

Karina Yuri Harada, Julieta Gonçalves Silva
Simone Schenkman, Eliana Tiemi Hayama
Francisco Roberto Gonçalves Santos
Mariângela Cainelli de Oliveira Prado
Regina Helena Petroni Pontes

Delimitation of homogeneous regions in the UNIFESP/EPM healthcare center coverage area based on sociodemographic indicators

Centro de Saúde Escola, Departamento de Medicina Preventiva, Universidade Federal de São Paulo/Escola Paulista de Medicina, São Paulo, Brazil

ABSTRACT

INTRODUCTION

Context: The drawing up of adequate Public Health action planning to address the true needs of the population would increase the chances of effectiveness and decrease unnecessary expenses.

Objective: To identify homogeneous regions in the UNIFESP/EPM healthcare center (HCC) coverage area based on sociodemographic indicators and to relate them to causes of deaths in 1995.

Design: Secondary data analysis.

Setting: HCC coverage area; primary care.

Sample: Sociodemographic indicators were obtained from special tabulations of the Demographic Census of 1991.

Main measures: Proportion of children and elderly in the population; family providers' education level (maximum: >15 years, minimum: <1 year) and income level (maximum: > 20 minimum wages, minimum: <1 minimum wage); proportional mortality distribution.

Results: The maximum income permitted the construction of four homogeneous regions, according to income ranking. Although the proportion of children and of elderly did not vary significantly among the regions, minimum income and education showed a statistically significant ($p < 0.05$) difference between the first region (least affluent) and the others. A clear trend of increasing maximum education was observed across the regions. Mortality also differed in the first region, with deaths generated by possibly preventable infections.

Conclusion: The inequalities observed may contribute to primary health prevention.

Key words: Censuses. Health status indicators. Mortality. Public Health. Regional Health Planning.

The Healthcare Center (HCC) of the area where the Escola Paulista de Medicina (EPM) is located is staffed by professionals qualified for promoting actions aimed at health and disease prevention, with the possibility of effective referral and counter-referral mechanisms with the São Paulo Hospital, linked to the Federal University of São Paulo (UNIFESP). However, there is no recent scientific literature studying the resident population in its healthcare coverage area. The drawing up of adequate Public Health action planning to address the true needs of the population would increase the chances of effectiveness and decrease unnecessary expenses. Knowledge of the population is thus of fundamental importance and forms the objective of this study.

The epidemiologic study performed in the city of São Paulo in 1984 by Ramos et al¹ was of great importance, conducted on a community-based sample of people over 60 years of age and stratified according to the socio-economic status of persons responsible for them. Some subsequent studies were conducted in an attempt to characterize the population of the city of São Paulo, most of them with participation of the IBGE

(Brazilian Institute of Geography and Statistics) and the SEADE² (State System for Data Analysis) Foundations.

This study sought to identify homogeneous regions within the EPM HCC coverage area based on sociodemographic indicators. It also aimed to relate the causes of death in 1995 with the delimitation of these homogeneous regions and to describe the distribution of its sociodemographic indicators.

METHODS

For this study, sociodemographic indicators were obtained from special tabulations of the Demographic Census of 1991. The IBGE database offers the information grouped into census sections of approximately 1000 residents each. The first step was to identify the 70 census sections belonging to the EPM HCC coverage area, defining them by the streets that serve as their boundaries. Six variables of interest were compiled for studying the population of each section, as follows:

- proportion of elderly: percentage of people aged 60 and over;
- proportion of children: percentage of children between 0 and 4 years;
- maximum education: percentage of family providers with more than 15 years of study;
- minimum education: percentage of family providers with less than 1 year of study;
- maximum income: percentage of family providers with income greater than 20 minimum wages; and
- minimum income: percentage of family providers with income below 1 minimum wage.

Social characterization of each census section is sought via these factors, but they do not specifically address the relationship between health and disease. Although it seems clear that the area with the lowest educational and income levels will present worse conditions of health than other areas, and therefore different Public Health actions should be directed towards them, it was decided that the relationship between social level and health conditions should be objectively

