CLUSTER ANALYSIS FOR CLASSIFICATION OF PATIENTS SUBMITTED TO FOBI-CAPELLA BARIATRIC SURGERY

Análise de agrupamentos para a classificação de pacientes submetidos à cirurgia bariátrica Fobi-Capella

Fernando **FREI**[1], Bráulio de Souza **LESSA**[2], Judith Cristina Gouveia **NOGUEIRA**[2],
Renata **ZOPELLO**[2], Samuel ribeiro da **SILVA**[2], Viviane A. Moreira **LESSA**[2]

From the Departamento de Ciências Biológicas FCLassis, Universidade Estadual Paulista "Júlio Mesquita Filho", UNESP, Assis, SP and ²Serviço de Cirurgia do Aparelho Digestivo, Hospital Heliópolis, São Paulo, SP, Brazil.

ABSTRACT - *Background* - Individuals classified as morbidly obese require effective treatments to promote weight reduction. Due to the inefficacy of conservative treatments and current medications, bariatric surgery (gastric bypass) has been indicated for this group of individuals. *Aim* - To identify subgroups of obese patients in order to find patterns that hinder weight loss and recovery after the operation. *Method* - Variables from the medical records of 50 patients were analyzed at two different times: before bariatric surgery and six months after the surgery, utilizing statistical cluster analysis. *Results* - The analysis led to dividing the patients into two groups. The body mass index - BMI variable did not influence the classification of the patients. The group in worse metabolic conditions was not necessarily composed of heavier subjects, but the cluster analysis grouped the patients according to four indicators of metabolic syndrome. Metabolic alterations and blood pressure were shown to have a clear correlation to obesity, relation to BMI was not confirmed by this study. *Conclusions* - Taking measures to manage metabolic syndrome and not focusing solely on weight loss could improve recovery after bariatric surgery, considering that the BMI was reduced in all patients and was not the distinguishing factor for the postoperative period.

Correspondence:
Fernando Frei
E-mail: ffrei@assis.unesp.br

RESUMO - *Racional* - Indivíduos classificados como obesos mórbidos necessitam de tratamentos eficientes para promover a redução de peso. Em consequência da ineficácia dos tratamentos conservadores e medicamentos atuais, a operação de Fobi-Capella (gastric bypass) tem sido indicada para esse grupo de indivíduos. *Objetivo* - Identificar subgrupos de pacientes obesos com a finalidade de encontrar padrões que dificultem a perda de peso e a recuperação dele após essa operação. *Método* - Foram analisadas variáveis dos prontuários de 50 pacientes em dois momentos distintos: antes da operação bariátrica e após seis meses utilizando a metodologia estatística de Análise de Agrupamentos. *Resultados* - A análise propiciou a divisão dos pacientes em dois grandes grupos. A variável IMC não influenciou na divisão dos pacientes. O grupo em piores condições metabólicas não foi necessariamente formado pelos indivíduos mais obesos, porém a Análise de Agrupamento associou os pacientes de acordo com quatro indicativos da síndrome metabólica. Houve clara relação entre as alterações metabólicas e de pressão arterial com a obesidade, porém neste estudo não se verificou dependência direta em relação ao IMC. *Conclusões* - As medidas que poderiam incrementar a recuperação após cirurgia bariátrica visam o controle da síndrome metabólica e não apenas a perda de peso, visto que o IMC foi reduzido em todos os pacientes e não foi o fator diferencial do pós-operatório.

INTRODUCTION

Obesity can be defined as a condition of abnormal or excessive accumulation of fat that may present a health risk[26]. Several comorbidities have been associated, among these, type 2 diabetes, hypertension and cardiac diseases stand out [10,25].
The simplest method of assessing weight excess and obesity is

through the anthropometric measurement of weight and height, known as BMI (body mass index). This measure is useful because it correlates with body fat, morbidity and mortality[2]. Thus, a BMI between 25.0 and 29.9 kg/m2 is defined as overweight and above 30.0 kg/m2 as obesity[24,17].

During the 1940s, due to the sixth revision of the International Classification of Diseases - ICD, the World Health Organization had already incorporated obesity as a health problem and this status has been maintained during the CID13 updating process.

