Can information about hospital mortality improve overall mortality figures?

Antony Stevens*
CGIAE/SVS – Ministry of Health, Brasília, Brazil - antony.stevens@saude.gov.br

1. Introduction

The Brazilian National Health Service (Sistema Único de Saúde (SUS)) was established by Federal Constitution in 1988. In this process the Hospital Information system was set up (Sistema de Informação Hospitalar (SIH/SUS)). The hospital authorisation record (Autorização de Internação Hospitalar (AIH)) is the basic document used by managers to record facts about hospital treatment within SUS. These records are sent to the Ministry of Health and most of the fields are subsequently made publicly available. There is also a private health sector.

The mortality information system (Sistema de Informações de Mortalidade (SIM)) is managed by the Data Processing Division of the Ministry of Health (DATASUS) and it provides individual mortality records for the whole country.

The Ministry of Health sponsors studies that attempt to improve the data recorded in SIM[França et al., 2014, L. et al., 2014].

The linkages here described are based on the AIHs and SIM records for the years 2008–2012 for the whole country. The tables and graphs however only represent data from the Northeast Region. This region is composed of nine states. This was a deliberate restriction to prevent the assumption that something could already be concluded about the country as a whole.

2. Linkage.

5.565 linkages were obtained, each representing a municipality of residence. This municipality is recorded in both types of file. The methodology for linkage follows that of Schnell and colleagues[Schnell et al., 2009]. A record of hospital treatment in one part of the country may be matched with a record of death in another part even with both municipalities being the same. The patient may have been travelling when he died or may have travelled to obtain treatment.

Figure 1 shows a typical distribution of scores obtained for a municipality, scores of 9.000 or above were assumed to represent true matches. When hospital records are matched with records of death we do not have to worry about whether having chosen the best match we may not at the same time fail to consider other relevant matches. A hospital record should only be matched with one death record, if it exists.

3. Tabulation of the Results.

The main result is a collection of matched records, the first a record of death and the second a hospital separation where it is recorded that the patient died. The cause of death in both systems is coded according to a table with more than 14.400 entries[Organization,]. These were transformed into broad disease groups defined by Mathers and colleagues[Mathers et al.,]. The meaning of the mnemonics of the disease groups can be seen in Table 1.

Table 2 is a contingency table for the disease groups for the Northeast Region. We observe that 24.829 deaths were classified as being due to cancer in both systems. The plot of the two first dimensions of a correspondence analysis of this table shown in Figure 2 suggests that among the associations those for cancer and external causes are the strongest[Lê et al., 2008]. But unfortunately the off–diagonal frequencies are not small. For example 10.501 *cardvasc* death records were matched with hospital records where the cause of death was *comm_mat*. Rather than answering a question another one has appeared which may only be settled with extensive fieldwork.

Table 3 raises another question. The place of death may be obtained from the death record and it may be observed that 425.613 deaths occurred in hospital but were not matched with hospital records. A proportion,

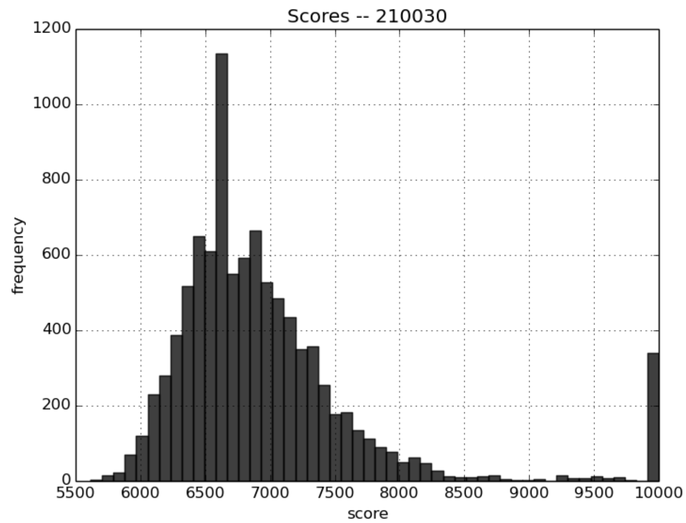

Figure 1: Typical Distribution of Values of the Dice Coefficient.