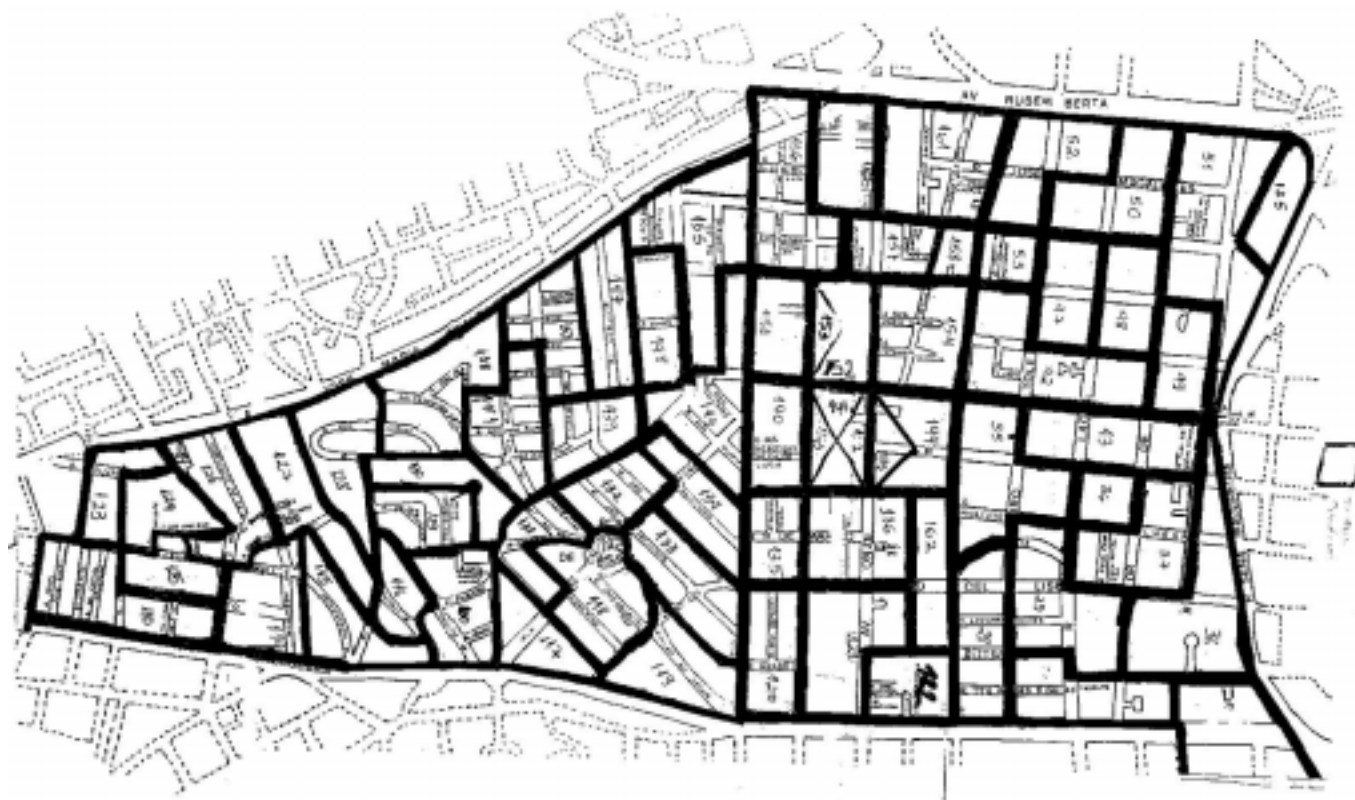


Figure 1 - Map of the UNIFESP/EPM Healthcare Center coverage area, 1996. The area is delimited by Av. José Maria Whitaker, Av. Rubem Berta, Av. Jabaquara and Av. Sena Madureira.

investigated. Thus, correlations were made between information on causes of death in the area during 1995, obtained from the PRO-AIM (program for improvement of mortality information in the city of São Paulo), and the sociodemographic variables from the corresponding census sections.

Characterizing each census section according to the proportions of elderly people and children, and the educational and income levels, provided the basic data for analyzing which of these parameters differentiated the sections better, so that similar census sections could be grouped to form homogeneous regions.

To evaluate the behavior of the other sociodemographic variables and to verify the structure of the causes of death in different social districts of the EPM HCC coverage area, the 70 census sections were characterized by ranking the proportions of maximum income (above 20 minimum wages). The proportions of elderly and children were equivalent in all sections, thus not being good indicators for distinguishing

homogeneous regions. Nor did minimum education and income distinguish between sections, since the percentage of family providers in those conditions was equally low. Maximum education and maximum income maintain a proportional and direct relationship to each other (the higher the education level, the higher the family income). Of these indicators, maximum income exhibited substantial contrasts among the sections and it distinguished between those which were similar to each other for defining the homogeneous regions. The higher the cut-off point of the variable, the better the discriminative power of the indicator was (for areas with similar socioeconomic levels), as in the case of maximum income, a positive indicator.