According to studies conducted in 2010 by the International Obesity Taskforce, one billion adults are overweight and another 475 million are obese worldwide. Globally it is estimated that up to 200 million school-age children are overweight or obese[12]. According to the Household Budget Survey conducted by the Brazilian Institute of Geography and Statistics, in 2002-2003, obesity affected 11.1% of adult Brazilians - 8,9% of all men and 13,1% of all women. In absolute numbers, obesity affected two million Brazilians[11].

Within the category of obese, there are individuals with a BMI higher than 40.0 kg/m[2], classified as morbidly obese. Such individuals should be treated as having a health problem that is life-threatening and that reduces quality of life[27].

Conservative treatments - low-calorie diets, behavioral therapy, and exercise - should be the first to be adopted. However, only 5-10% of individuals manage to lose significant amounts of weight over a period greater than a few years[3]. Other strategies combined with conservative treatments were tested, such as the use of medications that have demonstrated effective weight loss (between 5-10%). According to Zilberstein et al,[27] these drugs present acceptable side effects and have been indicated for prolonged periods in both the induction and maintenance of weight loss; however, it was observed that they cannot maintain weight loss, specially in patients with morbid obesity[3].

Due to the inefficacy of conservative treatments and current medications, bariatric surgery has been indicated for patients who were morbidly obese or those who were obese and presented comorbidities such as systemic arterial hypertension and diabetes mellitus[6,14]. Among various surgical modalities, the one that presents the best results is the operation called Fobi-Capella (gastric bypass), which combines gastric restriction with induction of malabsorption[27].

The objective of this study was to identify subgroups of obese patients in order to find the patterns that hinder their weight loss and recovery after bariatric surgery.

METHOD

This study was approved by the Committee on Ethics Research of Heliópolis Hospital (CEP Registry No. 785/2011).

Were analyzed data from medical records of patients treated at Heliópolis Hospital, located in São Paulo, SP, Brazil, between 2001 to 2010, at two different times: before bariatric surgery and six months after the surgery.

Fifty patients were selected out of 335 who underwent bariatric surgery in the specified period. The remaining records could not be included in the sample because they lacked one or more of the variables selected for the study. The 50-patient sample represented 15% of the records. Unlike inferential statistics techniques, in which the size of the sample is calculated in terms of significance levels and errors adopted, cluster analysis - an exploratory multivariate statistical technique -, does not define any rules for the minimum size required for its use. However, practical rules indicate that the sample size should be at least greater than the number of variables, or else calculated as 2xk, where k represents the number of variables being studied. Both cases apply here. The variables were: BMI, low-density proteins – LDL, high density proteins – HDL, very low density lipoproteins – VLDL, hemoglobin – Hb, platelets, leukocytes, triglycerides, glucose, and bilirubin.

Patients who underwent the surgery were classified through cluster analysis, a generic term encompassing the extensive variety of statistical methods that seek to develop criteria for grouping objects (human beings, animals, plants, cities). These are multivariate statistical techniques with exploratory overtones. The results of this set of techniques can contribute to: 1) defining a formal classification scheme, 2) suggesting a set of rules to classify new patients into new classes for diagnostic purposes; 3) suggesting statistical models to describe populations; 4) finding patients who may represent groups or classes, and identifying subgroups in the population at risk in relation to aggravation factors, in order to target preventive strategies according to geographical, social and health characteristics. Thus, the results obtained by cluster analysis can help medical and managerial teams make decisions regarding diagnoses, interventions, pre actions, and educational campaigns, among other measures[23]. Considering that the metrics differ among the variables being studied, standardization has been adopted using the standard normal function z and the Euclidean distance for obtaining dissimilarities between patients.

Measures of central tendency, variability and

confidence interval for the average were used. The variance analysis for a factor was used to determine the statistical significance between possible groups in various variables and post hoc Bonferroni test. Assumptions of normality and homoscedasticity were verified by the Shapiro-Wilk and Leneve test, respectively (0.05 significance level).

The statistical analysis employed the computer application SPSS® version 19.

RESULTS

The compound sample was composed of 48 women and two men between 21 and 61 years old.

