

Luciana Tricai Cavalini^I

Antonio Carlos Monteiro Ponce
de Leon^{II}

Correction approach for underreporting of deaths and hospital admissions due to ill-defined causes

ABSTRACT

OBJECTIVE: To propose a correction approach for underreporting and relocation of ill-defined causes of morbidity and mortality in the National Health System Mortality and Hospital Information Systems.

METHODS: Modified James-Stein empirical Bayes estimators for events in delimited geographic areas were applied as a correction approach for underreporting in Brazilian municipalities in 2001.

RESULTS: There was an increase of 55,671 deaths in the Mortality Information System, an underreporting correction of 5.85%. It was more effective at the age groups under five (8.1%) and 70 years old and more (6.4%); for neonatal (8.7%) and ill-defined (8.0%) causes of death; and in the states of Maranhão (10.6%), Bahia (9.5%) and Alagoas (8.8%). Relocation of ill-defined causes of mortality changed the structure of proportional mortality in the Northern and Northeastern regions, and increased the proportion of deaths due to cardiovascular diseases and reduced those due to external and neonatal causes. Relocation of ill-defined causes of hospital admissions did not affect hospital proportional morbidity.

CONCLUSIONS: The results of underreporting correction were consistent with previous studies, in terms of age groups, causes and geographic areas. Relocation of ill-defined causes of death was spatially consistent. The approach studied may be applicable on Brazilian Health Information since it can be implemented in computational algorithms. Some improvements, however, may be considered, like estimation approaches based on time-space event distribution.

KEYWORDS: Mortality. Morbidity. Underregistration. Cause of death. Information systems. National Health System (BR).

^I Departamento de Epidemiologia e Bioestatística. Instituto de Saúde da Comunidade. Universidade Federal Fluminense. Niterói, RJ, Brasil

^{II} Departamento de Epidemiologia. Instituto de Medicina Social. Universidade do Estado do Rio de Janeiro. Rio de Janeiro, RJ, Brasil

Correspondence:

Luciana Tricai Cavalini
Prédio Anexo do Hospital Universitário
Antonio Pedro
Rua Marquês de Paraná, 303 3º Andar
20241-263 Niterói, RJ, Brasil
E-mail: lutricav@vm.uff.br

Received: 11/3/2005 Reviewed: 7/4/2006
Approved: 8/28/2006

INTRODUCTION

Health information systems are strategic tools in the management of the Brazilian National Health System (SUS). They can contribute to define health priorities, composition health care organization and implement actions of control and assessment. The most frequently estimated indicators of morbidity and mortality in Brazil show in their numerators data from the following systems: *Sistema de Informações sobre Mortalidade* (SIM - Mortality Information System) and *Sistema de Informações Hospitalares* (SIH-SUS - SUS Hospital Information System).

SIM is the major source of information for making diagnoses on the health status in Brazil. Despite it has been long established, there is no consensus on this data system's value as an instrument of health surveillance and management, notwithstanding the high commitment of health services and providers in providing information. It is estimated an approximately 20% of underreporting of total deaths in Brazil in 1999.³ Incomplete information was also concerning, especially relating to the underlying cause of death.^{9,12} These limitations prevents wider utilization of SIM and, therefore, some experts advocate either its replacement or at the least the inclusion of data from population-based morbidity and mortality surveys.¹⁸ However, as these surveys are costly in such a vast territory as Brazil, they become in effect prohibitive especially currently when limited resources are allocated to social fields (including health) and academic research.

According to official estimates, SIH-SUS registers 70% of all hospital admissions in Brazil* but among its major limitations are the non-inclusion of private hospital admissions and incomplete information.¹¹

Despite its limitations, SIH-SUS has been widely used in studies on public hospital care⁵ and economical health analyses⁷ as a supporting tool for estimating indicators of mortality and¹⁵ disease incidence rates⁸ as well as a valuable instrument for health surveillance.¹³

Approaches for the correction of death underreporting have included demographic methods, such as Brass¹ and Coubage & Fargues,** which involve assumptions such as constant fecundity and linear variation of mortality. As they are census-based methods, the direct estimate of mortality is limited to a 10-year time period. In addition, mortality estimated through

these methods has not been used for estimating mortality due to specific causes.