The results are presented in the form of boxplots: graphs that present the distribution of the values, using the median as the reference. Thus, the "boxes" are drawn with the median represented by a line inside them. These "boxes" represent the values of the interquartile intervals that relate to the values included between the 25th and 75th

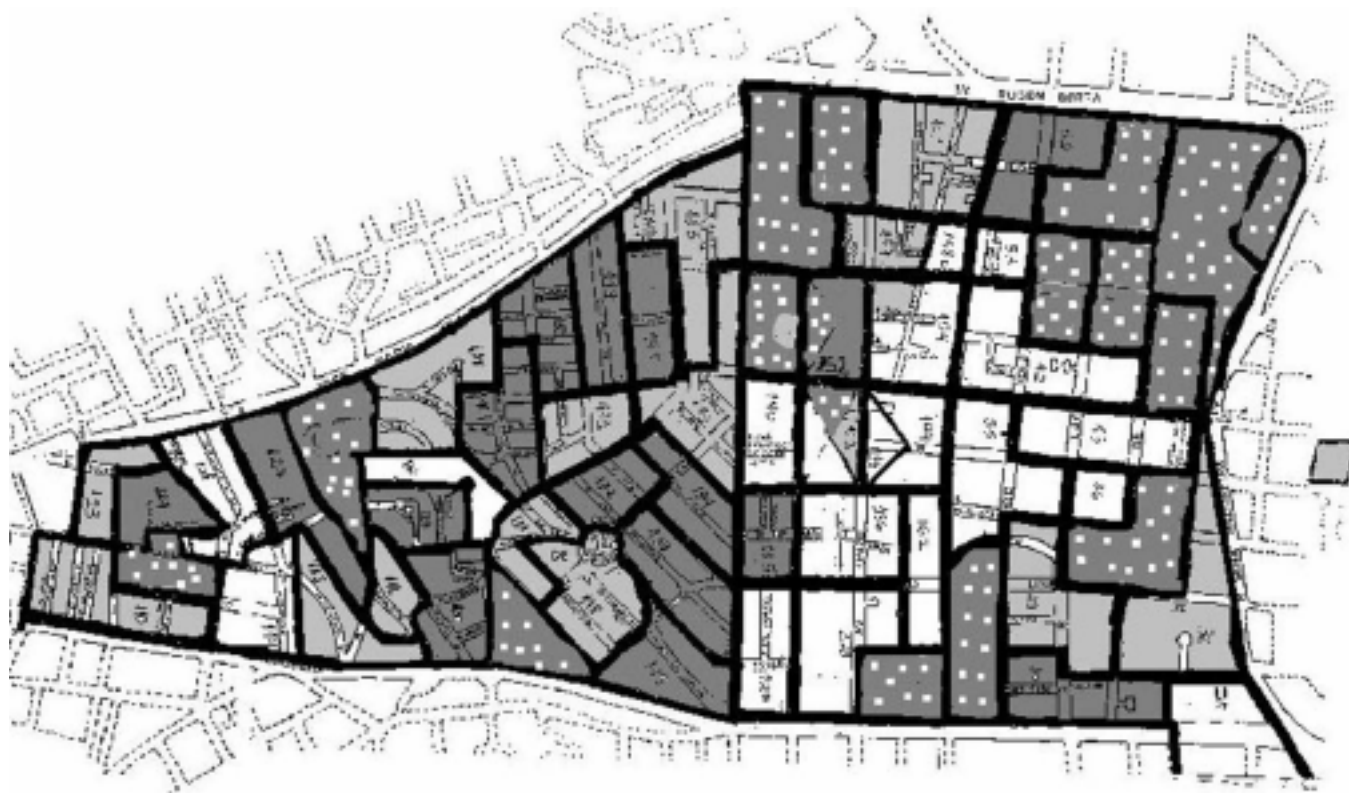


Figure 2 - Map of the UNIFESP/EPM Healthcare Center coverage area, 1996, divided into: homogeneous region I in light grey, region II in dark grey, region III in white and region IV in dark grey and white.

percentiles. Thus, 50% of the central values of the distribution around the median are included. So-called "fences" are traced out, extending from the tops and bases of these boxes, representing the upper and lower limits of the distribution, respectively. The upper limit is calculated from the sum of the 75th percentile value and 1.5 times the value of the interquartile interval (p25 to p75), being traced only as far as the highest value in the database that meets this condition. The lower limit, by analogy, is calculated by the subtraction of 1.5 times the value of the interquartile interval from the 25th percentile, and is traced as far as the lowest value in the database. Any value outside of these limits is represented by a small circle and classified as an outlier, i.e. a value very distant from the median and therefore rare.

Once the homogeneous regions of EPM HCC coverage area were defined, the structure of causes of death was examined in these areas, following the main guidelines of the ICD-1975 (International Classification of Diseases, ninth revision). This first glance at the area would allow different mortality structures to be identified which would not have been recognized at first, when the data was consolidated in great heterogeneous areas.

