

Effect of non-response bias on estimates of the numbers of patients with intractable diseases based on nationwide epidemiological surveys of Japan

Shuji Hashimoto, PhD¹, Miyuki Kawado, PhD¹, Takashi Kawamura, MD, PhD²,
Yosikazu Nakamura, MD, MPH³

¹Department of Hygiene, Fujita Health University School of Medicine, Toyoake, Aichi, Japan, ²Kyoto University Health Service, Kyoto, Japan, ³Department of Public Health, Jichi Medical University, Shimotsuke, Tochigi, Japan

Abstract

Objectives: Estimates of the numbers of patients with intractable diseases in Japan use data from nationwide epidemiological surveys under the assumption that the mean number of patients among hospitals that respond to the survey is equal to that among hospitals that do not respond ("the assumption of no response bias"). We examine the validity of this assumption.

Methods: Data from nationwide epidemiological surveys of eosinophilic granulomatosis with polyangiitis (EGPA) and familial Mediterranean fever (FMF) were used. Information from initial questionnaires and from second surveys, of hospitals that did not respond to the initial surveys, were combined in a statistical model to estimate rates of response.

Results: The proportion of hospitals that had patients with EGPA and FMF was higher among hospitals that responded to the initial survey than among those that responded to the second survey. The ratio of the response rate for hospitals without affected patients to that for those with affected patients was estimated to be 0.86 for EGPA and 0.90 for FMF. The ratio of the number of patients estimated under the assumption of no response bias to that based on the estimated ratio of response rate between in hospitals with and without affected patients was 1.11 for EGPA and 1.09 for FMF.

Conclusions: The number of patients with EGPA and FMF estimated under the assumption of no response bias was only about 10% greater than the true number, suggesting that the assumption was almost completely valid.

Keywords: Intractable diseases, Epidemiology, Prevalence, Bias

Introduction

Since 1972, the Japanese Ministry of Health, Labour and Welfare has promoted research into the causes and control of "intractable diseases"; diseases of unknown etiology for which no specific medical treatment is available.¹ Nationwide surveys have been carried out to investigate the epidemiological characteristics (e.g., prevalence, age distribution, and sex ratio) of many of these diseases.¹⁻⁴ The standard procedures of these surveys includes a postal survey that consists of two stages. The first stage is a survey of randomly selected Japanese hospitals, using a questionnaire that asks about the number of patients treated for a particular disease during the preceding year.⁵ The first stage of most of the nationwide epidemiological surveys of intractable diseases was carried out in two steps: a second questionnaire was sent to the hospitals that did not respond to the initial survey, to increase the total response rate.⁶ The response rate in most of these surveys was around 60%.^{7,8} The second stage is a survey limited to the hospitals that reported the number of patients, which uses a questionnaire to inquire about epidemiological and clinical information for individual patients.

The total numbers of patients treated for intractable diseases were estimated, using the data from these surveys,

under the assumption that the mean number of patients treated in hospitals that responded to the survey was equal to that in hospitals that did not respond. We will refer to this as "the assumption of no response bias". The validity of this assumption has been demonstrated for idiopathic thrombocytopenic purpura (ITP), in a study that compared data obtained from the nationwide epidemiological survey with the number of ITP patients who were financially subsidized for treatment⁹, but has not been investigated for any other diseases.

We investigated the validity of the assumption of no response bias by using a statistical model and data from the initial and second steps of nationwide epidemiological surveys of intractable diseases.

Methods

Data

We used data from the first stage of nationwide epidemiological surveys of eosinophilic granulomatosis with polyangiitis (EGPA) and familial Mediterranean fever (FMF), conducted by the Research Committee on Intractable Vasculitides and the Study Group of Familial Mediterranean Fever for Research on Intractable Diseases, respectively, in cooperation with the Research Committee on the Epidemiology of Intractable Diseases at the Japanese Ministry of Health, Labour and Welfare.^{3,4}

The initial survey of EGPA contacted 2,599 hospitals with departments of internal medicine, neurology and rheumatology/

Received 16 May, 2016. Accepted 3 August, 2016

Corresponding author : Shuji Hashimoto, PhD

Department of Hygiene, Fujita Health University School of Medicine, 1-98
Dengakugakubo, Kutsukake-cho, Toyoake, Aichi 470-1192, Japan

E-mail: hasimoto@fujita-hu.ac.jp

allergy in January 2009. The second step of this survey contacted the 1,426 hospitals that had not responded to the initial questionnaire in March 2009. A total of 956 patients with EGPA were reported.

The initial survey of FMF contacted 2,251 hospitals with departments of pediatrics, internal medicine and rheumatology/allergy in January 2010. The second step of this survey contacted the 1,188 hospitals that had not responded to the initial questionnaire was carried out in March 2010. In total, 170 patients with FMF were reported.