In regard to information on death causes and hospital admissions, simplified correction approaches for incomplete data do not take into account the fact that underreporting of the cause of an event is likely to be inconsistent for each specific cause; for instance, the allocation of events due to ill-defined causes according to the proportional distribution of events due to well-defined causes.

The present study aimed to propose an approach for the correction of death underreporting and relocation of ill-defined causes and hospital admission that can be systematically applied in computational routines. As a result, this approach could be applied in health care management without detriment to the continuing efforts towards reducing the actual underreporting and in improving the quality of data reported to health information systems.

METHODS

SIM and SIH-SUS data for the year 2001 was used as a reference in the study.

Correction of SIM underreporting

For the correction of underreporting, James-Stein empirical Bayes estimators,⁴ simplified by Marshall,¹⁰ were applied as they provide an *overall contraction* of estimators of the number of events in small areas related to the global mean in a large area, consisting of all smaller areas. The term *contraction* was defined as an approximation of incidence rates of a given event in small areas related to the reference rate in the large area, inversely proportional to the population size in the small area. This is a fitting characteristic in a scenario where the overall mean is strongly affected by rates of small areas with large population size, as in the case of the number of deaths by municipalities.

In other words, mortality rates of Brazilian federal units are mostly affected by its populous municipalities. Coincidentally, lower SIM underreporting is seen in the most developed municipalities¹⁶ that, in Brazil, are also the most populous ones: capitals, cities in metropolitan areas and midsize cities.¹⁴

Taking the federal unit, or a mesoregion, as the large area, and by contracting mortality rates of municipalities related to this mean, it was anticipated an

*Ministério da Saúde, Secretaria de Vigilância em Saúde. Guia de vigilância epidemiológica. 6ª ed. Brasília (DF): 2005. Sistemas de informação em saúde e vigilância epidemiológica. v. 1; p. 60-77.

**Vieira Jr L. Diferenciais dos riscos de morte infantil segundo aspectos sociais das mães no Brasil [dissertação de mestrado]. Rio de Janeiro: Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro; 2000.

increased number of deaths in the small municipalities, prone to higher underreporting.

The empirical Bayes estimator of specific mortality rate, by age group and ICD-10 chapter, is theoretically defined by the following formula:

Empirical Bayes estimator of mortality rate = Observed mortality rate in the large area + contraction factor \times (observed mortality rate in the municipality - observed mortality rate in the large area)

Taking the large area, the contraction factor for each municipality is defined by the following formula:

Numerator of contraction factor = Variance of mortality rate in the large area - (observed/expected death ratio in the large area \times variance of mortality rate in the small area)

Denominator of contraction factor = Numerator of contraction factor + [(observed/expected death ratio in the large area) / (number of expected deaths in the small area)]

Thus, in populous municipalities (with high number of expected deaths), the denominator and numerator of the contraction factor will tend to be the same and the contraction factor will be likely 1. As a result, the difference between mortality rate in the municipality and in the large area will remain unchanged in the estimator.

On the other hand, in small municipalities (with a small number of expected deaths), the denominator of the contraction factor will be greater than the numerator, and the contraction factor will likely be zero. Thus, the estimator of mortality rate in small municipalities will be close to the mortality rate observed in the large area.

When the mesoregion was assumed as the large area, correction of underreporting was more effective than when the federal unit was taken, because high variance of rates in small areas reduces the contraction factor ("c"), which in turn works by approximating the rate of the small area to that of the large area.