TABLE 1 - Results of the variables analyzed at pre and postoperative period

	Preoperative		Postoperative		
	$\overline{x}$	S	$\overline{x}$	S	IC 95% (difference between means)
BMI	47.37	6.45	33.31	6.14	(11.55 to 16.57)
LDL	134.38	46.36	92.74	24.05	(26.79 to 56.48)
HDL	46.36	10.87	49.32	13.19	(-7.77 to 1.85)
VLDL	27.72	13.45	18.84	8.13	(4.40 to 13.35)
Hemoglobin	13.51	1.14	13.07	1.05	(0.01 to 0.87)
Platelets	290.96	82.78	260.82	63.06	(0.79 to 59.48)
Leukocytes	8.24	2.23	6.74	2.04	(0.63 to 2.34)
Triglycerides	153.84	85.70	90.8	35.21	(37.38 to 90.05)
Glucose	105.76	24.65	83.72	11.08	(14.35 to 29.72)
Bilirubin	0.50	0.25	0.63	0.29	(-0.23 to - 0.01)

Cluster analysis using agglomerative hierarchical algorithms (Figure 1), allowed for the division of patients in the preoperative state into two groups. The variables that presented statistical differences for each group were: glucose, LDL, VLDL, triglycerides, and bilirubin. For the postoperative state, the variables that showed differences were glucose, triglycerides and VLDL.

FIGURE 1 - Cluster analysis (Ward method): preoperative (A) e postoperative (B)

The cluster analysis organized the patients into two distinct groups, in which the discriminatory variable was not the BMI value; a comparing results of Gr 1 Pre to Gr 2 Pre shows that the degree of obesity did not contribute to the formation of the groups, the same occurring with the postoperative groups (Figure 2A).
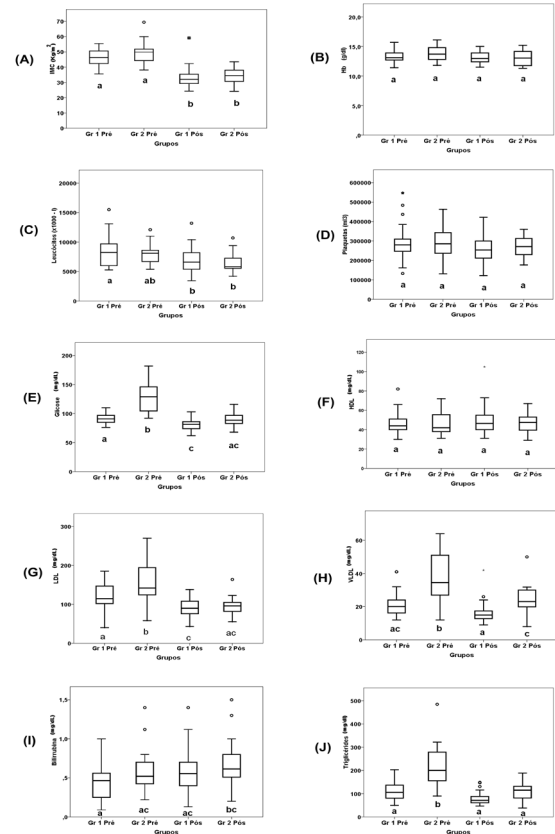

FIGURE 2 - Anova – Comparison of pre and postoperative groups: groups with the same letter indicate statistically insignificant averages (p > 0.05)
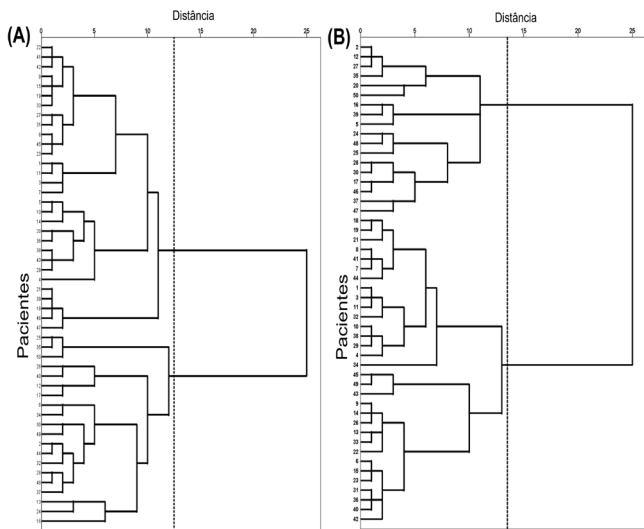
The Calinski-Harabasz and silhouette index confirmed that the best solution for the set of patients was their division into two groups.

