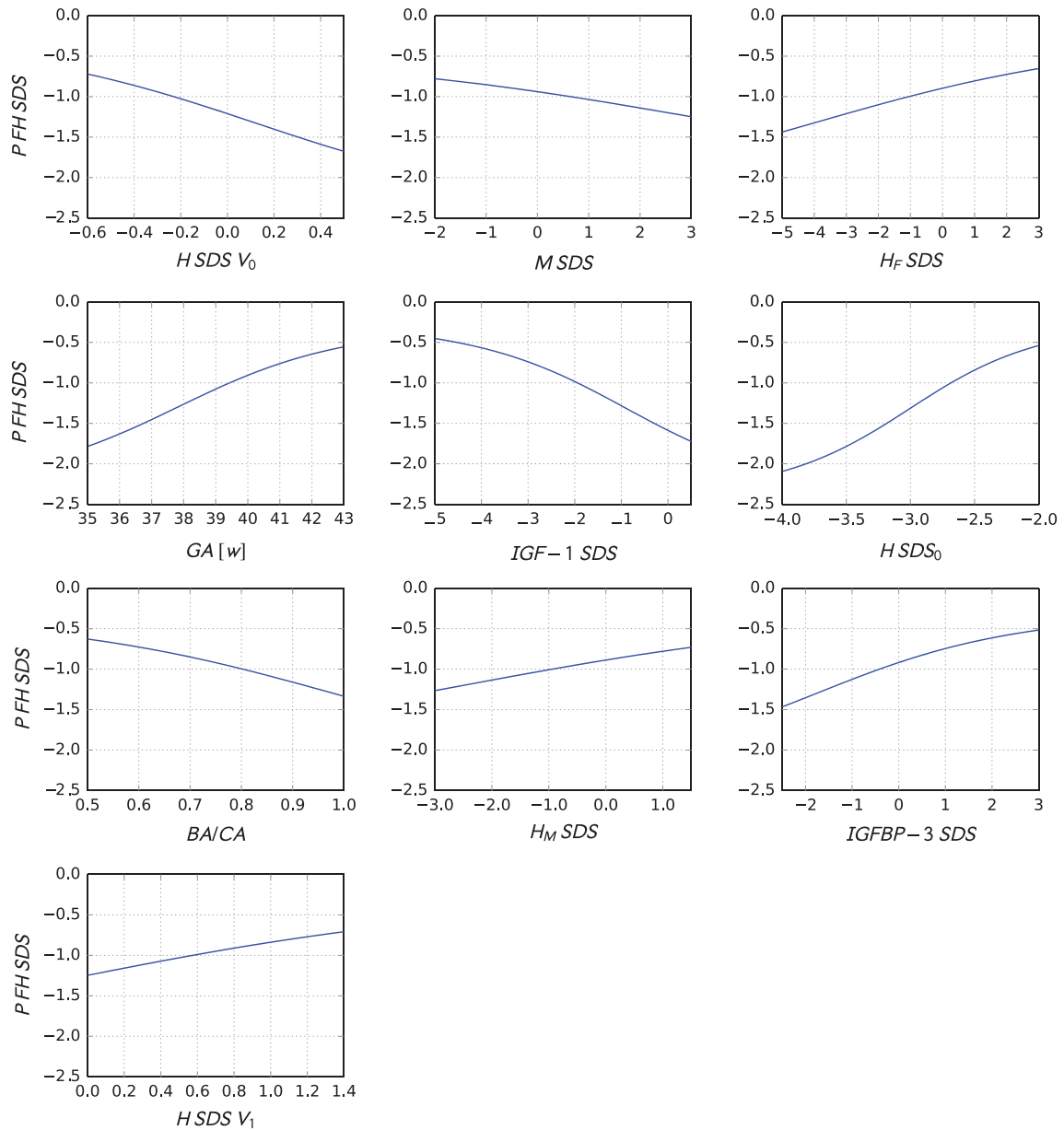


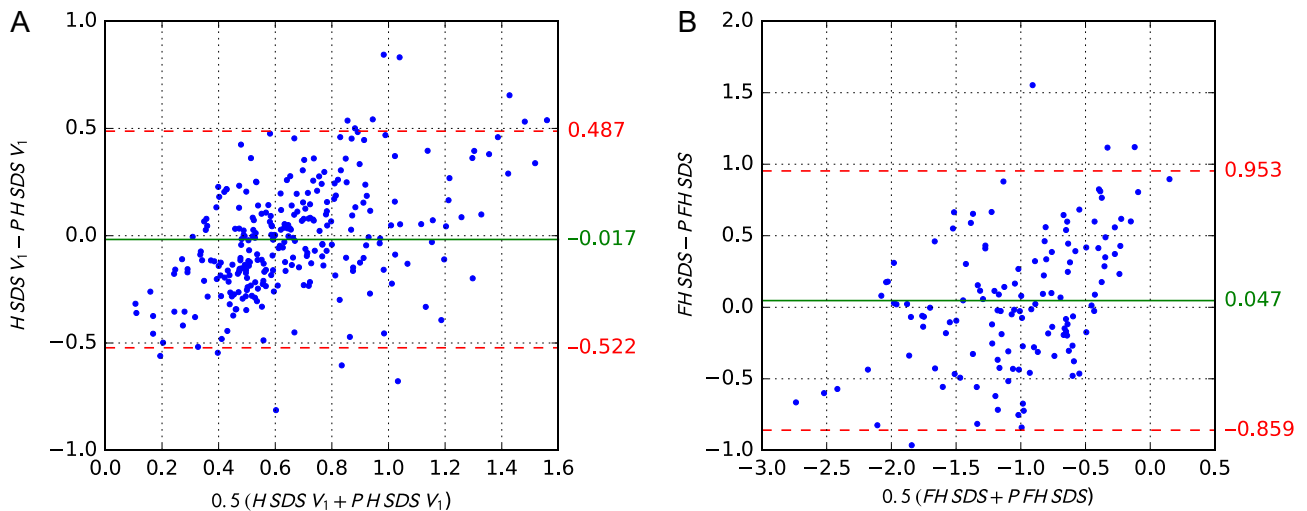
Figure 5
Sensitivity analysis for FH model – MLP 10:10-3-2-1:1.

potential predictors. In this model, the most important variables were patients' height at therapy onset (H_0 SDS), pre-treatment IGF-I SDS and IGFBP-3 SDS, but again, none of GH stimulation tests results.

The problem of eliminating the variables that are considered as crucial for the diagnosis of GHD has been observed also in our previous ANN model for FH prediction, based on pre-treatment data only (15), as well as in some of the previously published linear and non-linear models (6, 11, 13, 14). Moreover, the same situation was reported in the model for patients with ISS (29) and

in some models for patients with the broad spectrum of GH secretion capacities (8, 9). This finding requires further investigations on the significance of the results of GH stimulation tests for GH therapy effectiveness. It seems that the variability of GH response to stimulation should be taken into account, as previous studies have documented both early normalization of decreased GH secretion (30, 31, 32) and the converse situation (33).

In current classifications, GHD is defined as secondary IGF-I deficiency (34, 35). Nonetheless, IGF-I secretion was included as a potential predictor of GH therapy

**Figure 6**

Bland-Altman plots for 1st-year model (A) and FH SDS model (B). Horizontal axis presents mean of real and predicted values, while vertical one prediction error (difference between real and predicted value). Green horizontal line indicates mean difference and red horizontal lines 95% confidence interval.