The estimator of mortality rate needed to be adjusted in municipalities where mortality estimates have been reduced, i.e., when death estimates were lower than those reported in SIM, and in this case, the latter remained in the analysis. To reduce deaths would mean to arbitrarily eliminate SIM data, which would artificially reduce the number of deaths in municipalities with higher mortality rates than the mean rate in the

large area. As a result of the adjustment, the total number of deaths will be likely increased, since there will be only additions, and never reductions, to the number of deaths in each municipality.

The correction was applied, by municipality and by age group, for each chapter of the International Statistical Classification of Diseases and Related Health Problems - 10th Revision (ICD-10). The correction process was employed even to deaths due to ill-defined causes (underlying cause of death related in ICD-10 Chapter XVIII).

The corrected number of deaths was obtained using the following calculation:

Corrected deaths = SIM deaths - numerator of empirical Bayes estimator of mortality rate

It was chosen to weigh the corrected values obtained following quality criteria for SIM, previously established by Szwarcwald et al¹⁷ and adapted to the study. A maximum score of 4 was set to indicate maximum inadequacy of SIM, which was obtained through the sum of related scores according to the following criteria:

- a) Crude mortality rate, standardized by age, of less than 4 deaths per 1,000 inhabitants = 2; or between 4 and 6.75 deaths per 1,000 inhabitants = 1.
- b) Deviation of specific mortality rate in the municipality, by age group and cause of death, in relation to the mesoregion, greater than or equal to 10% = 1.
- c) Proportion of deaths due to ill-defined causes greater than or equal to 20% = 1.

The total number of deaths corrected was obtained using the following formula:

Deaths with underreporting correction = SIM deaths + [corrected deaths \times (score of SIM adequacy / 4)]

An example of the approach application can be obtained under request to the first author.

Relocation of SIM and SIH-SUS ill-defined causes

Likewise it was chosen to employ James-Stein estimators in the relocation of deaths and hospital admissions related to ICD-10 Chapter XVIII. For that, Brazilian municipalities were regrouped and a new configuration of the large area was created. It was sought to obtain clusters of municipalities with similar demographic characteristics, because, according to the theory of demographic and epidemiologic tran-

sition, disease development and dying processes are determined by a population's composition, since diseases are generally specific to each period of life.¹⁹

In order to generate the new clustering of municipalities, cluster analysis was used because it can indicate if a given unit of observation belongs to a given set, known as cluster. They are not a priori defined but rather during the analysis per se: units of observation with greater similarities tend to form a cluster.

In a simplified manner, this analysis can be thought to be carried out in a set of n units of analysis with two variables that will be used for generating clusters. These variables can be displayed on the abscissa and y-axis of a two-dimensional Cartesian coordinate system so that each unit of observation will be represented by a point and those closely arranged will form a cluster.

The mean method, which is the simplest and commonly used, was applied to generate clusters from variable means and square Euclidian distances between units of observation.⁶

Municipalities in each macroregion were considered units of observation and variables were population rates by sex and 5-year age group, multiplied by 100 (percentages), since cluster analysis does not satisfactorily discriminate very close observations, as in the case of rates between 0 and 1.

Based on the five sets of clusters generated, one for each Brazilian macroregion, the contraction approach was applied to mortality and hospital admission rates in the municipalities, by ICD-10 chapter, related to rates of their clusters. Since it was not required to correct underreporting at this step, all rates obtained for each municipality showed random variation. However, this same approach was not applied to deaths and hospital admissions related to Chapter XVIII, given that their final rates, obtained after underreporting correction, would be relocated, at this step, among all deaths and hospital admissions related to the remaining ICD-10 chapters.

At the end of this process, each municipality had a proportional distribution of deaths and hospital admissions due to well-defined causes that can be considered "adjusted" in terms of their population composition. Yet, the absolute rates cannot be used as a final result because: a) municipalities with extremely high rates of mortality or hospital admissions reported to SIM and SIH-SUS could also have their rates artificially reduced in a second process of contraction in relation to the cluster mean; b) up to this point, there

have not been included those deaths actually reported and with corrected underreporting, as well as those hospital admissions actually reported, due to causes related to ICD-10 Chapter XVIII.