Subsequently, the silhouette index was used for reclassifications through the analysis of the index for each subject. Thus, patients who were classified into one of two groups had their individual index evaluated. For those with negative values, a change of group was performed and after it an evaluation of the silhouette index, and so forth.

As a final result, two groups emerged from the analysis; "Group 1" had 30 patients, while "Group 2" had 20 patients in preoperative condition (denominated Gr 1 Pre and Gr 2 Pre). This same situation applied to the postoperative groups (denominated Gr 1 Pos and Gr 2 Pos). Postoperatively, four people who originally belonged to Group 2 were incorporated into Group 1. Group 2 was otherwise unchanged.

DISCUSSION

One of the major difficulties in cluster analysis is determining the number of groups. Moreover, different groupings arise when using different algorithms. Therefore, it is recommended that various algorithms be used. If the results present similar substructures, a natural partition can be obtained; otherwise, it is unlikely that the data will present distinct natural groups. For this, maximum distance algorithms are used (complete linkage), average distance (average linkage) and Ward's method.

Another strategy for determining the number of groups, in addition to the algorithms, is the use of indices for measuring the quality of the clusters obtained. In this present work, two indices were used, the Calinski-Harabasz index - also known as Pseudo F -, which performs the comparison between the sum of the quadratic distance within groups and compares it with the sum of the quadratic distance between the groups, and the silhouette index, which defines the quality of the clusters based on the proximity between the objects of a particular group and the proximity of these objects to the nearest group. In order to obtain the silhouette index, the groups obtained by applying one of the clustering algorithms and distance matrix between the objects should be used[21]. The result of the silhouette index value varies between -1 and 1. Values closer to 1 mean a better allocation of the object in the group; while values closer to -1 mean worse allocation, considering that the object is, on average, closer to the elements of the neighboring group. This way, a relocation of the object to its closest neighbor and verification of the index value can be done, improving the classification without changing the optimal number of groups[20].

There are a variety of methods and algorithms to perform the grouping of patients after assessing similarity or dissimilarity. In this study, the hierarchical clustering method was used, which has shown to be a useful tool in the discovery of substructures associated with a particular set of data. The hierarchical clustering algorithms belong to the recognition systems of unsupervised systems and are characterized by not having predetermined numbers of groups.

The reduction in BMI, glucose, total cholesterol, LDL and VLDL is expected, and is caused by the operation. Food intake is decreased, as are glucose and cholesterol absorption, and hormonal mechanisms are activated by the fact that food passes more quickly through the digestive tract. An unwanted result, which can be characterized as a "side effect", is a reduction of the number of erythrocytes, leukocytes and platelets. This occurs due to the lower absorption rate of iron and vitamins in the digestive tract. The downside of absorbing less fat and glucose is that one also absorbs less essential nutrients. Therefore, all postoperative patients must take multivitamins for the rest of their lives. This is why medical professionals vigilantly monitor these postoperative patients. In addition, the majority of Heliópolis Hospital patients that undergo bariatric surgery are female, and thus need more iron in their diet.

It is a fact that all the individuals analyzed in the study were morbidly obese; however, their BMI ranged from 35 to 69. The sample included grade 3 obese and superobese patients, whose BMI was higher than 50; however, these characteristics were not among those responsible for the discrimination of patients. The discriminatory variables responsible for uniting patients in Group 1 and in Group 2 were: fasting glucose (Figure 2E), triglycerides (Figure 2J), LDL cholesterol fractions (Figure 2G) e VLDL (Figure 2H).

HDL cholesterol (Figure 2F), hemoglobin (Figure 2B), leukocytes (Figure 2C), bilirubin (Figure 2I) and platelets (Figure 2D), were not discriminatory for the formation of Groups 1 and 2. Therefore, Group 2 is the one which presents the highest measures of glucose, triglycerides, VLDL and LDL, independent of body weight. The group in worse metabolic condition is not necessarily formed by the heavier subjects, however cluster analysis grouped patients according to four indicators of metabolic syndrome, a term used to designate a set of organic changes related to a higher risk of cardiovascular disease, according to the Third National Cholesterol Education Program (NCEP III)[1]. According to these criteria, metabolic syndrome includes obesity, with considerable abdominal visceral fat, triglycerides greater than or equal to 150 mg/dl, HDL lower than 40 mg/dl for males and lower than 50 mg/dl for females, fasting glucose greater than or equal to 110 mg/dl and arterial pressure greater than or equal to 130 mmHg x 85 mmHg. There is a clear relation between metabolic and arterial-pressure changes and obesity; however, this study did not find a direct relation to BMI. Thus, patients with a lower BMI could have a more severe metabolic syndrome, and the superobese could show no clinical or laboratorial evidence of the syndrome. This does not mean that obesity is not a risk factor for such changes, but shows that it is not the only element involved [19,22,7].

