Completing the Analysis of a Questionnaire About Pediatric Blood Pressure

M. FILOMENA TEODORO
CEMAT, Instituto Superior Técnico
Lisbon University
Av, Rovisco Pais, 1, 1048-001 Lisboa
PORTUGAL
and
CINAV, Naval Research Center
Portuguese Naval Academy
Base Naval de Lisboa
Alfeite, 1910-001 Almada
PORTUGAL
maria.alves.teodoro@marinha.pt

CARLA SIMÃO
Medicine Faculty
Lisbon University
Av. Professor Egas Moniz, 1600-190 Lisboa
PORTUGAL
and
Pediatric Department
Santa Maria's Hospital
Centro Hospitalar Lisboa Norte
Av. Professor Egas Moniz, 1600-190 Lisboa
PORTUGAL
carla.mail@netcabo.pt

Abstract: Pediatric arterial hypertension is a highly prevalent, silent, and perhaps because of this, under-diagnosed in most cases, with multiple repercussions on the health of children and adults. It is imperative that health professionals and family members be aware of HBP existence, the negative consequences associated with it, the risk factors and it prevention. Its expected the knowledge about this pathology increases with the level of education of the family members. A simple questionnaire was built with the aim of easy and quick answers. We can establish that, despite not rejecting the questionnaire in terms of internal consistency, it needs some improvement, for example, reformulation of some questions. The estimated models by GLM have reasonable explication of the data. Exploring the binary outcome issue, we extend this approach using GLMM, specifically, mixed logistics models. The process was unsuccessful, the algorithm did not converge. In [14] can be found a statistical data analysis using partial questionnaire's information. The present article completes such statistical approach estimating a model by Generalized Linear Models (GLM) for relevant remaining questions of questionnaire . Exploring the binary outcome issue, we intend to extend this approach using Generalized Linear Mixed Models (GLMM), but the process is still ongoing.

Key–Words: Pediatric hypertension, Caregiver, Knowledge, Generalized linear model, Multivariate techniques

1 Introduction

In medical, behavioral and social sciences it is usual to get a binary outcome. For example, in the present work is collected information related with the perception of pediatric patients caregivers about high blood pressure (HBP) in childhood where some of the outcomes are binary variables (1='yes'/ 0='no').

The hypertension is a serious public health problem present since childhood [4], silent in most cases and whose only form of early diagnosis is the regular blood pressure (BP) measurement in health monitoring consultations. It is a condition that interferes negatively with numerous functions in the body of the children still in development. Like all chronic diseases, not only interferes with the child but also with his family. For this reason, a clinical and therapeutic approach to the family, instead of a targeted intervention only for the patient, it is usually more effective [10].

The diagnosis of pediatric hypertension (PH) should be based on multiple measurements of BP in medical office environment and at different times [8]. The current recommendation of the European Society of Hypertension (ESH) [8] is the inclusion of the practice assessment arterial blood pressure (ABP) in all children above three years of age, in all health surveillance pediatric attendees, and below this age, if there exist risk factors. The publication of these and other guidance standards enabled the detection of cases of secondary asymptomatic hypertension, which otherwise would be hidden. On the other hand, moderate elevations in BP during childhood are more common than it was thought, especially in teenage years. In [1] Ahern evidences PH as a growing problem in health care, which is underdiagnosed [7]. This pattern is also presented in [6], where the trends and future challenges of this disease are explored.

The study [10] standardized the wide variety of published data and prepared reference tables of ABP values considered normal for each gender depending on the child's age and his height [8]. Regardless the limitations, such work is seen as a reference study and, according to its recommendations, the ABP in children is considered normal (see Table 1) when the systolic blood pressure (SBP) and/or diastolic blood pressure (DBP) are below than 90th percentile for gender, age and height of the child in question, while hypertension is defined as SBP and/or DBP persistently higher or equal to the 95th percentile, measured at least on 3 different occasions by auscultation [8]. The middle class, SBP and / or DBP greater or equal than 90^{th} percentile but less than 95^{th} percentile is classified as high normal BP [8]. Also teenagers with BP equal or greater than $120/80\ mmHg$, regardless the correction of percentile for height, are considered to have HBP. In cases where the measured values are very different from each other, the child should be classified based on the highest one [10].

