Ana Nilce Silveira Maia-Elkhoury[I]

Eduardo Hage Carmo[I]

Marcia Leite Sousa-Gomes[I]

Eduardo Mota[II]

Analysis of visceral leishmaniasis reports by the capture-recapture method

ABSTRACT

OBJECTIVE: To analyze the number of cases, deaths, incidence and fatality rate due to visceral leishmaniasis, and to estimate its underreporting, as well as the coverage of the national information systems.

METHODS: Confirmed cases of visceral leishmaniasis were analyzed, based on the following systems: the *Sistema de Informação de Agravos de Notificação* (SINAN – Information System on Disease Notification), the *Sistema de Informações sobre Mortalidade* (SIM – Mortality Information System) and the *Sistema de Informações Hospitalares* (SIH – Hospital Information System), between 2002 and 2003. The variables utilized in relationship for pair identification were: patient's name, mother's name, date of birth, gender, city of residence, and mailing address. The capture-recapture method was applied to calculate the estimates, by means of the Chapman formula.

RESULTS: The estimated underreporting of visceral leishmaniasis in the SINAN, in relation to the SIH and the SIM, was 42.2% and 45.0% respectively. The estimated underreporting of deaths was 53% and 46.5%, when compared to SINAN-deaths and SIH-deaths respectively. The estimated incidence was 2.9 per 100,000 inhabitants, from the comparison between the SINAN and the SIH, 70.5% higher than the one found when SINAN's data were the only ones utilized. Furthermore, when comparing data from SIM and SINAN-deaths, an estimated fatality rate of 8% was observed, representing an increase in 16% from the one initially registered in the SINAN-deaths.

CONCLUSIONS: The results show high estimated underreporting of cases and deaths due to visceral leishmaniasis in Brazil. The relationship between information systems and the capture-recapture method application enabled to know and improve the epidemiological estimates, making its utilization in health services feasible.

KEY WORDS: Leishmaniasis, visceral, epidemiology. Capture-recapture. Disease notification. underregistration. Information systems.

[I] Departamento de Vigilância Epidemiológica. Secretaria de Vigilância em Saúde. Ministério da Saúde. Brasília, DF, Brasil

[II] Instituto de Saúde Coletiva. Universidade Federal da Bahia. Salvador, BA, Brasil

Correspondence:
Ana Nilce Silveira Maia Elkhoury
SCS Quadra 4 Bloco A 2º Andar – Edifício Principal
70304-000 Brasília, DF, Brasil
E-mail: ana.elkhoury@saude.gov.br

INTRODUCTION

In Brazil, visceral leishmaniasis (VL) is endemic in 20 states, with 3,380 new cases, an incidence of 2/100,000 inhabitants and a fatality rate of 5.3% having been recorded annually in average, in the period between 1994 and 2003.

The availability of high-quality information on a continuous and systematic basis, in other words, far-reaching and reliable, is fundamental for VL surveillance and control. Expressive underreporting of cases referring to the different types of injury and diseases that comprise the information system of the *Sistema Único de Saúde* – SUS (National Health System) has been related. Studies

show that this underreporting varies among the different information systems, *,** of which the *Sistema de Informação de Agravos de Notificação* – SINAN (Information System on Disease Notification), the *Sistema de Informações sobre Mortalidade* – SIM (Mortality Information System) and the *Sistema de Informações Hospitalares* – SIH (Hospital Information System) are the most important and of interest to epidemiological surveillance. The SINAN gathers, transmits and disseminates data generated from the notification of cases by the epidemiological surveillance system on the three levels of government. The SIM collects data from death certificates, whereas the SIH registers data on hospitalizations conducted in the public health system and in the contracted-out system, representing from 70% to 80% of these events in Brazil.[3]

In the last years, SUS administrators have assessed the quality of data and information, due to its importance in the definition of public policies, planning, and decision-making. The information systems available have shown themselves to be important tools, especially when related to each other and when joint data analysis is performed, thus aiming at complementing and expanding knowledge about diseases and types of injury of public health interest.[11]