Therefore, in the second process of contraction of mortality and hospital admission rates in the clusters, a "relocation factor of ill-defined causes" was created and defined as the proportion of deaths or hospital admissions due to well-defined causes, by ICD-10 chapter.

The final number of deaths in an ICD-10 chapter due to well-defined causes was obtained using the following formula:

Final deaths in an ICD-10 chapter due to well-defined causes = deaths with corrected underreporting in this chapter + ["relocation factor of ill-defined causes" x total deaths with corrected underreporting in ICD-10 Chapter XVIII]

Final hospital admissions in an ICD-10 chapter due to well-defined causes = SIH-SUS hospital admissions in this chapter + ["relocation factor due to ill-defined causes" x total SIH-SUS hospital admissions in ICD-10 Chapter XVIII]

RESULTS

Correction of SIM underreporting

In 2001, there were reported 952,194 deaths in SIM. After correcting underreporting, there were a total of 1,007,865 deaths (5.9% correction) nationwide. The correction rate varied according to age group, cause of death and geographic area.

As for age, higher correction rates were found in the outmost age groups. The highest correction rate was seen in those under the age five (8.1%), and then correction rates dropped in the those aged five to nine, and then went up again in those aged between 10 and 19 years. Older age groups showed a growing curve of correction rates. This trend was seen nationwide with some differences: it was more remarkable in the North and Northeast regions and less evident in the Southeast region, especially in the population over 19 years of age (Table 1).

As for cause, there were differences among geographic areas. The North and Northeast regions showed equally high correction rates of underreporting of deaths due to congenital malformations (ICD-10 Chapter XVII) and perinatal causes (Chapter XVI). However, in the North region, a significant correction rate was found

Table 1 - Underreporting correction rate of Sistema de Informações sobre Mortalidade (SIM), by age group and macroregion. Brazil, 2001.

Age group (years)	North	Northeast	Macroregion (%) Southeast	South	Midwest	Total
0-4	8.4	10.2	6.7	6.3	5.6	8.1
5-9	3.9	2.0	4.1	0.6	0.9	2.7
5-9	3.9	2.0	4.1	0.6	0.9	2.7
10-19	5.8	5.3	5.2	5.2	3.7	5.2
20-49	4.7	5.8	4.4	3.8	3.6	4.6
50-64	6.1	8.3	4.2	5.0	4.1	5.3
65 and more	7.5	9.3	4.9	6.2	5.2	6.3
Total	6.7	8.4	4.7	5.5	4.6	5.9

for ill-defined causes of death (Chapter XVIII), and the same was seen in the Northeast region for diseases of the circulatory system (Chapter IX). In the South and Southeast regions, a parallel trend of underreporting correction was seen, with higher rates for perinatal and ill-defined causes and diseases of the respiratory system (Chapter X). The Midwest region showed a distinct trend, with higher correction rates for perinatal causes, diseases of the circulatory system and external causes (Chapter XX). In all macroregions, the rates of underreporting correction were low for the remaining causes, which can be considered either unusual or unlikely. Underreporting correction rates for causes of maternal death (Chapter XV) were also of small magnitude (Table 2).

The spatial distribution of underreporting correction shows that the highest rates were found in the North and Northeast regions. In addition, the mesoregion comprising the federal capital showed low correction rate, except in the states of Rondônia, Amazonas, Amapá, and Alagoas (Table 3).

Relocation of SIM deaths due to ill-defined causes

The relocation of SIM deaths due to ill-defined causes using the proposed approach produced changes in the epidemiology of mortality, expressed in the proportional mortality by cause. The magnitude of these changes, however, had regional differences. In fact, significant (and similar) changes were brought about

in the arrangement of proportional mortality by cause only in the North and Northeast regions, where rates showed a significant relative increase for diseases of the circulatory system (North = +8.5%; Northeast = +8.1%). In both macroregions, there was a significant relative reduction in deaths due to perinatal (North = -10.59%; Northeast = -13.19%) and external causes (North = -7.64%; Northeast = -10.85%).