Statistical analysis

The two diseases were considered separately. The proportions of hospitals with patients with the diseases among those that responded to the initial and second surveys were calculated. The ratio of response rate between hospitals with and without affected patients was estimated using the statistical model described in the appendix. Under the assumption of no response bias, this ratio is one. The numbers of patients with each disease in some strata of hospitals were then estimated by two methods: the method that has been used in most of the

nationwide epidemiological surveys of intractable diseases, which relies on the assumption of no response bias^{7,8}; and the method described in the appendix, which is based on the estimated ratio of response rate between hospitals with and without affected patients. The strata of hospitals were the combinations of department type (internal medicine, neurology, and rheumatology/allergy for EGPA; pediatrics, internal medicine, and rheumatology/allergy for FMF), hospital type (general hospitals, university hospitals, or specialized hospitals), and the number of beds (20–99, 100–199, 200–299, 300–399, 400–499, or 500 or more) in general hospitals.^{3,4}

Results

Figures 1 and 2 show the distributions of the dates on which the hospitals responded to the nationwide epidemiological surveys of EGPA and FMF, respectively. Both distributions had two peaks, shortly after the postal questionnaires were sent to the hospitals in the initial and second steps of the surveys.

Table 1 shows the number of hospitals that were contacted

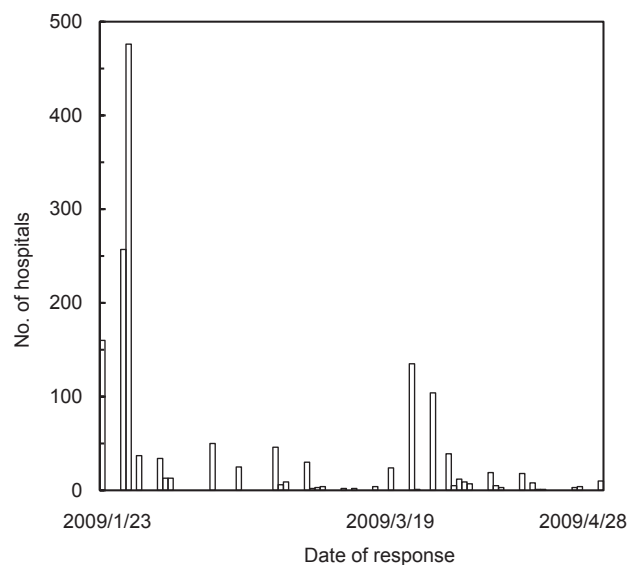
Table 1. Number of hospitals surveyed, the proportions that responded, and the estimated ratios of response rates between hospitals with and without affected patients in nationwide epidemiological surveys of eosinophilic granulomatosis with polyangiitis and familial Mediterranean fever

		Eosinophilic granulomatosis with polyangiitis			Familial Mediterranean fever		
		Number	% of total	% of responders	Number	% of total	% of responders
Initial survey of hospitals	Surveyed	2,599	100		2,251	100	
	Responded	1,173	45.1	100	1,063	47.2	100
	with patients	278	10.7	23.7	62	2.8	5.8
	without patients	895	34.4	76.3	1,001	44.5	94.2
Second survey of hospitals	Surveyed	1,426	100		1,188	100	
	Responded	408	28.6	100	340	28.6	100
	with patients	87	6.1	21.3	18	1.5	5.3
	without patients	321	22.5	78.7	322	27.1	94.7
Ratio of response rate in hospitals without affected patients to that for those with affected patients			0.86			0.90	

Table 2. Numbers of patients with eosinophilic granulomatosis with polyangiitis and familial Mediterranean fever estimated by two methods

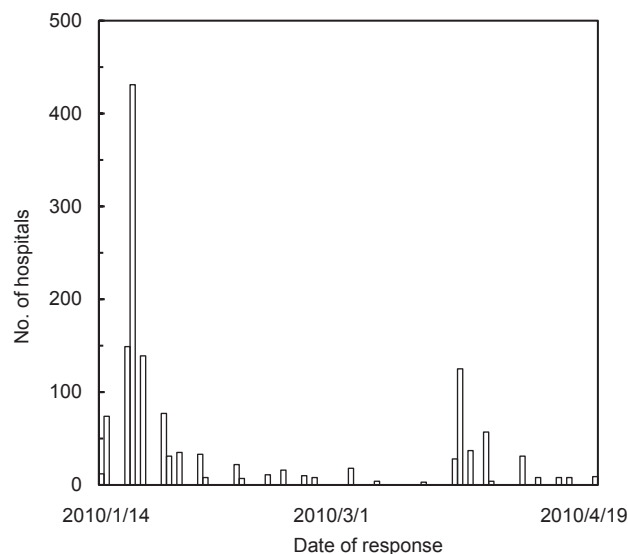
	Eosinophilic granulomatosis with polyangiitis	Familial Mediterranean fever
Reported no. of patients with the disease	956	170
Estimated no. of patients with the disease, under the assumption of no response bias	1,850	320
Estimated no. of patients with the disease, based on the estimate ratio of response rate between hospitals with and without affected patients	1,670	290
Ratio of the two estimates	1.11	1.09

Figure 1.