In the assessment of the several information systems that deal with the same type of injury or disease, the capture-recapture method has recently been employed in epidemiological studies. This method enables to assess the system's coverage, estimate underreporting, and thus calculate and correct epidemiological indicators.[1,5]

The capture-recapture method has been utilized since the 1930's in the field of ecology to assess and estimate animal populations, and subsequently in the field of demography to make population estimates.[6,10] As this method is considered to be a cheap and effective alternative to improve monitoring of diseases in the population, it has been used as a tool to promote increase in the coverage of health information systems, to reduce costs in order to estimate indicators, to assess the functioning of the services, and to implement its actions.[5,9]

The present study aimed to analyze the number of cases, deaths, incidence and fatality rate due to visceral leishmaniasis and to estimate its underreporting, as well as the coverage of the national information systems.

METHODS

The VL records present in the SINAN, SIM, and SIH for the Brazilian population, between 2002 and 2003, were utilized. The research covers all the states that had confirmed record of cases in the Sinan. The records of hospitalizations and deaths were in accordance with the International Classification of Diseases (ICD-10) and considered: visceral leishmaniasis (B55.0) or nonspecific leishmaniasis (B55.9) with subsequent codes that indicated VL, such as hepatosplenomegaly, hemorrhagic signs, malnourishment, anemia, septicemia, pneumonia, and splenomegaly followed by hemorrhage.[4,12]

The population studied comprises all the cases that are registered in at least one of the information systems, after data preparation: removal of duplicities and re-hospitalizations and selection of data on spreadsheets. From SINAN and SIH, VL cases that led to death constituted the groups named SINAN-deaths and SIH-deaths respectively. The total number of records analyzed was 9,481, which represented 64.7% of the total number of confirmed cases. The RecLink II program[2] was utilized to compare records and identify cases that were common among the sources, matching them in pairs: SINAN x SIH, SINAN x SIM, SIM x SINAN-deaths, and SIM x SIH-deaths. The variables utilized in the relationship for pair identification were: patient's name, mother's name, date of birth, gender, city of residence, and mailing address.

To estimate cases and deaths, the capture-recapture method was applied, using the Chapman formula to calculate the N value, its variance (estimated N) and the confidence interval of 95%.[13]

$$N = \frac{(R + 1)\,(S + 1)}{m + 1} - 1$$

$$Var \text{ (estimated N)} = \frac{(R + 1)\,(S + 1)\,(R - m)\,(S - m)}{(m + 1)\,(m + 1)\,(m + 2)}$$

$$95\% \text{ CI} = [N \text{ (estimated)} \pm 1,96 \sqrt{var \text{ (estimated N)}}].$$

From these estimates, the coverage and magnitude of underreporting in the information systems were calculated. The coefficient of incidence was also calculated, based on the cases found in the SINAN and the estimated country population for the year 2003. Regarding fatality rate, the number of deaths found in the SINAN-deaths was divided by the number of SINAN cases and multiplied by 100. From the estimated values of cases and deaths, involving the relationship SINAN x SIH, as well as SIM x SINAN-deaths, the indicators were recalculated with the same criteria, though the number of cases and deaths was divided by two, because the cases referred to a two-year period.

The study was approved by the Comitê de Ética e Pesquisa do Instituto de Saúde Coletiva da Universidade Federal da Bahia (Research and Ethics Committee of the Institute of Public Health at the Federal University of Bahia), on 5/8/2005, under register nº 036-05/CEP-ISC.

RESULTS

The VL records for the period studied were distributed according to the data source, before and after the preparation of these sources to perform the capture-recapture analysis (Table 1). A total of 10.9% of all the SIM records was included, as these met the criteria established for the basic causes of non-specific leishmaniasis.