The Southeast, South and Midwest regions showed a different, but parallel, trend compared to the North and Northeast regions. In these macroregions, the change of proportional mortality was small, especially for death causes of greater magnitude (Table 4).

Relocation of SIH-SUS hospital admissions due to ill-defined causes

The relocation of ill-defined causes of SIH-SUS hospital admissions by ICD-10 chapter did not produce any significant changes in the epidemiology of hospital morbidity in SUS when clustered data were analyzed at the level of macroregions. Hospital admissions rates by ICD-10 chapters with well-defined causes remained the same, and no change as in the epidemiology of mortality was detected (Table 5).

DISCUSSION

Higher rates of underreporting correction were seen in the outmost age groups, especially in deaths of children

Table 2 - Underreporting correction rate of Sistema de Informações sobre Mortalidade (SIM), by cause of death and macroregion. Brazil, 2001.

Cause of death (ICD-10 chapter)	North	Northeast	Macroregion (%) Southeast	South	Midwest	Total
I. Infecciosas	4.7	6.3	4.6	3.1	3.7	4.8
II. Neoplasms	7.4	9.1	4.0	5.1	3.5	5.1
IV. Endocrine	2.5	9.8	5.0	5.8	2.9	6.2
IX. Circulatory system	7.7	10.0	4.2	5.8	5.6	5.8
X. Respiratory system	5.8	8.0	5.2	7.3	4.7	6.1
XV. Maternal	0.0	1.1	2.8	0.0	0.0	1.3
XVI. Congenital	8.5	10.1	7.7	8.9	7.1	8.7
XVII. Perinatal	11.6	11.2	5.1	3.6	2.1	6.4
XVIII. Ill-defined	8.8	8.9	7.0	7.7	4.9	8.0
XX. External causes	5.9	7.6	4.8	5.4	5.5	5.7
Other causes	0.7	2.2	3.7	1.8	1.5	2.8
Total	6.7	8.4	4.7	5.5	4.6	5.9

Table 3 - Underreporting correction rate of Sistema de Informações sobre Mortalidade (SIM), by federal unit and mesoregions, after including or not the respective capital. Brazil, 2001.

Federation unit	Mesoregion (%)		Total
	Capital included	Capital not included	
Rondônia	6.0	4.6	5.2
Acre	4.7	5.0	4.8
Amazonas	8.5	7.1	8.3
Roraima	2.0	6.3	2.8
Pará	3.7	10.0	7.2
Amapá	2.7	1.5	2.7
Tocantins	4.6	6.3	5.8
Maranhão	10.0	11.1	10.6
Piauí	5.2	10.0	7.7
Ceará	3.4	10.8	8.0
Rio Grande do Norte	5.7	9.2	7.7
Paraíba	5.5	9.2	7.9
Pernambuco	3.1	10.7	7.4
Alagoas	9.3	7.7	8.8
Sergipe	5.8	8.3	6.6
Bahia	5.1	11.6	9.5
Minas Gerais	3.6	8.3	6.9
Espírito Santo	5.2	7.7	6.3
Rio de Janeiro	2.4	5.3	2.9
São Paulo	2.8	6.4	4.6
Paraná	0.0	6.0	4.2
Santa Catarina	4.1	7.1	6.7
Rio Grande do Sul	5.2	6.6	6.0
Mato Grosso do Sul	1.2	8.9	6.0
Mato Grosso	3.4	5.9	4.8
Goiás and Distrito Federal	2.9	7.5	4.0

under the age of five. This finding is consistent with the literature, since higher underreporting of deaths is anticipated in this age group.¹⁶ However, in intermediate age groups, a growing curve of the correction rates was seen for those aged between 10 and 24, especially for deaths due to external causes (6.9%) compared to other causes of death (2.0%). It is thus evidenced that the proposed approach for SIM underreporting correction allows to identifying social phenomena with impact on health indicators, since those deaths due to external causes among young people not reported to SIM could be largely due to homicides.