RESULTS

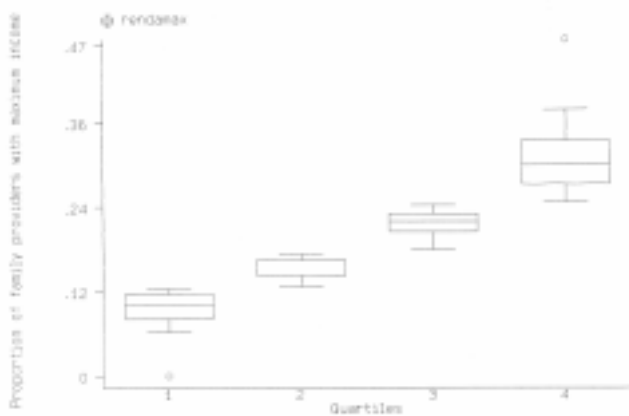
The total HCC coverage area (Figure 1), containing small parts of the Vila Mariana and Saude

Administrative Districts (Municipal Law no. 10,932 of January 15, 1991) has a generally good socioeconomic and cultural level, above the average for the total population of the city of São Paulo.

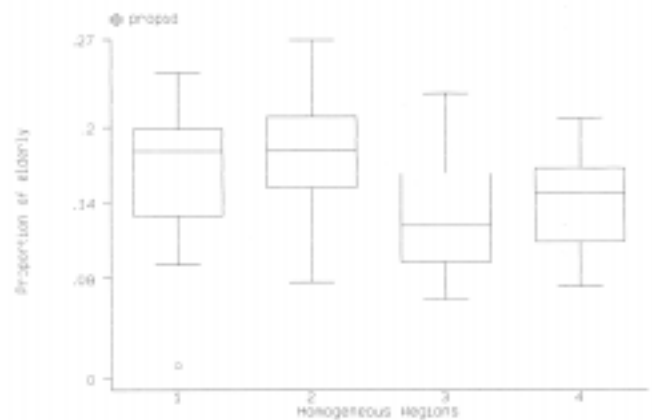
Graph 1 was compiled in accordance with the proportions of family providers with incomes greater than 20 minimum wages, dividing the coverage area into four regions with the same number of sections (quartiles). The first quartile (homogeneous region I) is made up of sections in which the proportion of family providers with maximum income was below 12.5%. The second quartile (homogeneous region II) contains the sections in which the proportion of family providers with maximum income varied between 12.5% and 17%. The third quartile (homogeneous region III) contains the sections in which the percentage of family providers with maximum income varied between 17% and 24%, and the fourth quartile (homogeneous region IV) is made up of sections in which the percentage of family providers with maximum income was over 24%. Thus, a map could be drawn indicating the homogeneous regions (Figure 2).

When the proportion of elderly people is analyzed for each of the four areas, it can be verified that there is no statistically significant difference among them. The elderly presented a median proportion of 16% in the total population of the four regions (Graph 2).

The same phenomenon is seen with the



Graph 1 - Distribution of proportions of family providers with monthly income of more than 20 minimum wages among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 2 - Distribution of proportions of elderly people among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.

proportion of children aged 0 to 4 years, which maintains a median of around 5% of the total population in the four regions (Graph 3).

Regarding minimum income, it can be seen that only the first quartile differs statistically from the others, presenting a median of 8% of family providers with income inferior to one minimum wage; in the other quartiles, the median was estimated to be around 5% (Graph 4).

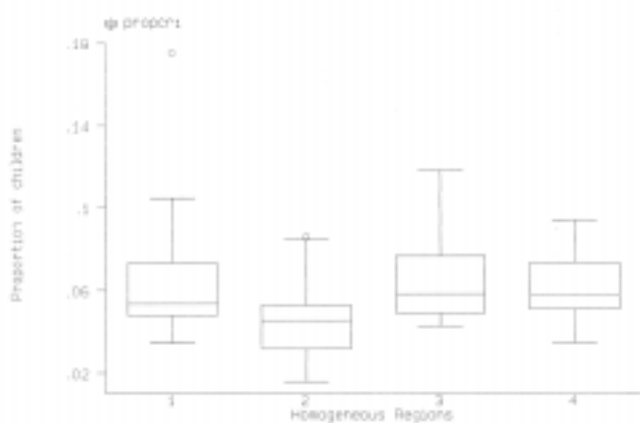
Minimum education is distributed similarly: the first quartile includes the sections whose distribution represents a median of 2.5% of family providers with less than a year of study; in the other quartiles, the median of the providers with minimum education is from 0.9% to 1.3%. Thus, a statistically significant difference between the first and the other quartiles was observed (Graph 5).