effectiveness only in a few models (7, 8, 9, 13, 14, 15). More severe IGF-I deficiency presented to be a significant predictor of better response to treatment in all the models except ones with incomplete data on IGF-I concentrations (8, 9). However, in the present study, IGF-I deficiency was not required as inclusion criterion, as in Poland, it is still not necessary to document decreased IGF-I secretion to diagnose GHD. Our results confirm the significance of IGF-I deficiency for better effectiveness of GH therapy. The importance of IGFBP-3 assessment in diagnosing GHD is still a matter of discussion (1, 36). This parameter presented to be a significant input variable in our FH model but not in 1st-year response one. In other studies, IGFBP-3 was included only in two models (7, 9), but it presented to be insignificant.

Creating FH model, we have entered as potential predictors the data obtained during 1st-year of treatment: an increase of H SDS ($H\ SDS\ V_1$), IGF-I SDS ($\Delta\ IGF\ SDS$) and $IGFBP-3\ SDS$ ($\Delta\ IGFBP-3\ SDS$); however, only $H\ SDS\ V_1$ has been included in the final model. With respect to IGF-I significance, similar findings have been presented by Schonau and coworkers. (7). In all published models in which an increase of growth rate during GH therapy in the time period preceding the forecast was included, it has been the most important input variable (6, 10, 11, 29) or even the only significant predictor (9).

