11. Swedish Institute for Infectious Diseases Control (SMI). Respiratory Syncytial Virus. Epidemiologisk Årsrapport. 2004-2005. Swedish Institute for Infectious Disease Control (SMI).
12. Sten A. Influensa – information och kampanjer ska öka vaccinationstäckningen I riskgrupper. Swedish Institute for Infectious Disease Control (SMI). Smittskydd. 2005;4:10-11.
13. Dannetun E, Tegnell A, Normann B, Garpenholt O, Giesecke J. Influenza vaccine coverage and reasons for non-vaccination in a sample of people over 65 years of age in Sweden 1998-2000. Scand J Infect Dis. 2003;35(6-7)389-93.
14. Euler G, Bridges CB, Brown CJ, Lu PJ, Singleton J, Stokley S *et al*. Estimated Influenza vaccination coverage among adults and children –United states, September 1, 2004– January 31, 2005. MMWR. 2005;54(12)304-7.
15. Szucs T. The socio-economic burden of influenza. J Antimicrob Chemother. 1999;44 Topic B 11-15.
16. Carrat F, Sahler C, Rogez S, Leruez-ville M, Freymuth F, Le Gales C *et al*. Influenza burden of illness: estimates from a national prospective survey of household contacts in France. Arch Intern Med. 2002;162:1842-48.

ORIGINAL ARTICLES

Surveillance report

COMPLETENESS OF MALARIA NOTIFICATION IN THE NETHERLANDS 1995-2003 ASSESSED BY CAPTURE-RECAPTURE METHOD

S Klein[1], A Bosman[2]

In 1999 in the Netherlands, the duty to notify malaria was transferred from physicians to laboratories by the new Infectious Diseases Law. To evaluate the effect of this change, we aimed to estimate completeness of malaria notification in the Netherlands from 1995-2003. We calculated it relative to sentinel laboratory and hospital admission data. Using the two-source capture-recapture method (CRM), we estimated the total number of cases to assess the completeness relative to this number.
The completeness of notification relative to sentinel data was 18.2 % (95% CI of 15.7-20.7) from 1995-1998 and 56.4 % (95% CI of 47.0-65.8) for 2000-2003. The completeness relative to the number of malaria cases admitted to the hospital was 35.1 % for the period 1995-2003. The estimated numbers of cases of malaria between 1995 and 1998 were 3123 (95% CI of 2796-3449) and 5043 (95% CI of 4343-5742) between 2000 and 2003. The completeness relative to this numbers changed from 35.5 % (95% CI of 32.1-39.7) in 1995-1998 to 36.1 % (95% CI of 31.7-41.9) for the years 2000-2003. Laboratory-based notification has significantly increased the absolute number of malaria notifications, but there was no change in completeness relative to hospital admissions. The increase in estimated malaria cases may be artificial, due to the extent of violation of CRM requirements over the study period.

Euro Surveill 2005;10(10): 244-6 Published online October 2005
Key words : Capture-recapture, malaria, notification system

Introduction

Since the new infectious diseases law was implemented in the Netherlands on 1 April 1999, laboratories are legally obliged to report malaria cases to the Municipal Health Service (GGD). Before this time, notification was only the responsibility of physicians. To evaluate this structural change in the Dutch notification system, this study, carried out in September 2004, aimed to estimate the completeness of malaria notifications in the Netherlands from 1995-2003. In this context, completeness refers to the proportion of cases detected by the notification system. It is generally assumed that malaria, like many other infectious diseases, is underreported [1,2]. Van Hest et al. investigated a total number of 774 malaria cases (95% CI of 740-821) and a completeness of notification on 40.2 % in the Netherlands in 1996 using three-source CRM [3].

1. Hamburg University of Applied Sciences, Hamburg, Germany
2. RIVM, Division of Public Health, Centre of Infectious Disease Epidemiology (CIE), Bilthoven, The Netherlands

Methods

Data Sources, Case Definition and Matching Algorithms

The sentinel register (12 voluntarily reporting laboratories at the moment) included the variables month and year of birth, gender, postal code, place of residence and day of onset/date of diagnosis. The variables which were reported to the notification register included the date of notification, year of birth, date of diagnosis, date of onset, postal code, gender, reporting GGD, method of diagnosis, and the species of plasmodium. The Dutch morbidity registration organisation provided hospital admission data on principal diagnosis malaria (ICD-9 code 084* - * meaning all species of malaria) with the variables pathogen, date of admission, date of discharge, year of registration, year of birth, gender, postal code and place of residence after discharge. A case of malaria in this study was defined as a person with a positive blood smear for a plasmodium species.

We matched data first by using the following identifiers: year of birth, gender, year of diagnosis/request/admission (if missing: year of onset/sample) and 4-digit postal code, using an algorithm in MS-Excel®. To correct for late notification, we used safety margin of 30 days around the date of diagnosis in the GGD data and these matching pairs we reviewed manually.