As for deaths due to maternal causes, the underreporting correction rate was low. Since these deaths occurred in age groups with low correction rates (except in the aforementioned case of violent deaths among young people, mostly males), it is likely that, at least in the Southeast and South regions, there is no residual underreporting. On the other hand, the relocation of deaths due to ill-defined causes did not

correct SIM inadequate information on maternal causes. Hence, low quality of reporting of maternal causes may be more substantial than the underreporting of deaths during pregnancy, labor, abortion or puerperium. Providing better quality of information on maternal mortality, especially in the North and Northeast regions, is still a challenge in Brazil.

The underreporting correction rate was consistent among geographic areas. The correlation between the estimated underreporting of deaths for the federal unit by Duarte et al³ (2002) and that found in the present study was 0.75, except for the North region. It is likely, in this macroregion, SIM deficiencies are of such great magnitude and so widely and heterogeneously distributed that the information available is not providing adequate input for the correction approach proposed in this study. For example, this can be evidenced by the fact that the three states of the North region (out of four nationwide) had underreporting correction rates higher in the mesoregion comprising their capitals than

Table 4 - Variation of proportional mortality due to well-defined causes resulting from relocation of ill-defined causes of death, by macroregion. Brazil, 2001.

Cause of death (ICD-10 chapter)	Macroregion (%)					Total
	North	Northeast	Southeast	South	Midwest	
I. Infections	-2.3	-2.6	+0.6	-1.0	+1.2	+0.7
II. Neoplasms	+1.0	-1.1	+0.0	-0.1	-2.0	-1.5
IV. Endocrine	+4.6	+5.2	+1.1	+0.6	+1.3	+3.9
IX. Circulatory system	+8.6	+8.2	+0.9	+0.6	+0.6	+2.5
X. Respiratory system	+1.4	+3.6	+0.8	+0.8	-1.3	+0.7
XV. Maternal	+2.4	-8.8	0.0	0.0	-4.0	0.0
XVI. Congenital	-10.5	-13.3	-4.3	-1.5	0.0	-4.3
XVII. Perinatal	-12.1	-12.8	-4.9	-1.7	0.0	-6.8
XX. External causes	-7.7	-10.9	-2.6	-1.8	0.0	-4.5
Other causes	+2.2	-1.4	+0.8	0.0	-0.3	-0.1

in to other mesoregions. This was not an anticipated result, since it was expected a higher correction rate in the areas not comprising the state capital, which presumably have more effective health information systems available. This finding evidences a need for applying a combination of several different approaches for underreporting correction in order to obtain estimates of specific mortality by cause in this region.

As for relocation of deaths due to ill-defined causes, the magnitude of change seen in the proportional mortality after correction was overall small. But these changes in the epidemiology of mortality were consistently produced in different geographic areas, as significant similar changes were seen in both the North and Northeast regions. However, more remarkable changes were expected in these macroregions given their higher rates of deaths due to ill-defined causes. In this sense, it proved correct the study assumption that, in areas with high mortality rates due to ill-defined causes, proportional relocation of these deaths might be inadequate.

The relative reduction of proportional mortality due to external causes, seen in the North and Northeast regions, corroborates the findings in the literature, that these deaths already had well-defined causes, in any case at the level of ICD-10 chapter information.² As a result, relocation of deaths due to ill-defined causes to this chapter was small, which consequently reduced its magnitude related to the total.