In the gathering analysis of VL cases, by connecting the sources in pairs (Figure), it can be observed that 7,382 cases were identified in the three sources assessed out of the 9,481 cases that were eligible for capture-recapture analysis in the different sources. A total of 5,344 cases (72.4%) out of 7,382 were captured by each of the sources only once: 3,897 by the SINAN, 1,312 by the SIH and 135 by the SIM. A total of 2,038 records, that is, 27.6%, were captured simultaneously by more than one source.

A total of 1,486 records (20.1%) out of the 7,382 VL cases found in the three data sources analyzed were not present in the SINAN: 135 of these were only in the SIM; 1,312 were in the SIH; and 39 were registered simultaneously in the SIM and in the SIH.

It was observed that 61 cases out of all the different VL records in the three information systems corresponded to the connection of the three sources assessed, which means 0.8% of the VL records were simultaneously present in the SINAN, SIH and SIM.

**Table 1.** Visceral leishmaniasis cases according to data source. Brazil, 2002-2003.

Data source	Initial cases N	Final cases* N	% variation**
SINAN	10,516	5,896	↓ 44.0
SIH	3,805	3,198	↓ 16.0
SIM	349	387	↑ 10.9
Total	14,670	9,481	↓ 35.3

* Final cases correspond to data after preparation of sources to be utilized in the analysis of the capture-recapture technique. Percentage values and arrows indicate variation between "initial cases" and "final cases"
** ↑ positive variation, ↓ negative variation
SINAN: *Sistema de Informação de Agravos de Notificação* (Information System on Disease Notification)
SIM: *Sistema de Informação sobre Mortalidade* (Mortality Information System)
SIH: *Sistema de Informações Hospitalares* (Hospital Information System)
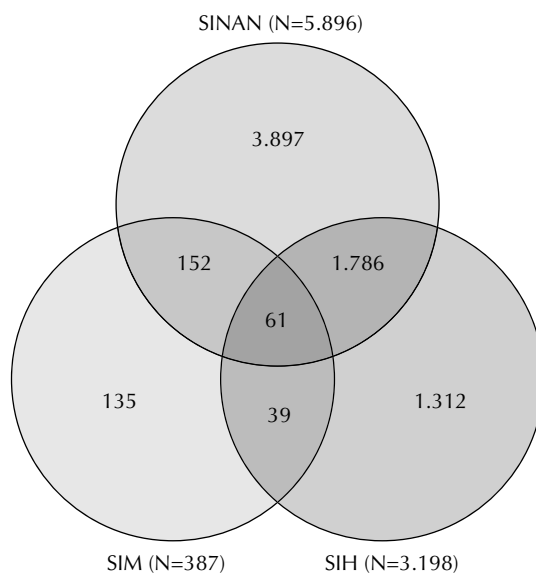
SINAN: *Sistema de Informação de Agravos de Notificação* (Information System on Disease Notification)
SIM: *Sistema de Informação sobre Mortalidade* (Mortality Information System)
SIH: *Sistema de Informações Hospitalares* (Hospital Information System)

Figure. Number of visceral leishmaniasis cases, according to their capture by the data sources in the SINAN, SIH and SIM. Brazil, 2002-2003.

Through the relationship between the different data sources and the finding of common records among the sources, utilizing the simple probabilistic model (Chapman formula), the VL cases and deaths were estimated (Table 2).

Regarding the cases, point estimates of 10,207 cases were found when the SINAN x SIH sources were assessed, and of 10,691 cases for the SINAN x SIM sources. In relation to deaths, the estimates found were 824, when the SIM x SINAN-deaths was analyzed, and 723 for the SIM x SIH-deaths. The N variance (estimated) and respective confidence intervals were obtained for both cases and deaths.

After analysis of the sources in pairs, taking into account the fact that the system analyzed for the cases was the SINAN and for the deaths was the SIM, it was possible to estimate the coverage for each one of the information systems (Table 3). For the assessment of cases, the SINAN showed the best rates of coverage when compared to the other systems. Regarding deaths, the SINAN-deaths presented coverage that was similar to the SIM's.