Mnemonic	**Disease Group**
cancer	Malignant neoplasms
cardvasc	Cardiovascular diseases
comm_mat	Communicable, maternal, perinatal and nutritional conditions
diabetes	Diabetes mellitus
external	Injuries
other_chronic	Chronic Disease not Included in the Other Groups
resp	Respiratory diseases
unclassified	Hospital codes that are not used in mortality records and, rarely, mortality codes that are specific to Brazil
undefined	Symptoms, signs and ill-defined condition

Table 1: Description of the mnemonics of the major disease groups.

AIH Cause of Death	cancer	cardvasc	comm_mat	diabetes	external	other_chronic	resp	unclassified	undefined
SIM Cause of Death									
cancer	24829	3930	8565	245	250	6540	8704	610	4170
cardvasc	465	50101	10501	1018	493	8555	9301	1472	8751
comm_mat	357	6159	32842	430	328	4545	7083	617	4024
diabetes	99	6414	5158	4801	168	2895	2094	413	1780
external	164	2439	1311	57	350	1480	1549	11157	2321
other_chronic	949	7151	13233	498	355	18807	4975	1097	5671
resp	149	3071	6294	236	132	1024	7271	242	1495
unclassified	0	12	14	3	0	9	1	7	12
undefined	105	1188	839	97	32	685	471	279	819

Table 2: Distribution of disease groups among records of death that have been matched with hospital records where a death is recorded, for the years 2008–2012 for the Northeast Region.

perhaps large, of these records may be related to treatment obtained in private hospitals. But an estimate of the reasonableness of his figure will have to be obtained as well as a decision as to whether the distribution of deaths in private hospitals is the same.

We may observe in Table 4 that the disease groups frequencies are different according to whether they have

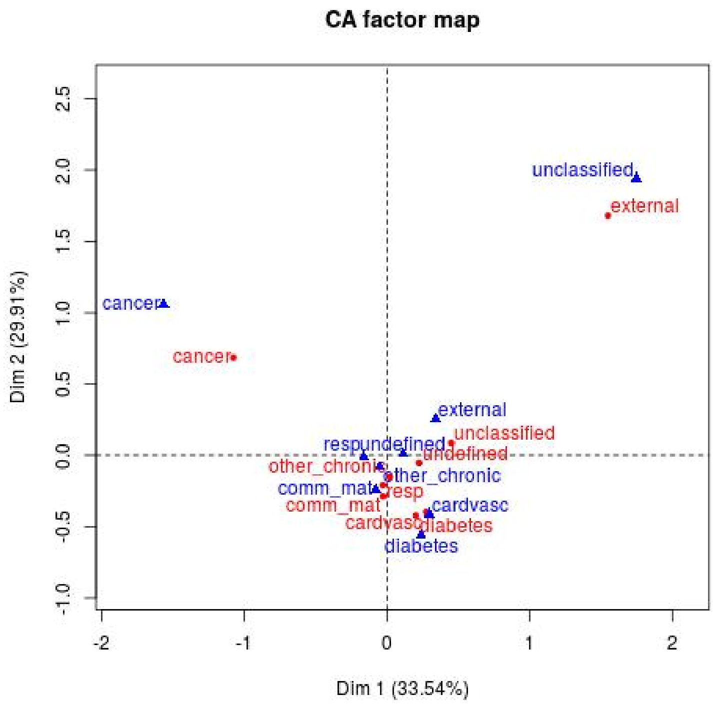

Figure 2: Plot of the first two dimensions of a correspondence analysis of Table 2. The row variables are in red.