In regard to perinatal mortality, this same finding indicates the need for further investigation of potential determinants. Underreporting correction was significant, as already anticipated, but relocation of deaths due to ill-defined causes to this ICD-10 chapter was small. It is possible that deaths actually registered in SIM as occurring perinatally (that could be reported as deaths from perinatal causes) are more likely to have correct information of the cause related to Chapter XVI, since the cause is suggested by the age of death. This finding should be further investigated as well as its determinants, identifying whether better quality of

information on the cause of perinatal deaths reported to SIM, in any case in the ICD-10 chapter, would be attained primarily due to the fact it is primarily provided by physicians reporting deaths, and secondarily through child mortality surveillance systems, which would result in improved quality of information through investigation of deaths in this age group. Anyhow, the study results reinforce the notion that proportionally relocating deaths due to ill-defined causes, based on the proportional mortality reported to SIM, can be inadequate. Based on what was learned from perinatal causes, causes with high proportional mortality in a given age group can be added by a small number of deaths due to ill-defined causes in the non-proportional relocation, provided that other variables associated to the mortality epidemiology are considered, for instance, the population pyramid and geographic area, as described in the present study.

About SUS hospital morbidity, the process of relocation of ill-defined causes of hospital admission did not produce any significant changes in the epidemiology of this indicator. This is most probably due to the fact that there is no underreporting in SIH-SUS, and that the proportion of hospital admissions due to ill-defined causes is lower than that of deaths. Accordingly, in relative terms, a smaller basis is utilized in the process of non-proportional relocation of hospital admissions due to ill-defined causes, preventing from obtaining the same variation in SIM as seen in the North and Northeast regions. Besides, it may be detecting the established difference between mortality and morbidity trends. This suggests that the process of relocation of hospital morbidity due to ill-defined causes using James-Stein estimators should include other determining elements for small areas clustering, and not exclusively the population pyramid.

In conclusion, the approach for underreporting correction and relocation of ill-defined morbidity and mortality causes proposed in the present study allows improving the quality of national health information systems, especially SIM. And it can be accomplished by translating the proposed approach into computational

Table 5 - Variation of proportional SUS hospital morbidity due to well-defined causes resulting from relocation of ill-defined causes of hospital admission, by macroregion. Brazil, 2001.

Cause of admission (ICD-10 chapter)	Macroregion (%)					Total
	North	Northeast	Southeast	South	Midwest	
I. Infections	-0.1	-0.2	0.0	0.0	-0.2	-0.1
II. Neoplasms	0.0	+0.4	+0.3	+0.3	+0.4	+0.3
IV. Endocrine	0.0	+0.4	+0.3	+0.4	+0.4	0.0
IX. Circulatory system	+0.4	+0.4	+0.2	+0.1	+0.3	+0.2
X. Respiratory system	-0.1	-0.1	+0.1	-0.1	-0.1	-0.1
XV. Maternal	-0.1	-0.2	-0.4	-0.1	-0.2	-0.3
XVI. Congenital	0.0	-0.5	0.0	-0.3	-0.6	0.0
XVII. Perinatal	0.0	0.0	0.0	0.0	0.0	0.0
XIX and XX. Injuries + External causes	0.0	+0.2	+0.2	+0.2	0.0	+0.2
Other causes	+0.2	+0.2	+0.1	+0.0	+0.1	+0.2

algorithms. For instance, SUS Department of Information and Information Technology (Datusus) initiative for the development of an interface between the application *R* for statistical analysis and TabWin (application for tabulation and processing of health information system data) illustrates the operational capacity of the approach studied. Routines in *R* can also be developed for the estimates required in the correction of health information system data. Therefore, the proposed adjustment can be implemented and used not

only in this research field but also in the development of indicators by SUS management. Its automation through routines in the application *R*, integrated to the TabWin+*R* interface, would make its use easy by health surveillance and planning authorities, enabling them to rely their decision making on more validated information. Thus, the study primary aim was to recognize the value of and promote continuous and increasing utilization of health information systems for diagnoses in health in Brazil.