Through the coverage of the information systems, the magnitude of underreporting of SINAN cases in relation to the SIH and SIM was estimated (Table 4). It was observed that the underreporting was high for

Table 2. Estimate of cases and deaths due to visceral leishmaniasis, according to data source, compared in pairs. Brazil, 2002-2003.

Data source	Estimates of cases and deaths*		
	N (estimated)	Variance (N)	CI 95%
SINAN x SIH	10,207	16,342	9,957;10,457
SINAN x SIM	10,691	229,786	9,752;11,630
SIM x SINAN-deaths**	824	953	764;884
SIM x SIH-deaths**	723	2,958	616;830

* Calculations performed by means of the Chapman formula.
** Group of visceral leishmaniasis cases that progressed to death.

Table 3. Coverage of information systems, according to data sources. Brazil, 2002-2003.

Data source	System coverage	
SINAN x SIH	SINAN	SIH
	57.8%	31.3%
SINAN x SIM	SINAN	SIM
	55%	3.61%
SIM x SINAN-deaths*	SIM	SINAN deaths
	47%	49%
SIM x SIH-deaths*	SIM	SIH deaths
	53.5%	17.2%

* Group of visceral leishmaniasis cases that progressed to death.

Table 4. Underreporting of visceral leishmaniasis, according to data source. Brazil, 2002-2003.

Data source	Estimated underreporting for visceral leishmaniasis	
	SINAN (%)	SIM (%)
SINAN x SIH	42.2	-
SINAN x SIM	45.0	-
SIM x SINAN deaths*	-	53.0
SIM x SIH deaths*	-	46.5

* Group of visceral leishmaniasis cases that progressed to death.

both sources, with percentages of 42.2% and 45.0% respectively. However, the estimated underreporting of SIM deaths, when compared to SINAN-deaths (53.0%), was greater than the one found when SIM x SIH-deaths (46.5%) were related.

From the estimates of VL cases and deaths in Brazil, it was possible to correct the main epidemiological indicators for the year 2003. It was observed that the coefficient of incidence varied from 1.7 cases/100,000 inhabitants to 2.9 cases/100,000 inhabitants. Regarding deaths, after recalculating the fatality rate, it was possible to verify an increase from 6.9% to 8.0%.

DISCUSSION

The present study identified the high percentage of underreporting of VL cases and deaths in Brazil. It was verified that 45% of deaths captured by the SIM as basic cause of VL were not notified in the SINAN. The same thing occurred in the SIH, as 42.2% of hospitalizations paid by the *SUS* and classified as VL (B55.0) were not included in the cases notified in the SINAN either.

The high underreporting of cases does not differ from other studies on assessment of disease records with obligatory notification. The underreporting of AIDS cases was 42.8% in the city of Rio de Janeiro, 46.3% in the city of Belo Horizonte, and 17.3% in the city of Recife.[8],*,** Together, these studies show the need to develop analyses on the magnitude of diseases, utilizing different information systems.

The percentage of underreporting found is alarming, because it does not refer to administrative information exclusively, but rather having direct implications for VL control and surveillance practices in Brazil. Since 2003, the recommendation of such practices has been based on the epidemiological situation of each area, classified as of sporadic, moderate or intense transmission.***

Regarding the quality of SINAN, SIM and SIH records, some studies showed that these information sources are recommended for application of the capture-recapture method to assess information referring to *diabetes mellitus* and AIDS.**,**** More recently, epidemiologists have improved its use to estimate prevalences and incidences of diseases in the population. Nonetheless, the utilization of this methodology in epidemiological research has limitations, as it is necessary to maintain some premises to obtain valid estimates such as: fixed population, information quality, valid diagnosis, and independence among factors.