HBP, although considered an adult disease appear very often in infancy, tending to increase with age and with body mass [8]. The same author establishes a strong correlation between body mass and BP levels. The diagnostic criteria for PH have as their main reference the distribution of blood pressure (BP) in healthy children [3]. These criteria are based on the concept that the pediatric BP increases with age and with body mass. In [9, 12] is highlighted that childhood obesity is a great risk for developing HBP.

An experimental questionnaire was built and applied to evaluate the caregivers perception of Portuguese pediatric attendees of the HSM so we could get their perception about PH and possibles risks. We can found in [14] a preliminary analysis of such questionnaire. The collected information was statistically analyzed, where a descriptive analysis and a predictive model were performed. Also the fulfillment about regular measurement of pediatric BP guidelines was checked.

This article is the continuation of such work. We complete the results presented in [14]. It is performed a statistical approach estimating by GLM a logistic model for remaining questionnaire information (questions 2 and 4). Exploring the binary outcome issue, we tried to extend this approach using GLMM, specifically, mixed logistics models but the process was not successful until the moment (the algorithm did not converge). The GLMM approach is still going on.

The outline of this article is developed in four sections. In next Sect. is detailed some information about the questionnaire and is done an appointment with the adopted statistical methodo: GLM. Section 3 deals with the data. A random sample of 128 individuals were given out and 107 were correctly filled. The models estimated by GLM and the estimates of odds ratio considering some individual characteristics are presented relatively to the questions of interest. In last Sect. we discuss the results, get some conclusions and describe some ongoing work.

2 Methodology

2.1 Questionnaire

An experimental questionnaire was applied to caregivers of children aged between 3 and 18 years. It is intended to get the perception of children and young people caregivers about HBP in children.

The questionnaire is an attempt to design a "minimal" scheme to be finished quickly. It required the information described bellow.

- Some socio-demographic characteristics such as age, gender, race, residence area, level of education, profession;
- Five dichotomous questions in order to be answered (Yes/No) quickly:
    1. May the hypertension arise during the pediatric age?
    2. Is hypertension a silent disease?
    3. Should a BP measurement be made in health routine visits from 3 years age?
    4. Are there risk factors such as obesity, lack of practice physical activity or overeating that may be associated with hypertension?
    5. Is it usual to measure your child's blood pressure in health surveillance visits?

2.2 GLM Approach

In the classical linear model, a vector X with p explanatory variables $X = (X_1, X_2, \dots, X_p)$ can explain the variability of the variable of interest Y (response variable), where $Y = Z\beta + \epsilon$. Z is a specification matrix with size $n \times p$ (sometimes $Z = X$, considering an unitary vector in first column), β a parameter vector and ϵ a vector of random errors ϵ_i, independent and identical distributed to a reduced Gaussian.

The data are in the form (y_i, x_i), $i = 1, \dots, n$, as result of observation of (Y, X) n times. The response variable Y has expected value $E[Y|Z] = \mu$.

In GLM, where the model is an extension of classical model, the response variable, following an exponential family distribution [15], does not need to be Gaussian, and a transformation of expected value of response variable is related with explanatory variables. A detailed description of GLM ca be found in [15, 11].

Another extension from the classical model is the fact that the function which relates the expected value and the explanatory variables can be any differentiable function. Each Y_i has expected value $E[Y_i|x_i] = \mu_i = b'(\theta_i)$, $i = 1, \ldots, n$.

Also, it is defined a differentiable and monotone link function g that relates the random component with the systematic component of the response variable. The expected value μ_i is related with the linear predictor $\eta_i = z_i^T \beta_i$ through the relation

$$\mu_i = h(\eta_i) = h(z_i^T \beta_i), \qquad \eta_i = g(\mu_i) \tag{1}$$

where h is a differentiable function, $g = h^{-1}$ is the link function, β is a vector of parameters with size p (the same size of the number of explanatory variables) and Z is a specification vector with size p.

We can found different link functions in GLM. When the random component of response variable has a Poisson distribution, the link function is logarithmic and the model is log-linear. In particular, when the linear predictor $\eta_i = z_i^T \beta_i$ coincides with the canonical parameter θ_i, $\theta_i = \eta_i$, which implies $\theta_i = z_i^T \beta_i$, the link function is denominated canonical link function.