REFERENCES

1. Abreu DMX, Rodrigues RN. Diferenciais de mortalidade entre as regiões metropolitanas de Belo Horizonte e Salvador, 1985-1995. *Rev Saúde Pública*. 2000;34:514-21.
2. Drumond Jr M, Lira MMTA, Freitas M, Nitrini TMV, Shibao K. Avaliação da qualidade das informações de mortalidade por acidentes não especificados e eventos com intenção indeterminada. *Rev Saúde Pública*. 1999;33:273-80.
3. Duarte EC, Schneider MC, Paes-Sousa R, Ramalho WM, Sardinha LMV, Silva Jr JB, et al. Epidemiologia das desigualdades em saúde no Brasil: um estudo exploratório. Brasília (DF): Organização Pan-Americana da Saúde; 2002.
4. Efron B, Morris C. Data analysis using Stein's estimation rule and its competitors: an empirical Bayes approach. *J Am Stat Assoc*. 1975;70:311-9.
5. Escosteguy CC, Portela MC, Medronho RA, Vasconcellos MTL. O Sistema de Informações Hospitalares e a assistência ao infarto agudo do miocárdio. *Rev Saúde Pública*. 2002;36:491-9.
6. Everitt BS. Cluster analysis. 2nd ed. London: Heineman Educational Books; 1980.
7. Feijó MCC, Portela MC. Variação no custo de internações hospitalares por lesões: os casos dos traumatismos cranianos e acidentes por armas de fogo. *Cad Saúde Pública*. 2001;17:627-37.
8. Ferreira VMB, Portela MC. Avaliação da subnotificação de casos de Aids no município do Rio de Janeiro com base em dados do sistema de informações hospitalares do Sistema Único de Saúde. *Cad Saúde Pública*. 1999;15:317-24.
9. Laurenti R, Mello Jorge MHP, Gottlieb SLD. Mortes maternas no Brasil: análise do preenchimento da variável da declaração de óbito. *Inf Epidemiol SUS*. 2000;9:43-50.
10. Marshall R. Mapping disease and mortality rates using empirical Bayes estimators. *Appl Stat*. 1991;40:283-94.
11. Mathias TAF, Soboll MLMS. Confiabilidade de diagnósticos nos formulários de autorização de internação hospitalar. *Rev Saúde Pública*. 1998;32:526-32.
12. Mello Jorge MHP, Gottlieb SLD, Laurenti R. O sistema de informações sobre mortalidade: problemas e propostas para o seu enfrentamento: mortes por causas naturais. *Rev Bras Epidemiol*. 2002;5:197-211.
13. Mendes ACG, Silva Jr JB, Medeiros KR, Lyra TM, Melo Filho DA, Sá DA. Avaliação do Sistema de Informações Hospitalares - SIH/SUS como fonte complementar na vigilância e monitoramento de doenças de notificação compulsória. *Inf Epidemiol SUS*. 2000;9:67-86.
14. Pochmann M, Amorim R. Atlas da exclusão social no Brasil. São Paulo: Cortez; 2003. A exclusão social nas regiões e nos estados brasileiros. v. 1; p. 35-9.
15. Schramm JMA, Szwarcwald CL. Sistema hospitalar como fonte de informações para estimar a mortalidade neonatal e a natimortalidade. *Rev Saúde Pública*. 2000;34:272-9.
16. Szwarcwald CL, Leal MC, Castilho EA, Andrade CLT. Mortalidade infantil no Brasil: Belíndia ou Bulgária? *Cad Saúde Pública*. 1997;13:503-16.
17. Szwarcwald CL, Leal MC, Andrade CLT, Souza Jr PRB. Estimativa da mortalidade infantil no Brasil: o que dizem as informações sobre óbitos e nascimentos do Ministério da Saúde? *Cad Saúde Pública*. 2002;18:1725-36.
18. Viacava F. Informações em saúde: a importância dos inquéritos populacionais. *Ciênc Saúde Coletiva*. 2002;7:607-21.
19. Wood CH, Carvalho JAM. A demografia da desigualdade no Brasil. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada; 1999. Desigualdade de renda e expectativa de vida. p. 101-20.

LT Cavalini received a doctoral grant by Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ - Grant n. E-26/151.246/2002)

Based on doctoral thesis by LT Cavalini presented at the Instituto de Medicina Social of Universidade do Estado do Rio de Janeiro in 2005.