Differently, the graph that describes the distribution of the percentage of family providers with maximum education demonstrates a clearly increasing trend from the first to the fourth homogeneous region. The first region includes those sections where the median of family providers with more than 15 years of study is 30.5%, and in the second region, 32%. In the third and fourth regions the median only varied slightly: 47.1 and 48.3% respectively, but with one difference: in the fourth region the distribution of the sections was more concentrated around the median, that is to say, the distribution was less dispersed than that of the third region (Graph 6).

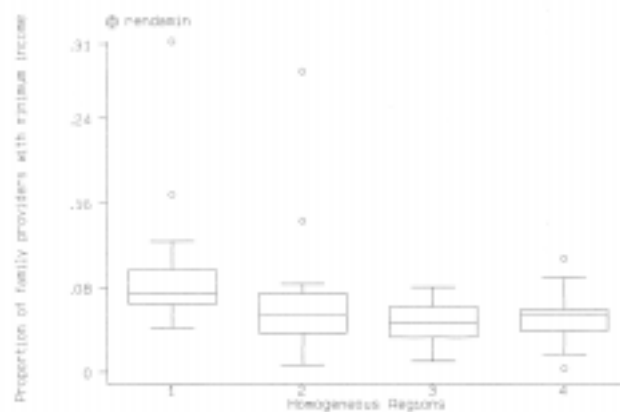
With regard to the death data (PRO-AIM, 1995), it could again be verified that the first region differed from the others. While the leading causes of death in the second, third and fourth regions were cardiovascular disorders, neoplasms, external causes and AIDS (not always in that same order), as is the overall pattern in the city of São Paulo, in the first area respiratory disorders were the second biggest cause of death, preceding neoplasms and these being followed by deaths from non-alcoholic cirrhosis. It is also interesting to observe that AIDS does not appear as an important cause of death in the first region. Among the respiratory disorders, 70% were due to bronchopneumonia and tuberculosis. As for non-alcoholic cirrhosis, it was noticeable that it was especially due to viral hepatitis affecting mainly children and newborns (Graphs 7 to 10).

DISCUSSION

Sociodemographic indicators were studied in order to characterize the resident population in the UNIFESP HCC coverage area. For this purpose, various criteria³ were employed in choosing these indicators, such as: validity (adaptation of the indicator to characterize the population), reliability (reproducibility of the data in new measurements), representativity (guarantee of having good coverage of the area), as well as being attentive to



Graph 3 - Distribution of proportions of children aged 0-4 years among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 4 - Distribution of proportions of family providers with monthly income less than one minimum wage among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.

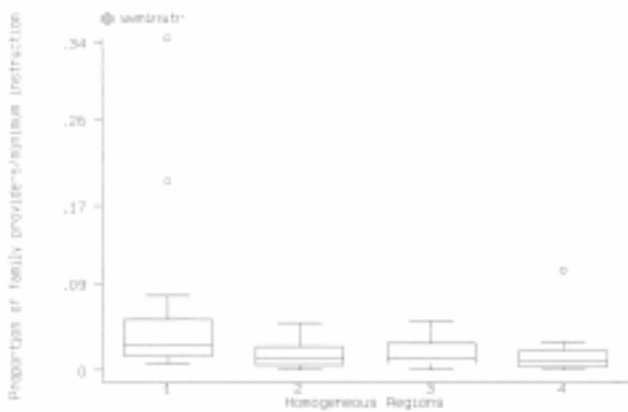
ethical considerations matters (ensuring that data collection would not result in harm to the population and that confidentiality towards individuals was maintained). Thus, the identification of the proportions of children and of the elderly in the area was the first pertinent classification, since regions with predominantly young or elderly populations may need different health actions. All census sections had similar proportions of these age groups, whose sizes were close to the values seen in the developed world, with proportions of children between 3% and 5% and proportions of the elderly between 20% and 25%.