Cluster analysis presents results that are, from a biological standpoint, coherent, considering that Groups 1 and 2 were defined according to metabolic syndrome, which can be characterized even if not all the changes that define it are present[19]. Group 2 presented the worst metabolic parameters (Figures 2E, 2H and 2J), maintaining itself as worse than Group 1 even in the six-month period after the operation, when both groups had already lost weight.

In relation to glycemia, Group 2 presented, in

the postoperative period, levels similar to those of Group 1 before the operation (Figure 2E). It is clear that there has been important improvement for the subjects, but the results show an interesting pattern: 80% of the patients of Group 2 remained in Group 2, discriminated by the same variables that, nevertheless, decreased in both groups.

The literature review shows that bariatric surgery is indicated not only for weight loss, but for the treatment of type 2 diabetes mellitus (non-insulin dependent)[9,4]. There is no doubt that obesity triggers metabolic syndrome in many individuals[19] and that weight loss is critical for its control [4].

The results obtained here do not contradict the benefits of the operation, nor do they bring into question its efficacy in the treatment of metabolic syndrome. Even considering the statistical differences, all patients have had significant improvement in all aspects studied.

Another important fact to note is the six-month length of the postoperative observation period. It is likely that longer-term evaluation would show another statistical profile, considering that patients continue to lose weight and could present other changes in their metabolic profiles. According to Lemanu et al [16], literature surveys on increased postoperative recovery have demonstrated that there are few studies investigating this specific issue, although there is extensive literature on recovery from non-bariatric operations [16].

It should be noted, as a limitation of the study, that the sample includes a preponderance of female patients; therefore, the results cannot be generalized for men. However, observing these results and comparing them to the study of Lemanu et al, [16] it can be suggested that the early identification of patients discriminated as Group 2, and intensive preoperative measures in order to minimize the effects of metabolic syndrome may both be useful in improving postoperative recovery, especially in the short and medium term. Furthermore, there were no characteristics found that hindered weight loss in both groups (Figure 2A); thus, this aspect was not included among the measures discussed as potentially improving postoperative recovery [18].

Regarding the issue of the scarcity of short-term studies on measures aimed to improve the postoperative period, the cluster analysis provided information on the clinical evolution of patients and, consequently, on their prognosis in terms of improvement of the metabolic syndrome. As previously discussed, the recovery profile for patients is known in the long term; this study, based on data from the six months immediately after the operation, may add data about the time period required for complete control of the metabolic syndrome, that is, how long Group 2 would need to reach levels that were considered normal. In this regard the grouping may qualify the prognosis: both groups showed improvement, but Group 1, in the six months immediately following the operation, approached more of the desired metabolic parameters.

Based upon this study, cluster analysis could again be performed on the same subjects, two years after the operation, as observations at different periods would help reveal the chronology of metabolic change in these patients, indicating interventions that could optimize postoperative recovery

CONCLUSION

The measures taken to enhance bariatric surgery recovery should include control of metabolic syndrome and not focus only on weight loss, considering that the BMI of all patients was reduced and was not the distinguishing factor in the postoperative period.