Usually, in presence of dichotomous data, discrete models are usually estimated by logistic or probit regression [15].

The logistic regression is widely applied in models with dichotomous data and/or proportions. It is supposed that the response variable Y verifies $Y_i \cap Bin(1, \pi_i)$, i.e.

$$f(y_i|x_i) = (\pi_i)^{y_i} (1 - \pi_i)^{1-y_i}, \qquad y_i = 0, 1, \tag{2}$$

where each individual i as associated a specification vector z_i which results from p explanatory variables $x = (x_1, x_2, \ldots, x_p)$. Knowing that

$$E[Y_i] = \pi_i, \qquad \theta_i = ln\left(\frac{\pi_i}{1-\pi_i}\right)$$
$$and \qquad \theta_i = \eta_i = z_i^T \beta_i\ , \tag{3}$$

we conclude that the logistic function is the link function g. The probability of success π_i is given by

$$\pi_i = \frac{e^{z_i^T \beta_i}}{1 + e^{z_i^T \beta_i}}. \tag{4}$$

Notice that $E[Y_i] = \pi_i = \mu_i \in [0, 1]$, another distribution function can be a good candidate to the link function inverse. One can suppose that the success probability π_i and the covariate vector are related by

$$\pi_i = \Phi(\eta_i) = \Phi(z_i^T \beta_i), \tag{5}$$

where $\phi(.)$ is the Gaussian distribution $N(0, 1)$. In this case, we obtain the link function $g(\mu_i) = \Phi^{-1}(\mu_i)$, the probit function, and the estimated model is named the probit model.

3 Empirical application

We developed an observational and prospective study, with all participants selected from the external patient Pediatric Department of HSM, from May to December 2014, having been obtained for this, Department director approval. It was taken a random sample of 128 individuals, all of them caregivers of children and/or adolescents (first-degree relatives or their legal representatives), users of the National Health System (NHS), which resorted to external consultation (general pediatrics and/or sub-specialties) for reasons not related to changes in the BP. Only 107 questionnaires were considered correctly filled.

The first step was to organize the data using descriptive statistic techniques (for example see [13]). In a preliminary data analysis and taking into account the non-quantitative nature of the involved variables, were calculated measures of association, nonparametric Spearmann correlation coefficient, nonparametric test of Friedman for paired samples.

In Fig. 1 is summarized the questionnaire answers count of the selected 107 participants. About 66.6% of participants know that the pediatric hypertension may arise during the pediatric age, 82% take into account that, frequently, hypertension is a silent disease, 61.3% know that a BP measurement should be made from 3 years age in health surveillance visits and 91.9% are aware about some risk factors that may be associated with hypertension. The question 5 was implemented to verify if regular measurement of the pediatric BP was fulfilled according to current recommendations. This issue is not fulfilled by 40.5% the participants. Initially, the questionnaire statistical inference took into consideration the questions 1, 2, 3 and 4.

We measure the homogeneity and internal consistence of questionnaire and respective validation. The alpha-Cronbach coefficient was lower than expected, giving the indication of sufficient, bad not good internal consistency. We also performed several tests (Table 2) to compare the answers of questions 1, 2, 3 ad 4: the paired T-test, McNemar's test for frequencies comparison, Crochan's Q test for binary variables comparison.

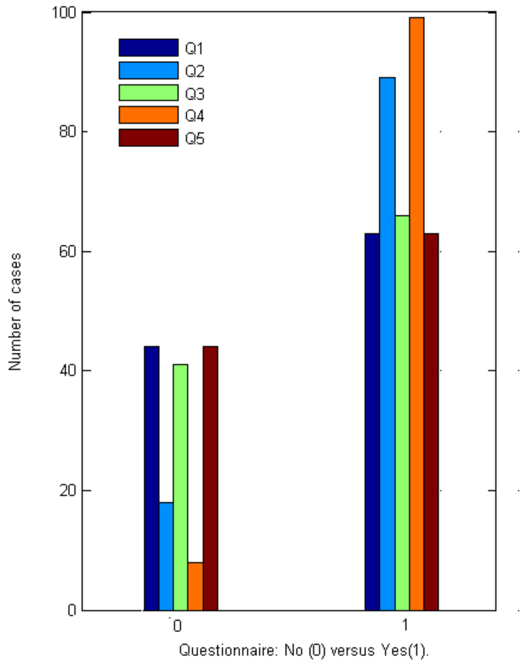

Figure 1: Questionnaire responses count. Questions 1 to 5. '1'=Yes; '0'=No.