We searched for additional matches in the remaining non-matched cases, using a second algorithm. This algorithm used the same identifiers, but without postal code, for the GGD records without a valid postal code (e.g. unknown, missing, abroad, homeless). To be confident we reviewed these matching pairs manually, comparing the date of diagnosis with the date of admission (+/- 3 days, hospital data) and the species (hospital data and laboratory data).

Completeness of notification

The completeness of notification was assessed by searching for cases which were also on the notification register as in the sentinel laboratory register or hospital admission register, respectively. To calculate the completeness (C) of notification relative to sentinel laboratory data we used the formula: $C = a/b*100\%$ where a is the

number of cases in both data sources (notification and laboratory data base) and b is the number of cases in the laboratory data base. The 95%CI of C is: C +/- 1.96* √a*(b-a)/b. To calculate the completeness of notification relative to hospital admission data we used the same formulas.

Two-source-capture-recapture Method (CRM)

CRM is adapted from biology and is applied on overlapping incomplete data sources. The two-source method is a relatively simple, feasible and reproducible method used to estimate the number of total cases, including the ones which were not observed, and subsequently to assess the completeness of the sources. We used the hospital admission data and the notification data to estimate the number of malaria cases.

On the basis of previous literature, Hook and Regal [4] conclude that the maximum likelihood estimator (MLE) (if the numbers are high enough) for the real number of cases (N) is: N = (a + b)*(a + c)/a where a is the number of cases in both registers and b and c are the numbers in only one of the registers. The 95% CI of N is: N +/- 1.96*√(a+b)*(a+c)*b*c/a3.

To calculate the completeness of the notification register relative to the estimated number of cases (C) we used the formula: C = (a+b)/N*100% where b is the number in only the notification register.

Results

Completeness of notification relative to sentinel laboratory detected cases and to hospital-determined cases.

The completeness of the notification relative to the laboratory diagnosed cases was 18.2 % (95% CI of 15.7-20.7) for the period from 1995-1998 and increased significantly to 56.4 % (95% CI of 47.0-65.8) for the years 2000-2003 [Table 1].

The completeness of notification relative to the cases admitted to the hospital was 35.7 % (95% CI of 17.7-53.7) for the period 1995-1998 and 37.7 % (95% CI of 21.3-54.0) for the period 2000-2003 [Table 2]. This change is not significant.

Completeness relative to estimated total number of cases

The first algorithm contributed 861 matching pairs, which was 84.9 % of the final matching pairs. The second algorithm contributed 153 (15.1 %) of all matched cases. Between 1995 and 2003, 2886 patients with malaria were admitted to hospital and 3382 cases were notified to the GGDs. 1014 of these cases could be found in both sources. These numbers gave a CRM estimate of 9626 (95% CI of 9226-10 025) malaria cases in these nine years, while 3123 (95% CI of 2796-3449) cases were estimated before 1999 compared to 5043 (95% CI of 4343-5742) cases after 1999. Based on the estimated numbers of total cases, as demonstrated in table 3, the completeness of notification increased minimally from 35.5 % (95% CI of 32.1-39.7) for the years 1995-1998 to 36.1 % (95% CI of 31.7-41.9) for the years 2000-2003.

Discussion

• The increase in estimated malaria cases is assumed to be artificial, due to that the introduction of the new law enhanced the violation of the basic assumptions underlying CRM [2,4,5]:

- Same case definition in each source: Hospital admission uses high severity of malaria case by implication. As the notification register also includes outpatients with a lower severity, the case definitions are different.
- Same probability to be ascertained for each case: Outpatients have no probability of being on the hospital admission register.
- Source independency: Since 1999 the dependency between cases notified from hospitals and those notified from laboratories was reduced.

Table 1

Number of sentinel laboratory detected malaria cases and completeness of notification relative to this number, the Netherlands, 1995-2003

Year of diagnosis	Laboratory-confirmed	Laboratory-confirmed and notified	Laboratory total	Completeness	Aggregated completeness (95% CI)
1995	3	1	4	25	18.2 (15.7-20.7)
1996	1	1	2	50	
1997	3	0	3	0	
1998	2	0	2	0	
1999	9	6	15	40	
2000	8	18	26	69.2	56.4 (47.0-65.8)
2001	11	16	27	59.3	
2002	15	10	25	40	
2003	7	9	16	56.3	
Total	**59**	**61**	**120**	**50.8**	

Table 2

Hospital admissions with malaria and completeness of notification relative to this number, the Netherlands, 1995-2003

Year of diagnosis	Hospitalised	Hospitalised and notified	Hospitalised total	Completeness	Aggregated completeness (95% CI)
1995	243	124	367	33.8	35.7 (17.7-53.7)
1996	196	144	340	42.4	
1997	254	105	359	29.2	
1998	215	131	346	37.9	
1999	252	80	332	24.1	
2000	236	126	362	34.8	37.7 (21.3-54.0)
2001	185	125	310	40.3	
2002	179	66	245	26.9	
2003	112	113	225	50.2	
Total	**1872**	**1014**	**2886**	**35.1**	