Another indicator that seemed of great relevance was the degree of education among the population, because it may be inferred that educational level and awareness of risk factors in disease prevention maintain a proportional and direct relationship.³ Thus, the proportions of family providers with minimum and maximum education revealed a good educational level in the vast majority of sections, with the number of family providers with tertiary education (more than 15 years of study) reaching almost half of the total and the proportion of providers with less than one year of study not reaching 5%. The indicator also revealed a small group of sections where such characteristics were not valid. In spite of the differences it revealed, this indicator did not prove itself sensitive enough to

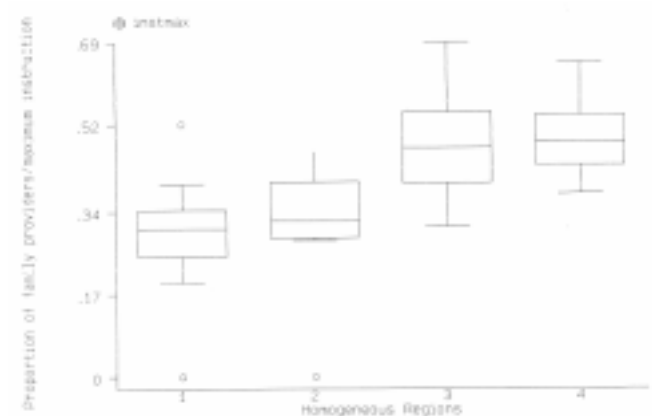
define homogeneous regions, since the educational level of the area is rather uniform.

Studying the relationship between education and income in the area, a strong statistically significant correlation ($r = 0.7183$; $p < 0.0001$) between these two indicators was verified. In the evaluation of minimum income a similar pattern to that of education was noticed. A small group of sections stand out from the rest of the population because of larger proportions of family providers without education and with minimum income. On the other hand, the pattern of the proportion of family providers with maximum income differed from all the other indicators, exhibiting an increasing trend across the four groups of sections.

The fundamental step in this study was to choose the best sociodemographic indicator from among the available ones in order to discriminate between sections and group them into homogeneous regions. The proportion of providers with maximum income turned out to be the most adequate indicator for this purpose. When the other indicators were analyzed in relation to maximum income, it was noticed that the patterns found throughout the HCC coverage area reproduced themselves when divided into homogeneous regions (Graph 11), giving support to the idea that this was most appropriate socioeconomic indicator for the intended



Graph 5 - Distribution of proportions of family providers with education less than one year of study among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 6 - Distribution of proportions of family providers with education greater than 15 years of study among the four homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.

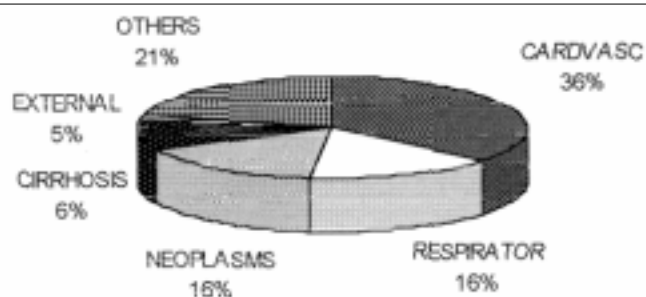
division. It can be observed that the distribution of maximum income presents a larger dispersion of values around the median, only being exceeded by the distribution of maximum education. However, the distribution by income presented values as low as 0% at its lower limit, different from the distribution by education, with values around 30% at its lower limit. Thus, a positive indicator was chosen, with significant dispersion of values around its median and containing values with good discriminative power between sections, provided these were not outliers, as in the case of maximum education.

After dividing the area into four homogeneous regions according to maximum income, we defined the subdivisions on the map of the UNIFESP HCC coverage area (Figure 2). The first region (region I), least affluent, with larger proportions of providers without education and with minimum income had a pattern of diffuse geographical distribution, gathering together the sections where the Mário Cardim, José Maria Whitaker (no longer in existence) and Onze de Junho slums were located, as well as other sections, mainly more distant from the EPM and the UNIFESP HCC.

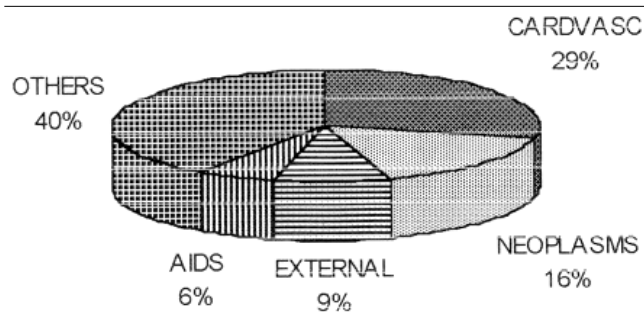
The second region (region II) is located close to region I and has a similarly diffuse distribution pattern.

The third region (region III) is very concentrated around EPM, close to the UNIFESP HCC, and joins with census sections located predominantly in the Vila Clementino district.