The results in Table 2 suggest the exclusion of question 3 from our analysis, all tests show evidence that question 1 - question 3 are not significantly different. Also the non-parametric Spearmann coefficient of association is significant for the pair question 1 - question 3. The Cochran's Q tests rejects that the distributions of question 1, question 2, question 3 and question 4 are the same. Initially, we compute the association of the questionnaire answers with socio-demographic characteristics of participants. The Spearmann coefficient was significant for question 1 and question 3, question 2 and age, question 2 and profession, the Kruskal-Wallis test was significant between question 4 and age ($\eta - coefficient = 0.561$) and question 4 and level of education ($\eta - coefficient = 0.224$).

The internal consistency analysis was another key step: T-tests were carried out for paired samples and calculated the Cronbach's alpha coefficient. Due the qualitative non-ordinal variables nature, these measures were considered as indicative.

The next step of process consists in estimation, validation and selection of the models by GLM. These three steps are summarized as follows: 1- Models formulation: identify response variable distribution, select the preliminaries explanatory variables and specification matrix, select the link function g; 2- Models adjustment: estimation of model parameters, application of suitability measures of estimates; 3- Selection and validation of models: selection of variables, diagnostics, residual analysis and interpretation.

Different link functions g were used to estimate several models. The log-log, logistic and probit regression were taken into consideration. The stepwise selection was used to get the best explanatory variables in each case.

As expected, the dichotomous data conduced to discrete models. The logistic models were selected. Almost all estimated and selected models for question 1 have validation tests with $p - values$ between 0.01 and 0.10. The estimated model for question 2 has the associate validation tests with $p - values$ less than 0.05.

Some levels of residential area and/or race had few individuals. We didn't consider these individual characteristics as explanatory variables or effects. The codification of selected variables is on Table 3.

Below, we present the estimation of logistic models in several steps. Initially, the selection of the best set of explanatory variables that conduces to the "best" model, was done by stepwise, based on the $p - values$ from Wilks test likelihood ratio relatively to inclusion/exclusion of explanatory variables.The impact of a explanatory variables is measured by the obtained $p - value$: a small $p - value$ means a significant influence. By first are selected the principal effects, by last, the second order interaction.

Question 1 was discussed previously, we will detail the results associated to question 2 and 4. For each question (1,2,4) we have estimated several models: some of them just considering the principal effects (gender, age, level of education and profession), others taking into account the principal and secondary effects. The parameters estimates of such models relatively to questions 1, 2 and 4 are detailed in Table 4.

Considering question 1, model A consider the main effects and has a deviance global test with $p - value = 0.0995$, by other hand, model B considers the main and secondary effects and conduces to a deviance global test with $p - value = 0.0979$. The residual analysis evaluation of all models was done.

When all validation and selection techniques were complete, we need to be able to choose correctly between different models with close explanatory performance. The best models with similar performance conduce to estimates which are combined using a weighted mean using the $p - values$ of global test.

Question 1 (the knowledge about the possibility of occurrence of PH) and its relation with socio-demographic characteristics of participants was established and discussed previously in [14]. The odds ratios were estimated from models A and B for some

Table 1: Classification of pediatric hypertension.

Class	**SBP and/or DBP Percentile**
normal	< 90
normal-high	≥ 90 until $< 95\ mmHg$
	Teenagers: AP $> 120/80\ mmHg$, including percentile < 90
HBP- class 1	$95 \leq$ Class $1 \leq 99 + 5\ mmHg$
HBP - class 2	$> 99 + 5\ mmHg$

Table 2: Confidence Interval and paired T-test for question mean difference. McNemar's test for frequencies comparison. Crochan's Q test for binary variables comparison.

Pair	**C.I. (95%)**	t-test **p-value**	McNemar's test **p-value**	Crochan's Q test **p-value**
Q1-Q2	$(0.056, 0.261)$	0.003	0.005	0.003
Q1-Q3	$(-0.17, 0.058)$	0.333	0.429	0.343
Q1-Q4	$(0.157, 0.348)$	0.000	0.000	0.000
Q2-Q3	$(-0.33, -0.10)$	0.000	0.001	0.001
Q2-Q4	$(0.017, 0.170)$	0.018	0.025	0.016
Q3-Q4	$(0.202, 0.415)$	0.000	0.000	0.000

Table 3: Codification of variables.