The complexity of neural models can seem to be an obstacle in their daily use in clinical practice. However, they could be implemented as user-friendly software

as it was done for instance in the case of the models presented by Elias and coworkers (37). Still, application of models in the field of GH treatment response prediction requires some caution due to usage of various methods of hormonal assessment and differences in growth between populations. Certainly, any model should require calibration before it can be applied to children of different ethnicity or examined by laboratory methods with different reference ranges than those used in patients whose data we used to derive that model.

The undoubted contribution of our study to the deepening knowledge of the importance of particular auxological indices and results of hormonal tests for GH therapy effectiveness are presentations of non-linear dependencies between the analyzed input and output variables. This result is clearly a consequence of application of neural models that are more flexible in finding dependencies in data than other computational methods. Moreover, the progress of information technologies observed in recent years allows relatively easy translation of such complex tools to clinical practice.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

The study was partially supported by statutory funds from Polish Mother's Memorial Hospital Research Institute, Lodz, Poland.

References

- Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, Rossi WC, Feudtner C, Murad MH & Drug and Therapeutics Committee and Ethics Committee of the Pediatric Endocrine Society. Guidelines for growth hormone and insulin-like growth factor-i treatment in children and adolescents: growth hormone deficiency, idiopathic short stature, and primary insulin-like growth factor-I deficiency. *Hormone Research in Paediatrics* 2016 **86** 361–397. (<https://doi.org/10.1159/000452150>)
- Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, Chernausek SD, Savage MO, Wit JM & 2007 ISS Consensus Workshop participants. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the growth hormone research society, the Lawson Wilkins pediatric endocrine society, and the European society for paediatric endocrinology workshop. *Journal of Clinical Endocrinology and Metabolism* 2008 **93** 4210–4217. (<https://doi.org/10.1210/jc.2008-0509>)
- Ranke MB. New paradigms for growth hormone treatment in the 21st century: prediction models. *Journal of Pediatric Endocrinology and Metabolism* 2000 **13** 1365–1370. (<https://doi.org/10.1515/jpem-2000-s609>)
- Wit JM, Ranke MB, Albertsson-Wikland K, Carrascosa A, Rosenfeld RG, Van Buuren S, Kristrom B, Schoenau E, Audi L, Hokken-Koelega AC, et al. Personalized approach to growth hormone treatment: clinical use of growth prediction models. *Hormone Research in Paediatrics* 2013 **79** 257–270. (<https://doi.org/10.1159/000351025>)
- Ranke MB, Lindberg A, Mullis PE, Geffner ME, Tanaka T, Cutfield WS, Tauber M & Dunger D. Towards optimal treatment with growth hormone in short children and adolescents: evidence and theses. *Hormone Research in Paediatrics* 2013 **79** 51–67. (<https://doi.org/10.1159/000347121>)
- Ranke MB, Lindberg A, Chatelain P, Wilton P, Cutfield W, Albertsson-Wikland K & Price DA. Derivation and validation of a mathematical model for predicting the response to exogenous recombinant human growth hormone (GH) in prepubertal children with idiopathic GH deficiency. *Journal of Clinical Endocrinology and Metabolism* 1999 **84** 1174–1183. (<https://doi.org/10.1210/jcem.84.4.5634>)
- Schonau E, Westermann F, Rauch F, Stabrey A, Wassmer G, Keller E, Brämwig J, Blum WF & German Lilly Growth Response Study Group. A new and accurate prediction model for growth response to growth hormone treatment in children with growth hormone deficiency. *European Journal of Endocrinology* 2001 **144** 13–20. (<https://doi.org/10.1530/eje.0.1440013>)
- Dahlgren J, Krström B, Niklasson A, Nierop AFM, Rosberg S & Albertsson-Wikland K. Models predicting the growth response to growth hormone treatment in short children independent of GH status, birth size and gestational age. *BMC Medical Informatics and Decision Making* 2007 **7** 40. (<https://doi.org/10.1186/1472-6947-7-40>)
- Albertson Wikland K, Krström B, Rosberg S, Svensson B & Nierop A. Validated multivariate models predicting the growth response to GH treatment in individual short children with a broad range in GH secretion capacities. *Pediatric Research* 2000 **48** 475–484. (<https://doi.org/10.1203/00006450-200010000-00010>)