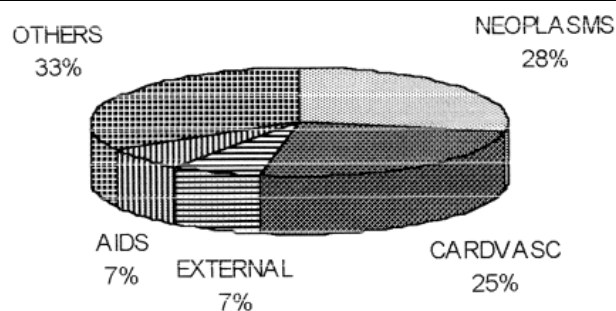
The fourth region (region IV), with larger proportions of maximum education and maximum income, is concentrated close to Av. Sena Madureira, where the best dwelling conditions are located. The Neide Solito slum is also located in this region and the best explanation for this fact is that census data from apartment blocks around the slum cause the characterization of that section as high educational and income level. The Neide Solito slum is in fact more organized in relation to other slums; it has houses built with concrete, electric power, water supply and availability of sewers, certainly not presenting sociodemographic



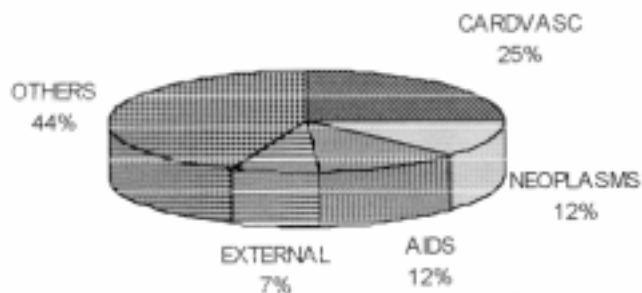
Graph 7 - Distribution of the principal causes of death in the first homogeneous regions. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 8 - Distribution of the principal causes of death in the second homogeneous region. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 9 - Distribution of the principal causes of death in the third homogeneous region. UNIFESP/EPM Healthcare Center coverage area, 1996.



Graph 10 - Distribution of the principal causes of death in the fourth homogeneous region. UNIFESP/EPM Healthcare Center coverage area, 1996.

indicators like those of other slums in the area belonging to the homogeneous region I.

Health and disease are complex concepts that involve not only physical but also mental and social³ well-being; thus there may be difficulties in measuring disease. Morbidity is a generic term used for designating groups of cases of a given disease, hazards or the sum of health problems that involve a group of people.⁴ If on the one hand morbidity tends to be the best indicator for evaluating a population's health conditions (bringing direct information on diseases without the examiner's⁴ personal conclusions), on the other hand, great limitations impeded its use in this study. The sources for morbidity measurements are routine health service registrations (giving information on attendance requirements, service usage or clinical exam results), surveillance³ system registrations and disease notifications,³ and the population's enquiries about health problems.

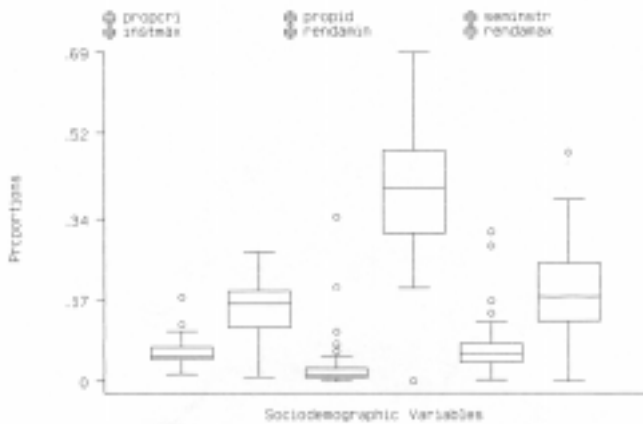
Besides its high cost due to the involvement of qualified health professionals, an enquiry that takes into account diagnostic complexity and types of health hazards can have its validity compromised by communication deficiencies, omissions or inaccuracies in disease diagnosis (especially for chronic diseases, which may be difficult to confirm on a single visit, like arterial hypertension and diabetes melitus).³

Sample variation and the coverage of the population are other factors that place the use of morbidity at a disadvantage, both in primary data collection for studies and in research of secondary data through routine registrations, since the demand for the UNIFESP HCC cannot be characterized as a random sample of the sick population: the population of the nearest sections has easiest access; the population of sections with better economic levels probably seeks private clinics or the facilities of health insurance schemes, avoiding public services.

Another important issue that hinders the real representation of morbidity is that relationships between need, demand and offer have multiple combinations and the interpretation of the results of morbidity studies depends on data collection being population-based, on health service demand registration, or on clinical results registration.³

Historically, mortality was the first indicator used in evaluating collective health. The greatest limitation on this indicator is that it does not give a global vision of the population's health conditions. As death represents the last event in the health-disease process, the mortality statistics reflect an incomplete picture of the health situation.³ In spite of this limitation, mortality data are still used more frequently than other sources of data today. This can be explained by the obligation of death registration, attributing good coverage and representativity to the data.³ The reliability of the information is reinforced by the PRO-AIM, a program that not only monitors mortality data, but also requires quality data, returning to the physician erroneously filled-out death declarations and/or those in which there remain doubts about the causes of death, which might have arisen from previous conditions, such as AIDS, pregnancy or childbirth.