- Accurate classification as a case: Positive predictive value is assumed to be high for malaria, due to inclusion of laboratory diagnostics in the case definition.
- Suitability of matching: Because of incomplete notification the identifiers have not always been available; also the reliability of the given identifiers is supposed not to be optimum and to diminish, leading to under-assessment of the number of matching pairs. It is not likely to confuse cases with other cases in a rare disease like malaria in the Netherlands.
- Closed study population: Stable 'catchability' of cases exists because it is not likely that one would travel in or out of the Netherlands while suffering from acute malaria. Fatality is also low. 'Catchability' between years may have changed because of risen number of immigrants from endemic countries.

• Laboratory-based notification has significantly increased the absolute number of malaria notification in the Netherlands [Table 3]. The increase in the numbers of travellers to and immigrants from endemic countries cannot explain this increase.

• No change in completeness of notification relative to hospital admission data was observed [Table 2]. We therefore conclude that the increase of notified cases was mainly due to non-hospitalised cases.

• The overestimation of the unobserved number of cases after 1999 creates the impression of a low completeness of notification by the laboratories.

Over-all Conclusion

The number of malaria cases and incidence is still much higher than notified, but it is likely that this study is overestimating the number of cases. Even if we consult only the number of cases which are recorded by at least one source, the surveillance system of malaria notification does not provide a realistic description of the incidence in the Netherlands.

Regarding CRM, the violation of the basic assumptions underlying the method leads to the overestimation of malaria cases and even a three-source investigation could not estimate the number of total cases because of a high dependency between notification register and laboratory reports after 1999.

Recommendations

• In order to facilitate the CRM as a tool in evaluating surveillance systems in general, we would recommend the reintroduction of common personal identifiers in the malaria reporting system. This might also be a benefit in other surveillance systems.

• A more complete evaluation of malaria surveillance based upon the CDC Guidelines would facilitate future CRM studies, by providing answers to some important questions on data quality, sensitivity and specificity that arose in our study.

Acknowledgements

We thank Rob van Hest, Mary Ward and Sabine de Greeff for their suggestions.

References

1. Thacker S B, Choi K, Brachmann PS. The surveillance of infectious diseases. JAMA. 1983 Mar 4;249(9):1181-5.
2. Grein T in: Kraemer A, Reintjes R. Infektionsepidemiologie. Methoden, Surveillance, Mathematische Modelle, Global Public Health. Springer; Berlin 2003:123-127.
3. van Hest N A H, Smit F, Verhave J P. Underreporting of malaria incidence in the Netherlands: results from a capture-recapture study. Epidemiol Infect. 2002 Oct;129(2):371-7.
4. Hook E B, Regal R R. Capture-Recapture Methods in Epidemiology: Methods and Limitations. Epidemiol Rev. 1995;17(2):243-64.
5. Reintjes R, Termorshuizen F, van de Laar M J W. Assessing the sensitivity of STD surveillance in the Netherlands: an application of the capture-recapture method. Epidemiol. Infect. 1999 Feb; 122(1): 97-102.

Table 3

Distribution of observed malaria cases, estimated number of unobserved cases, total number of estimated cases and completeness, The Netherlands, 1995-2003

Year of diagnosis	Hospital admissions	Only H *	Both, H/N ¥	Only N	Notifications	No. of unobserved cases	CRM ± MLE # (95% CI)	Completeness (95% CI)	Aggregated CRM MLE and aggregated completeness (95% CI)
1995	367	243	124	194	318	380	941	33.8	3123 (2796-3449) 35.5 (32.1-39.7)
							(836-1046)	(30.4-38.0)	
1996	340	196	144	157	301	214	711	42.4	
							(647-774)	(38.9-46.5)	
1997	359	254	105	122	227	295	776	29.2	
							(685-868)	(26.2-33.2)	
1998	346	215	131	132	263	217	695	37.9	
							(628-761)	(34.6-41.9)	
1999	332	252	80	375	455	1181	1888	24.1	
							(1561-2216)	(20.5-29.1)	
2000	362	236	126	411	537	770	1543	34.8	5043 (4343-5742) 36.1 (31.7-41.9)
							(1353-1733)	(31.0-39.7)	
2001	310	185	125	419	544	620	1349	40.3	
							(1189-1509)	(36.0-45.8)	
2002	245	179	66	331	397	898	1474	26.9	
							(1196-1751)	(22.7-33.2)	
2003	225	112	113	227	340	225	677	50.2	
							(605-749)	(45.4-56.2)	
Total	**2886**	**1872**	**1014**	**2368**	**3382**	**4372**	**9626**	**35.1**	
							(9226-10025)	**(33.7-36.7)**	

* Hospitalised ± Capture-recapture method
¥ Notified # Maximum likelihood estimator