- Salah N, Abd El Dayem SM, El Mogy F, Amin M & Ibrahim M. Egyptian growth hormone deficient patients: demographic, auxological characterization and response to growth hormone therapy. *Journal of Pediatric Endocrinology and Metabolism* 2013 **26** 257–269. (<https://doi.org/10.1515/jpem-2012-0091>)
- de Ridder MAJ, Stijnen T & Hokken-Koelega ACS. Prediction of adult height in growth-hormone-treated children with growth hormone deficiency. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 925–931. (<https://doi.org/10.1210/jc.2006-1259>)
- Carel JC, Ecosse E, Nicolino M, Tauber M, Leger J, Cabrol S, Bastié-Sigeac I, Chaussain JL & Coste J. Adult height after long term treatment with recombinant growth hormone for idiopathic isolated growth hormone deficiency: observational follow up study of the French population based registry. *BMJ* 2002 **325** 70–76. (<https://doi.org/10.1136/bmj.325.7355.70>)
- Lee PA, Germak J, Gut R, Khutoryansky N & Ross J. Identification of factors associated with good response to growth hormone therapy in children with short stature: results from the ANSWER Program®. *International Journal of Pediatric Endocrinology* 2011 **2011** 6. (<https://doi.org/10.1186/1687-9856-2011-6>)
- Ranke MB, Lindberg A, Martin DD, Bakker B, Wilton P, Albertsson-Wikland K, Cowell CT, Price DA & Reiter EO. The mathematical model for total pubertal growth in idiopathic growth hormone (GH) deficiency suggests a moderate role of GH dose. *Journal of Clinical Endocrinology and Metabolism* 2003 **88** 4748–4753. (<https://doi.org/10.1210/jc.2003-030600>)
- Smyczyńska J, Hilczer M, Smyczyńska U, Stawerska R, Tadeusiewicz R & Lewiński A. Neural network models – a novel tool for predicting the efficacy of growth hormone (GH) therapy in children with short stature. *Neuroendocrinology Letters* 2015 **36** 348–353.
- Palczewska I & Niedźwiecka Z. Somatic development indices in children and youth of Warsaw. *Medycyna Wieku Rozwojowego* 2001 **5** (2 Suppl 1) 18–118.
- Greulich WW & Pyle SI. *Radiographic Atlas of Skeletal Development of the Hand and Wrist*, 2nd ed. Stanford, CA, USA: Stanford University Press, 1993.
- Niklasson A & Albertsson-Wikland K. Continuous growth reference from 24th week of gestation to 24 months by gender. *BMC Pediatrics* 2008 **8** 8. (<https://doi.org/10.1186/1471-2431-8-8>)
- Rosenfeld RG, Albertsson-Wikland K, Cassorla F, Frasier SD, Hasegawa Y, Hintz RL, Lafranchi S, Lippe B, Loriaux L, Melmed S, et al. Diagnostic controversy – the diagnosis of childhood growth-hormone deficiency revisited. *Journal of Clinical Endocrinology and Metabolism* 1995 **80** 1532–1540. (<https://doi.org/10.1210/jc.80.5.1532>)
- Webb EA & Dattani MT. Diagnosis of growth hormone deficiency. *Endocrine Development* 2010 **18** 55–66. (<https://doi.org/10.1159/000316127>)
- Elmlinger MW, Kühnel W, Weber MM & Ranke MB. Reference ranges for two automated chemiluminescent assays for serum insulin-like growth factor I (IGF-I) and IGF-binding protein 3 (IGFBP-3). *Clinical Chemistry and Laboratory Medicine* 2004 **42** 654–664. (<https://doi.org/10.1515/CCLM.2004.112>)
- Ministry of Health Republic of Poland. Leczenie niskorosłych dzieci z somatotropinową niedoczynnnością przysadki. Warsaw, Poland: Ministry of Health, 2016. (available at: www.mz.gov.pl/leki/refundacja/programy-lekowe)
- Rumelhart DE, Hinton GE & Williams RJ. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, pp 318–362. Eds DE Rumelhart, JL McClelland & CPR Group. San Mateo, CA, USA: Morgan Kaufmann Publishers Inc, 1986. (<https://doi.org/10.1016/B978-1-4832-1446-7.50035-2>)
- Charalambous C. Conjugate gradient algorithm for efficient training of artificial neural networks. *IEE Proceedings G: Circuits, Devices and Systems* 1992 **139** 301–310. (<https://doi.org/10.1049/ip-g-2.1992.0050>)
- Robbins P, Soper A & Rennolls K. Use of genetic algorithms for optimal topology determination in back propagation neural networks. In *Artificial Neural Nets and Genetic Algorithms*, pp 726–730. Eds RF Albrecht, CR Reeves & NC Steele. Vienna, Austria: Springer, 1993. (https://doi.org/10.1007/978-3-7091-7533-0_105)
- Satizábal MHF & Pérez-Urbe A. Relevance metrics to reduce input dimensions in artificial neural networks. In *International Conference on Artificial Neural Networks*, pp 39–48. Eds JM de Sa, LA Alexandre, W Duch & D Mandic. Porto, Portugal: Springer-Verlag, 2007. (https://doi.org/10.1007/978-3-540-74690-4_5)