Considering that death data may help in understanding local situations and consequently in aiding the decision-making process regarding measures to be applied for the benefit of the population, the mortality profile of the population of the HCC coverage area was investigated. The results strikingly indicated



Graph 11 - Distribution of proportions of children, elderly people, family providers without education or with minimum income, and family providers with maximum education or income. UNIFESP/EPM Healthcare Center coverage area, 1996.

a great number of deaths in the first area, generated by infections that could have been prevented or treated earlier, such as tuberculosis, pneumonia and hepatitis, suggesting either a lack of early health care, perhaps on account of access difficulties, or lack of quality in health services. Are there alternative stances regarding risk situations generating these health hazards? Are there alternative concepts and perceptions concerning the health-disease process? These are important discussion topics proposed for further study after a more profound investigation of the population under the responsibility of the UNIFESP HCC.

CONCLUSION

The present study offered the opportunity to describe in detail the UNIFESP HCC coverage area. It can be concluded that, even though it is a privileged area, when submitted to income distribution analysis, inequalities can be observed and they are reflected in the mortality data, the final event of the health-disease process. The delineation of the area into smaller, homogeneous regions makes it possible to adapt and prioritize the HCC's actions, so as to attempt to modify the health-disease process, aiming at the primary prevention of health hazards.

REFERENCES

1. Ramos LR, Gohman S. Geographical stratification by socio-economic status: methodology from a household survey with elderly people in São Paulo, Brazil. *Rev Saúde Públ* 1989;23:478-92.
2. Bussab WO, Dini NP. Pesquisa de emprego e desemprego SEADE/DIEESE: regiões homogêneas da Grande São Paulo. *Rev Fund SEADE* 1985;1(3):5-11.
3. Pereira GP. *Epidemiologia: teoria e prática*. Rio de Janeiro: Editora Guanabara Koogan; 1995.
4. Laurenti R, et al. *Estatísticas de saúde*. São Paulo: Editora EPU; 1985.

From Centro de Saúde Escola, Departamento de Medicina Preventiva, Universidade Federal de São Paulo/Escola Paulista de Medicina, São Paulo, Brazil

Authors

Karina Yuri Harada - Medical student.

Julieta Gonçalves Silva - Medical student.

Simone Schenkman - Doctor, MSc in epidemiology, supervisor of residency in preventive and social medicine.

Eliana Tiemi Hayama - Psychologist.

Francisco Roberto Gonçalves Santos - Assistant professor and head of discipline.

Mariângela Cainelli de Oliveira Prado - Assistant professor and PhD from FSP/USP.

Regina Helena Petroni Pontes - Educator in public health and PhD from FSP/USP.

Sources of Funding: Not declared

Conflict of interest: Not declared

Last received: 21 May 1998

Accepted: 30 July 1998

Address for correspondence:

Karina Yuri Harada - A/C Dra. Simone Schenkman
Disc. Epidemiologia, Deptto. de Medicina Preventiva
Rua Pedro de Toledo, 675
São Paulo/SP - Brasil - CEP 04039-032

RESUMO

Objetivo: Delimitar regiões homogêneas na área de abrangência do Centro de Saúde-Escola-EPM/UNIFESP (CSE) com base em indicadores sociodemográficos, bem como relacioná-los às causas de mortalidade de 1995.

Tipo de estudo: Análise de dados secundários.

Localização: Área de abrangência do CSE; atenção primária.

Variáveis estudadas: Populacionais: proporções de crianças e idosos; chefes de família: instrução (máxima: >15 anos; mínima: <1 ano) e renda (máxima: > 20 salários mínimos; mínima: <1 salário mínimo). Distribuição de mortalidade proporcional

Resultados: A renda máxima permitiu a construção de quatro regiões homogêneas, de acordo com sua escala por postos. Embora a proporção de crianças e idosos não tenha variado significativamente entre as regiões, a renda e instrução mínimas apresentaram uma diferença estatisticamente significativa ($p < 0,05$) entre a primeira região (menos favorecida) e as demais. Gradiente crescente foi igualmente observado (instrução mínima). A mortalidade também diferiu na primeira região, com óbitos por causas preveníveis.

Conclusões: As desigualdades observadas podem contribuir para a prevenção primária da saúde.