Variables	**Description**	**Codification**	**Variables**	**Description**	**Codification**
Y_1	Question 1	1=Yes 0=No	x_2	Profession	11=Other 10=Without qualification
Y_2	Question 2	1=Yes 0=No			. . . . . .
Y_3	Question 3	1=Yes 0=No	x_3	age	1=High qualification $'1' = ([18, 29])$
Y_4	Question 4	1=Yes 0=No			$'2' = ([30, 49])$ $'3' = (\geq 50)$
x_1	Education level	1=primary 2=High School 3=University	x_4	Gender	0=Female 1=Male

values of explanatory variables. It was concluded that the knowledge about the possibility of the occurrence of hypertension in pediatric age is lower for a man than for women, when they have similar ages, education and profession; when the qualification decreases 5 levels, the odd ratio is clearly reduced from values greater than one for values less than one; when the level of education increases one degree, the youngest individuals have greater chance to know about the occurrence of pediatric hypertension.

Relatively to question 2 (appearance of the disease without notice), the estimated model C has a deviance global test with $p-value = 0.012$ (see Table 4). Notice that there is no principal effect in the model C, the interaction between profession and education level is the only significant (secondary interaction) effect. The estimated model D has a deviance global test $p-value = 0.015$ and has a unique principal effect, the profession.

The deviance residuals of model C and D, for the 107 individuals are represented in Fig. 2; on left we have the deviance residuals of model C, on right, are represented the deviance residuals of model D.

We describe some relations between the idea of hypertension as a silent disease and socio-demographic characteristics of participants. From model D, when we consider a man or woman with a median qualification profession (between oldest or youngest people) there is more than double chance to know that pediatric hypertension is a silent disease. We get $\Psi_{Q2,prof_D} = 0.969$ if we increase one unit of profession level value. It means that, if the level of qualification decreases, the odds ratio decreases slightly.

When we consider question 4 (knowledge about several risks), the estimated model E has no significant interaction and keeps only one principal effect, the education level, for a deviance global test of $p-value = 0.051$. The residuals analysis was also performed. The odds ratio for a person with high school level education (any gender, age or profession) we get $\Psi_{Q4,educ_E} = 1.047$. People with medium level of education have more than double chance to know that there exist several risk factors which contribute to HBP. This odds ratio does not change when the education level increases or decreases one unit.

Exploring the binary outcome issue, we extended this approach using GLMM, specifically, mixed logistic models. Considering question 1, we have tested several random effects such as gender, education level (slip in two levels: medium-high and low). The process was unsuccessful, the algorithm did not converge. We are still working on this approach.

4 Discussion and final remarks

Pediatric arterial hypertension is a highly prevalent, silent, and perhaps because of this, under-diagnosed in most cases, with multiple repercussions on the health of children and adults. It is imperative that health professionals and family members be aware of HBP existence, the negative consequences associated with it, the risk factors and it prevention. Its expected the knowledge about this pathology increases with the level of education of the family members.

A simple questionnaire was built with the aim of easy and quick answers. We can establish that, despite not rejecting the questionnaire in terms of internal consistency, there is some evidence that it needs some improvement, for example, reformulation of some questions. The estimated models by GLM have reasonable explication of the data. Exploring the binary outcome issue, we extend this approach using GLMM, specifically, mixed logistics models. The process was unsuccessful, the algorithm did not converge.

The use of some additional statistical methods, similarly to [5], correspondence analysis or other multivariate method, can give another extra contribution to the data variability explanation. Taking into consideration the idea of simplicity, our ongoing work consists in the application of another statistical multivariate techniques [2] to complete this preliminary study.

Accordingly with the questionnaire, caregivers know that there is pediatric hypertension, is silent in most cases and that there exist several risk factors which contribute to it. It is possible to conclude that age , educational level and profession influence the knowledge that relatives have about this pathology.The individuals from higher age groups are those who Know the possibility of this pathology be able to arise quietly. Higher levels of education are associated with greater knowledge about the existence of hypertension risk factors. Finally, the higher level of profession, greater the knowledge about the existence recommendation for the measurement of BP in children from 3 years of age.