Place of Death	**Unmatched**	**Matched**
Hospital	425613	467562
Other Health Establishment	10795	3894
Home	287103	139713
Public Place	94444	10965
Others	42904	7128
Unknown	5231	948

Table 3: Place of Death versus “Matched with Hospital Record of Death”.

or not been matched with a hospital record. Questions arise, for example, as to whether we are as likely to identify cancer among the unmatched records.

The box plots of Figure 3 may offer suggestions as to how to deal with *ill_defined* causes of death. For each of the nine states of the region the percentage was obtained of *ill_defined* death records matching hospital records for the different disease groups. For example, a mean of approximately 25% of *ill_defined* death records were matched with *cardvasc*. The current suspicion that *cancer* and *external* causes are unlikely to be missed, at least in the context that generated the data, seems reasonable[Silva et al., 2011]. To these two *diabetes* may be added.

The percentage of *ill_defined* mortality records that are also *ill_defined* in the hospital records is not small and it is possible that some sort of supplementary generic rule will end up being devised to take this into account.

4. Conclusions

Information about hospital mortality may improve overall mortality figures. As the logical consequences of each new tabulation are appreciated new questions will require answers. In the end the improvement may

matched	**Unmatched**	**Matched**
who_sim		
cancer	72187 (39)	111216 (61)
cardvasc	234193 (56)	184171 (44)
comm_mat	142312 (62)	86605 (38)
diabetes	39446 (44)	49987 (56)
external	172228 (82)	38991 (18)
other_chronic	93017 (51)	89095 (49)
resp	35511 (50)	36166 (50)
unclassified	61 (40)	93 (60)
undefined	77135 (69)	33886 (31)

Table 4: Distribution of disease groups in mortality records according to whether or not they were matched, row percentages in brackets.

not be a number with a desirable variance but knowledge of what is really going on in the health service shared by a wider group of people.

References

[França et al., 2014] França, E., Teixeira, R., Ishitani, L., Duncan, B. B., Cortez-Escalante, J. J., Morais Neto, O. L. d., and Szwarcwald, C. L. (2014). Ill-defined causes of death in Brazil: a redistribution method based on the investigation of such causes. *Revista de Saúde Pública*, 48:671 – 681.

[L. et al., 2014] L., S. C., de Frias P. G., deSouza, J. P. R. B., da Silva de Almeida W., and de M., N. O. L. (2014). Correction of vital statistics based on a proactive search of deaths and live births: evidence from a study of the North and Northeast regions of Brazil. *Population Health Metrics*.

[Lê et al., 2008] Lê, S., Josse, J., and Husson, F. (2008). FactomineR: An R package for multivariate analysis. *Journal of Statistical Software.*, 25.

[Mathers et al.,] Mathers, C., Bernard, C., Iburg, K., Inoue, M., Fat, D., Shibuya, K., Stein, C., Tomijima, N., and Xu, H. Global burden of disease in 2002: data sources, methods and results.[http://www.who.int/healthinfo/paper54.pdf].

[Organization,] Organization, W. H. *ICD-10 : international statistical classification of diseases and related health problems*.

[Schnell et al., 2009] Schnell, R., T., B., and J., R. (2009). Privacy-preserving record linkage using Bloom filters. *BMC Med Inform Decis Mak*.

[Silva et al., 2011] Silva, G. A. e., Gamarra, C. J., Girianelli, V. R., and Valente, J. G. (2011). Tendência da mortalidade por câncer nas capitais e interior do Brasil entre 1980 e 2006. *Revista de Saúde Pública*, 45:1009 – 1018.

[Stevens, 2015] Stevens, A. (2015). Relatório contendo resultados de estudo que identifica as causas de internação das pessoas que foram a óbito e que constam com código mal-definido ou indeterminado, Brasil, Regiões, UF, 2008 a 2012. Contrato OPAS Número BR/CNT/1400502.001. CGIAE/SVS – Ministry of Health.