* Brito AM, Menezes AB, Abafgaro ACP. Subnotificação para a AIDS e Tuberculose em Doentes com AIDS Internados no Hospital Correia Picanço, Recife, no Ano de 1994. In: IV Congresso Brasileiro de Epidemiologia, 1998, Rio de Janeiro. Livro de Resumos: Revista Brasileira de Epidemiologia, Rio de Janeiro: Armazém das Letras, 1998. p. 192
** Oliveira MTC. A subnotificação de casos de aids em Belo Horizonte, Minas Gerais: uma aplicação da técnica de captura-recaptura [master's dissertation]. Belo Horizonte: Faculdade de Medicina da Universidade Federal de Minas Gerais; 2000.
*** Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Manual de vigilância e controle da leishmaniose visceral. Brasília: Ministério da Saúde; 2003. (Série A. Normas e Manuais Técnicos).
**** Coeli CM Vigilância do Diabetes Mellitus em uma população idosa: aplicação da metodologia de captura-recaptura [doctorate thesis]. Rio de Janeiro: Instituto de Medicina Social da Universidade do Estado do Rio de Janeiro; 1998.

In the present study, the premises could be fulfilled in their most part, as a fixed population was utilized in a short period of time, without any change or variation. The utilization of the variables patient's name, date of birth, gender, mother's name, city of residence and mailing address enabled the information systems to be related to each other and the pairs to be identified, with the exclusion of those pairs that were doubtful. Furthermore, the utilization of the probabilistic record linkage program (RecLink) minimized the possibility to include possible false pairs in the analysis.[2]

By assessing the independence among sources, it has been verified that, in the manner they were used in the present study, these sources are generally independent in the capture-recapture method, as it is not possible to determine the level of dependence among them when only two sources are related to each other.[7]

One possible limitation of the present study refers to the VL diagnosis accuracy in the SIH records, as in the SINAN only the cases that met the confirmation criteria were included. In the SIM, deaths that had VL as their basic cause remained and the associated causes confirmed the disease. Regarding the SIH information, there was no guarantee that the diagnosis was actually VL, because in many situations only the initial diagnosis is registered in the *Autorização de Internação Hospitalar – AIH* (Authorization for Hospital Admittance), and is not altered afterwards. Thus, a limitation is observed in relation to the capture-recapture method application for the constant information in the SIH without an investigation and confirmation of the aggravation diagnosis.

Some difficulties arose in the utilization of the methodology applied. One of them refers to the different variables available in the information systems and necessary for the relationship of the sources. A total of 12% of the constant records in the SINAN were excluded, with the variable "final classification" in blank or ignored. A high number of records in the SIM (16%) and in the SIH (15%) with the ICD code for non-specific leishmaniasis could not be included in the study as they did not meet the established criteria. In some studies conducted for quality assessment of the SINAN records concerning other diseases with obligatory notification, though the same methodology was not utilized, high percentages of the variable "final classification" without completion were found, such as 18.5% for meningitis* and 44.0% for dengue fever.**

These information systems have different objectives and logic,[3] and it is necessary to adopt strategies that would enable a relationship among databases. Thus, a higher proportion of true pairs would be guaranteed, such as the utilization of the patient's name and the patient's mother's name. However, difficulties concerning data collection were identified, especially when it refers to the SIH, where the patient's mother's name is not a variable available in the system, as the variable "name of the responsible adult" is utilized. These difficulties can partly explain the high percentage of excluded records (35%) for the capture-recapture analysis.

Another aspect to be considered is that the variable case "evolution" is still not characterized in the SINAN if the death occurred as a result of VL or due to any other cause, such as an accident. In this case, a limitation can be assumed regarding the quality of information for mortality. However, this is not the only limitation, as several studies point to deficiencies in the quality of death certificate completion and in the establishment of the basic causes, thus representing a major challenge for the mortality system.[14]

In spite of these limitations, guarantee of quality of information was the choice, aiming to maintain the premises to obtain valid estimates, in this case, to have an efficient database to relate cases in the different sources. In a way, the exclusion of some of these records may have influenced the results, altering the estimates for cases and deaths, underreporting and coverage of systems. This is because possible pairs of cases and deaths were not informed, thus decreasing the number of common records between two sources analyzed, and increasing, as a result, the estimated N.