The results are promising, but there are some details which need to be extended. To complete this work, was applied recently a new questionnaire, with the same objective but with different sample, structure and questions:

- Socio-demographic characterization of the population of participants, which represent the variables under study);

Table 4: Logistic models-parameter estimates. Questions 1 ($\hat{\beta}_i$ and $\hat{\alpha}_i$); Question 2($\hat{\gamma}_i$ and $\hat{\delta}_i$); Question 4 ($\hat{\epsilon}_i$).

Effect	**Parameter** Q1 (A)	**Estimate** Q1 (A)	**Parameter** Q1 (B)	**Estimate** Q1 (B)	**Parameter** Q2 (C)	**Estimate** Q2 (C)	**Parameter** Q2 (D)	**Estimate** Q2 (D)	**Parameter** Q4 (E)	**Estimate** Q4 (E)
Intercept	$\hat{\beta}_0$	0.634	$\hat{\alpha}_0$	-9.915	$\hat{\gamma}_0$	1.013	$\hat{\delta}_0$	0.976	$\hat{\epsilon}_0$	0.769
x_1	$\hat{\beta}_1$	-0.093	$\hat{\alpha}_1$	3.507					$\hat{\epsilon}_1$	0.072
x_2	$\hat{\beta}_2$	-0.051	$\hat{\alpha}_2$	0.612			$\hat{\delta}_2$	-0.031		
x_3	$\hat{\beta}_3$	-0.151	$\hat{\alpha}_3$	4.888						
x_4	$\hat{\beta}_4$	-0.769	$\hat{\alpha}_4$	-0.363						
$x_1 * x_2$					$\hat{\gamma}_{13}$	0.019				
$x_1 * x_3$			$\hat{\alpha}_{13}$	-1.764						
$x_2 * x_3$			$\hat{\alpha}_{23}$	-0.297						
$x_2 * x_4$			$\hat{\alpha}_{24}$	-0.091						
$p-value$		0.0995		0.0979		0.012		0.015		0.051

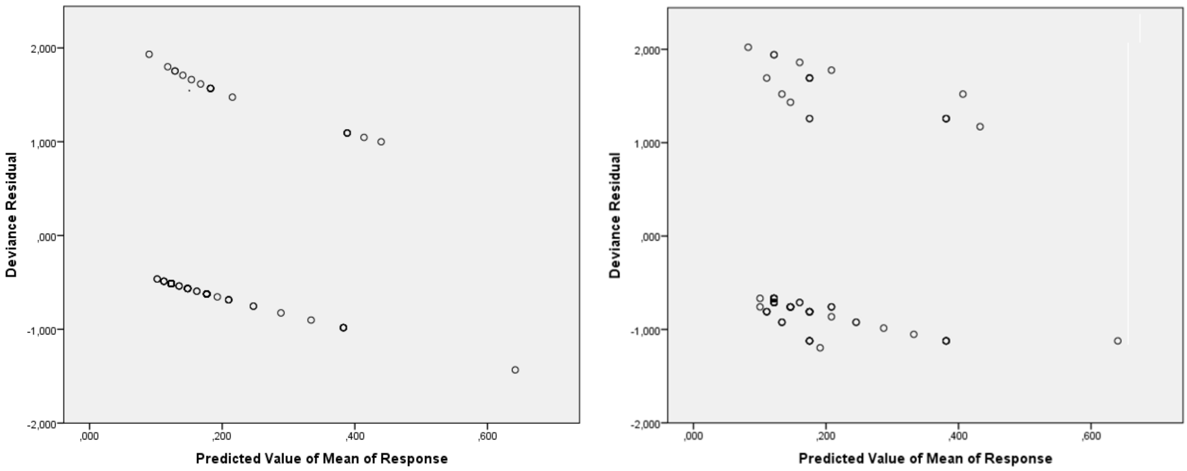

Figure 2: Question 2. Deviance residuals versus fitted values. Models C (on left) and B (on right).