[Stevens et al., 2014] Stevens, A., Schmidt, M. I., and Duncan, B. B. (2014). Information-processing methods for mortality surveillance in the presence of varying levels of completeness and ill-defined codes of causes of death – the case of Brazil. *Population Health Metrics*.

[Teixeira et al., 2006] Teixeira, C. L. d. S., Klein, C. H., Bloch, K. V., and Coeli, C. M. (2006). Reclassificação dos grupos de causas prováveis dos óbitos de causa mal definida, com base nas Autorizações de Internação Hospitalar no Sistema Único de Saúde, Estado do Rio de Janeiro, Brasil. *Cadernos de Saúde Pública*, 22:1315 – 1324.

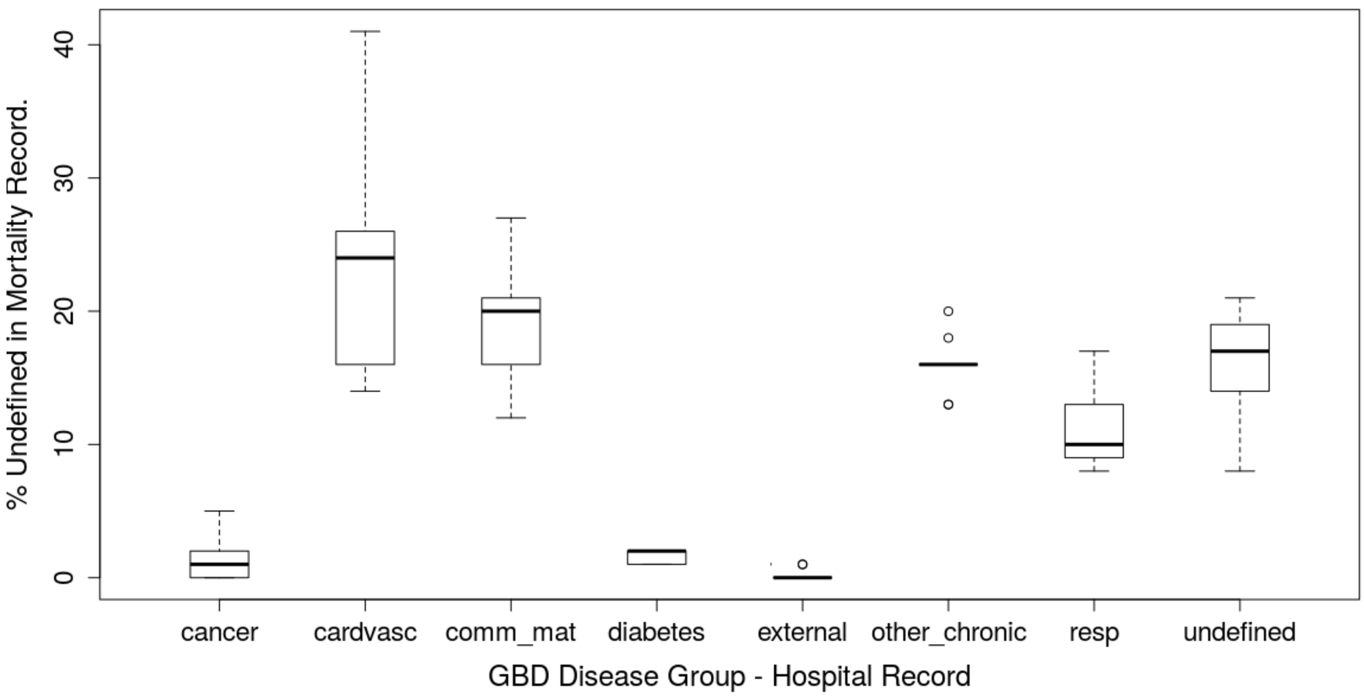

Figure 3: Box Plots of the percentage of *ill_defined* mortality records matched with the different disease groups in the hospital records. Each plot is based on nine readings, one for each state in the region.