In the present study the analysis performed to correct the main epidemiological indicators pointed to an increase in the incidence coefficient and fatality rate for VL in Brazil. The increase in fatality rate occurred due to underreporting of deaths in the SIM, which was higher than the one referring to cases in the SINAN. In general, the correction of epidemiological indicators may suggest to health administrators and the surveillance service the need to prioritize and readapt the actions for the improvement of assistance, surveillance and control of a disease or aggravation.

Epidemiological surveillance systems must seek alternatives to decrease underreporting of the information systems. New studies are needed to understand possible differences, to know and correct, if necessary, the databases, and consequently, to improve their quality so that the information utilized by administrators is efficient in terms of defining public policies, planning and decision making.

---

* Teles, MDS. Análise da qualidade dos registros do sistema de informação de agravos de notificação para meningites-Salvador-BA [master's dissertation]. Salvador: Universidade Federal da Bahia; 2004.
** Toledol ALA, Escosteguy CC, Medronho RA, Andrade FC, Oliveira KB, Valência LI. Qualidade do SINAN – Dengue: Consistência dos dados e do diagnóstico final na epidemia 2001-2002, Município do rio de Janeiro. In: VI Congresso Brasileiro de Epidemiologia, 2004, Recife. Livro de Resumos: Rev Bras Epidemiol.2004. Suplemento Especial (versão eletrônica).

REFERENCES

1. Brenner H. Use and limitations of the capture-recapture method in disease monitoring with two dependent sources. *Epidemiology*. 1995;6(1):42-48.

2. Camargo Junior, KR, Coeli, CM. Reclink: aplicativo para o relacionamento de bases de dados, implementando o método *probabilistic record linkage*. *Cad Saude Publica*. 2000;16(2):439-47.

3. Carvalho DM. Grandes sistemas nacionais de informação em saúde: revisão e discussão da situação atual. *Inf epidemiol SUS*. 1997;4:7-46.

4. Collin S, Davidson R, Ritmeijer K, Keus K, Melaku Y, Kipngetich S, et al. Conflict and kala-azar: determinants of adverse outcomes of kala-azar among patients in southern Sudan. *Clin Infect Dis*. 2004;38(5):612-9.

5. Corrão G, Bagnardi V, Vittadini G, Favilli S. Capture-recapture methods to size alcohol related problems in a population. *J Epidemiol Community Health*. 2000;54(8):603-10.

6. Darroch JN. The multiple-recapture census l: estimation of a closed population. *Biometrika*. 1958;45:343-59.

7. Dunn J, Andreoli SB. Método de captura e recaptura: nova metodologia para pesquisas epidemiológicas. *Rev Saude Publica*.1994;28(6):449-53.

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 Saude Publica*. 1999;15(2):317-24.

9. Laporte RE. Assessing the human condition: capture-recapture techniques. *BMJ*. 1994;308(6920):5-6.

10. Otis DL, Burnham P, White GC, Anderson DR. Statistical inference from capture data on closed animal populations. *Wildlife Monographs*.1978;62.

11. Scatena JHG, Tanaka OY. Utilização do Sistema de Informações Hospitalares (SIH-SUS) e do Sistema de Informações Ambulatoriais (SIA-SUS) na análise da descentralização da saúde em Mato Grosso. *Inf epidemiol SUS*. 2001;10(1):19-30.

12. Seaman J, Mercer AJ, Sondorp HE, Herwaldt BL. Epidemic visceral leishmaniasis in southern Sudan: treatment of severely debilitated patients under wartime conditions and with limited resources. *Ann Inter Med*. 1996;124(7):664-72.

13. Stroup DF. Special analytic issues. In: Teutsch SM, Churchill RE, editors. Principles and practice of public health surveillance. New York: Oxford University Press; 1994. p.136-49.

14. Vanderlei LC, Arruda BKG, Frias PG, Arruda S. Avaliação da qualidade de preenchimento das declarações de óbito em unidade terciária de atenção à saúde materno infantil. *Inf epidemiol SUS*. 2002;11(1):7-14.