- Fifteen theoretical questions about pediatric hypertension, on which the evaluation of the object of the study is based: Characterization of the knowledge of parents and / or caregivers on Pediatric Hypertension;
- One final question regarding the periodicity of evaluation of BP in periodic health surveillance visits of the pediatric relatives of the participants;
- The questionnaire was sent to 182 participants via e-mail (containing the link to the online form), previously sent by the respective class directors of the ten schools of the Group.

For improved survival and quality of life of hypertensive patients an early diagnosis is necessary and an early and appropriate therapeutic intervention. Prevention is possible by controlling known risk factors and modifiable. Prevention requires a higher degree of awareness and possibly the implementation of awareness campaigns promoting regular assessment of arterial pressure in children.

Summarizing, GLM models [15, 11] were developed to relate the questionnaire questions and sociodemographic explanatory variables. To improve this statistical approach it is ongoing a statistical multivariate analysis [2] using all relevant available data. The remaining ongoing work using other statistical methods, discriminated and analyzed will be found in a future extended version of this article. We intend to compare the two questionnaires submitted to different populations with distinct structures and methods of response.

We also intend to continue this study in a different scale, identifying the risks of the disease and its relationship with inherited and acquired changes in the state of potential patients. Following the objective to get a picture of pediatric HTA in Portugal, under the aim of a project proposed by Portuguese Pediatric association, we are collecting data at national level (simple size=5500).

Acknowledgements: This work was supported by Portuguese funds through the *Center of Naval Research* (CINAV), Naval Academy, Portuguese Navy, Portugal and the *Center for Computational and Stochastic Mathematics* (CEMAT), *The Portuguese Foundation for Science and Technology* (FCT), University of Lisbon, Portugal, project UID/Multi/04621/2013.

References:

[1] D. Ahern and E. Dixon, *Pediatric Hypertension. A Growing Problem*, Prim. Care Clin. Office Pract. **42(1)**, 143–150 (2015).

[2] T.W. Anderson, *An Introduction to Multivariate Analysis* (Jonh Wiley & Sons, New York,2003).

[3] H. Andrade et al., *Hipertensão arterial sistémica em idade pediátrica*, Revista Portuguesa de Cardiologia **29(3)**, 413–432 (2010).

[4] P.P. Bassareo and G. Mercuro, G., *Pediatric hypertension: An update on a burning problem* World J. Cardiol. **6(5)**, 253–259 (2014).

[5] M.G. Costa et al., *Conhecimento dos pais sobre alimentação: construção e validação de um questionário de alimentação infantil*, Revista Enfermagem Referência **3(6)**, 55–68 (2012).

[6] J. Flynn, *Pediatric hypertension: recent trends and accomplishments, future challenges*, Am. J. Hypertens **21**, 605-612 (2008).

[7] M.L. Hansen and P. Gunn, *Underdiagnosis of hypertension in children and adolescents*, Journal of American Medical Association **298**, 874–879 (2007).

[8] E. Lurbe and R.F. Cifkovac, *Management of high blood pressure in children and adolescents: recommendations of the European Society oh Hypertension*, Journal of Hypertension **27**, 1719-1742 (2009).

[9] P. Muntner and J. He, *Trends in blood pressure among children and adolescents* Journal of the American Medical Association **291**, 2107–2113 (2004).

[10] National High Blood Pressure Education Program Working Group on High Blood Pressure in Children And Adolescents, *The fourth report on the diagnosis, evaluation and treatment oh high blood pressure in children and adolescents*, Pediatrics **114**, 555-576 (2004).

[11] J.A. Nelder and R.W.M. Wedderburn, *Linear Models*, Journal of Royal Statistical Society **35**, 370–384 (1972).

[12] S. Stabouli and V. Kotsis, *Adolescent obesity is associated with high ambulatory blood pressure and increased carotid intimal-medial thickness*, Journal of Pediatrics **147**, 651-656 (2005).

[13] A.C. Tamhane and D.D. Dunlop, *Statistics and Data Analysis: from Elementary to Intermediate* (Prentice Hall, New Jersey, 2000).

[14] M.F. Teodoro and C. Simão, *Perception about Pediatric Hypertension*, Journal of Computational and Applied Mathematics **312**, 209-215 (2017).

[15] M.A. Turkman and G. Silva, *Modelos Lineares Generalizados da teoria a prática* (Sociedade Portuguesa de Estatística, Lisboa, 2000).