- 27 Özsesmi SL & Özsesmi U. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling* 1999 **116** 15–31. ([https://doi.org/10.1016/S0304-3800\(98\)00149-5](https://doi.org/10.1016/S0304-3800(98)00149-5))
- 28 Olden JD & Jackson DA. Illuminating the ‘black box’: a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling* 2002 **154** 135–150. ([https://doi.org/10.1016/S0304-3800\(02\)00064-9](https://doi.org/10.1016/S0304-3800(02)00064-9))
- 29 Ranke MB, Lindberg A, Price DA, Darendeliler F, Albertsson-Wikland K, Wilton P, Reiter EO & KIGS International Board. Age at growth hormone therapy start and first-year responsiveness to growth hormone are major determinants of height outcome in idiopathic short stature. *Hormone Research* 2007 **68** 53–62. (<https://doi.org/10.1159/000098707>)
- 30 Loche S, Bizzarri C, Maghnie M, Faedda A, Tziella C, Autelli M, Casini MR & Cappa M. Results of early reevaluation of growth hormone secretion in short children with apparent growth hormone deficiency. *Journal of Pediatrics* 2002 **140** 445–449. (<https://doi.org/10.1067/mpd.2002.122729>)
- 31 Bizzarri C, Pedicelli S, Boscherini B, Bedogni G, Cappa M & Cianfarani S. Early retesting by GHRH+arginine test shows normal GH response in most children with idiopathic GH deficiency. *Journal of Endocrinological Investigation* 2015 **38** 429–436. (<https://doi.org/10.1007/s40618-014-0205-3>)
- 32 Vurali D, Gonc EN, Ozon ZA, Alikasifoglu A & Kandemir N. Clinical and laboratory parameters predicting a requirement for the reevaluation of growth hormone status during growth hormone treatment: retesting early in the course of GH treatment. *Growth Hormone and IGF Research* 2017 **34** 31–37. (<https://doi.org/10.1016/j.ghir.2017.05.003>)
- 33 Hilczer M, Smyczynska J, Stawerska R & Lewinski A. Stability of IGF-I concentration despite divergent results of repeated GH stimulating tests indicates poor reproducibility of test results. *Endocrine Regulations* 2006 **40** 37–45.
- 34 Wit J-M, Ranke MB & Kelnar CJH. The ESPE classification of paediatric endocrine diagnoses. *Hormone Research* 2007 **68** 1–120.
- 35 Savage MO, Burren CP & Rosenfeld RG. The continuum of growth hormone-IGF-I axis defects causing short stature: diagnostic and therapeutic challenges. *Clinical Endocrinology* 2010 **72** 721–728. (<https://doi.org/10.1111/j.1365-2265.2009.03775.x>)
- 36 Murray PG, Dattani MT & Clayton PE. Controversies in the diagnosis and management of growth hormone deficiency in childhood and adolescence. *Archives of Disease in Childhood* 2016 **101** 96–100. (<https://doi.org/10.1136/archdischild-2014-307228>)
- 37 Elias KM, Fendler W, Stawiski K, Fiascone SJ, Vitonis AF, Berkowitz RS, Frenzl G, Konstantinopoulos P, Crum CP, Kedzierska M, *et al.* Diagnostic potential for a serum miRNA neural network for detection of ovarian cancer. *eLife* 2017 **6** 1–28. (<https://doi.org/10.7554/eLife.28932>)

Received in final form 28 November 2017

Accepted 14 December 2017

Accepted Preprint published online 14 December 2017