ACKNOWLEDGEMENTS

REFERENCES

1. American Heart Association. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report. [acessado em 19/04/208.] Disponível em:http://circ.ahajournals.org/content/106/25/3143.full.pdf.
2. Aronne LJ. Classification of Obesity and Assessmentof Obesity-Related Health Risks. Obesity Research. 2002 December; 10 Suppl 2:S105-115.
3. Balsiger BM, Murr MM, Poggio JL, Sarr MG. Bariatric Surgery: Surgery for Weight Control in Patients With Morbid Obesity. Medical Clinics of North America 2000; 84(2): 477 -489.
4. Buchwald H, Avidor Y, Braunwald E, Jensen MD, Pories W, Fahrbach K, Schoelles K. Bariatric Surgery. A Systematic Review and Meta-analysis. JAMA. 2004 October; 292 (14): 1724-37.
5. Everitt BS. Cluster Analysis. 3a edição, John Wiley & Son. New York; 1993.
6. Ferraz EM, Arruda PCL, Bacelar TS, Ferraz AAB, Albuquerque AC, Leão CS. Tratamento Cirúrgico da Obesidade Mórbida Rev. Col. Bras. Cir. 2003; 30(2):98-105.
7. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. JAMA. 2002;287:356-9.
8. Frei, F. Introdução à Análise de Agrupamentos: Teoria e Prática. Editora Fundação UNESP. São Paulo; 2006.
9. Geloneze B, Pareja JC. Cirurgia bariátrica cura a síndrome metabólica? Arq Bras Endocrinol Metab. 2006; 50(2): 400-7.
10. Hurt RT, Kulisek C, Buchanan LA, McClave SA. The Obesity Epidemic: Challenges, Health Initiatives, and Implications for Gastroenterologists. Gastroenterology & Hepatol. 2010 December; 6(8):780–792.
11. Instituto Brasileiro de Geografia e Estatística. Pesquisa de Orçamentos Familiares 2002-2003: análise da disponibilidade domiciliar de alimentos e do estado nutricional no Brasil. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 2004.
12. International Obesity Taskforce – IOT. The Global Epidemic. [Acessado em 11/04/208]Disponível em: http://www.iaso.org/iotf/obesity/obesitytheglobalepidemic/.

13. James WPT. WHO recognition of the global obesity epidemic. International Journal of Obesity.2008;32:S80–86.

14. Karmali S, Stoklossa CJ, Sharma A, Stadnyk J, Christiansen S, Cottreau D, Birch DW. Bariatric surgery: A primer. Can Fam Physician. 2010 Sep;56(9):873-9.

15. Kaufman L, Rousseeuw PJ. Finding groups in data: An introduction to cluster analysis. New York. John Wiley & Sons; 1990.

16. Lemanu DP, Srinivasa S, Singh PP, Johannsen S, Maccormick AD, Hill AG. Optimizing perioperative care in bariatric surgery patients. Obes Surg. 2012, Jun;22(6):979-90.

17. Lyznicki JM, Young DC, Riggs JA, Davis RM. Obesity: Assessment and anagement in primary care. Am Fam Pract Physician. 2001;63:2185–2.

18. Marcelino LF, Patrício ZM. A complexidade da obesidade e o processo de viver após a cirurgia bariátrica: uma questão de saúde coletiva. Ciênc. Saúde Coletiva, Dez 2011, vol.16, no.12, p.4767-4776.

19. Ramos MEM, Ramos AM, Moriguchi EH, Ramos SS, Zimmer PM. Prevalência de sobrepeso e obesidade e sua associação com a síndrome metabólica em uma amostra de idosos. Revista da AMRIGS, Porto Alegre. 2006;50 (4): 307-38.

20. Rendón E, Abundez I, Arizmendi A, Quiroz EM. Internal versus External cluster validation indexes. International Journal Of Computers And Communications. 2011; 1 (5): 27-34.

21. Rousseeuw, PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 1987; 20: 53–65.

22. Saely CH, Aczel S, Marte T, LANGER P, HOEFLE G, DREXEL H. The metabolic syndrome, insulin resistance, and cardiovascular risk in diabetic and nondiabetic patients. J Clin Endocrinol Metab. 2005;90(10):5698-703.

23. Sneath PHA, Sokal RR. Numerical taxonomy. W.H. Freeman and Company. San Francisco; 1973.

24. World Health Organization (WHO). BMI classification. [Acessado em 11/01/208]. Disponível em: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.

25. World Health Organization (WHO). Diet, nutrition andprevention of chronic diseases. Report of a Joint WHO/FAO Expert Consultation. Geneva: WHO; 2003. WHO Technical Report Series, 916.

26. World Health Organization (WHO). Obesity and overweight. [acessado em 08/01/2011]. Disponível em: http://www.who.int/topics/obesity/en/.

27. Zilberstein B, Neto MG, Ramos AC. O papel da Cirurgia no Tratamento da Obesidade. Revista Brasileira de Medicina. São Paulo. 2002 abr; 59(04): 258-264.