Distribution of the dates on which hospitals responded to the nationwide epidemiological survey of eosinophilic granulomatosis with polyangiitis.

Figure 2.



Distribution of the dates on which hospitals responded to the nationwide epidemiological survey of familial Mediterranean fever.

in and responded to each survey, and estimated ratios of the response rates between hospitals with and without patients with the diseases. The response rates in the initial and second surveys were 45.1% and 28.6% for EGPA, and 47.2% and 28.6% for FMF, respectively. The proportion of the hospitals that responded that had patients with the disease was higher in the initial survey than in the second survey for both EGPA and FMF. The estimated ratio of the response rate in hospitals without affected patients to that in those with affected patients was 0.86 for EGPA and 0.90 for FMF.

Table 2 shows the numbers of patients with EGPA and FMF estimated by the two methods. The number of patients with EGPA was estimated to be 1,850 under the assumption of no response bias, and 1,670 based on the estimated ratio of

response rate between hospitals with and without affected patients. The ratio of the former estimate to the latter one was 1.11. The numbers of patients with FMF estimated by these two methods were 320 and 290, and the ratio of two estimated numbers of patients was 1.09.

Discussion

There were higher proportions of hospitals with affected patients among those that responded to the initial surveys than in those that responded to the second steps of the nationwide epidemiological surveys of EGPA and FMF. This suggests that the response rates were higher in the hospitals with affected patients than in those without affected patients—a finding that is consistent with previous reports^{7,8}—and that the assumption of no response bias is unlikely to be completely valid.

In the present study, the ratio of the response rate in hospitals without affected patients to that in those with affected patients was estimated to be 0.86 for EGPA and 0.90 for FMF. The assumption of no response bias assumes this ratio to be one. The ratio of the estimated number of patients with the disease under the assumption of no response bias to that based on the estimated ratio of response rate was 1.11 for EGPA and 1.09 for FMF. Therefore, it was suggested that the estimated number of affected patients under the assumption of no response bias would be only about 10% greater than the true number.

Previously, there has been only one study on the validity of the assumption of no response bias in the nationwide epidemiological surveys of intractable diseases.⁹ That study, using the data of ITP patients who were financially subsidized for treatment (FST-patients), indicated that the mean number of ITP FST-patients among the hospitals that responded to the nationwide epidemiological survey of ITP was higher than the mean number across all hospitals. It also suggested that the ratio of the mean number in the hospitals that responded to the mean across all hospitals was between 1.0 and 1.1. Although not all patients with ITP are FST-patients, the previous study suggested that the number of patients with ITP estimated under the assumption of no response bias was between 1.0 and 1.1 times the true number, a relatively small bias. Although the assumption of no response bias is unlikely to be completely valid, the results of both the previous and present studies support the use of this assumption in estimation of the prevalence of many intractable diseases based on data from nationwide epidemiological surveys.

This study has some limitations. The data came from the first stage of two nationwide epidemiological surveys of EGPA and FMF. The data were slightly modified using the information obtained by the survey for individual patients in the second stage.^{3,4} Using the method with the assumption of no response bias, the numbers of patients with EGPA and FMF estimated by the modified data (1,870 for EGPA and 320 for FMF) were almost the same as our estimates. Additionally, these two surveys of EGPA and FMF may not be representative of the other surveys that are carried out. However, the response rate of the two surveys was about 60%, the normal level for such surveys of intractable diseases.^{7,8} The estimated number of patients was more than 1,000 for EGPA, but only about 300 for FMF.

We used the data of the initial survey and from the second survey, of hospitals that had not responded to the initial survey, to provide information to evaluate non-response bias in

the surveys.^{10,11} The essential assumption in the method used for estimating the ratio of response rate between hospitals with and without affected patients was that this ratio in the initial survey was equal to that in the second survey, a similar assumption to that used in a well-known multiplicative model.¹² The frequency of patients with intractable diseases is likely to vary with hospital characteristics, such as department type (internal medicine and pediatrics), hospital type (general hospitals and university hospitals), and the number of beds; the response rate of nationwide epidemiological surveys might also be expected to vary with these features.^{5,6,9} We did not consider these features in estimating the ratio of response rate between hospitals with and without patients with EGPA and FMF because of the small number of hospitals with such patients that responded to the surveys.

In conclusion, we suggest that estimates of the number of patients with EGPA and FMF under the assumption of no response bias will be about 10% greater than the true number, and that the assumption would be almost completely valid.

Acknowledgements

This study was supported by a Grant-in-Aid from the Ministry of Health, Labour and Welfare, Japan, for Research on Rare and Intractable Diseases.

Conflict of interests

The authors have no conflicting interests that could affect this work.

References

- Ohno Y, Kawamura T, Tamakoshi A, Wakai K, Aoki R, Kojima M, Lin Y, Hashimoto T, Nagai M, Minowa M. Epidemiology of diseases of unknown etiology, specified as "intractable diseases". *J Epidemiol* 1996;6(3 Suppl): S87-94.
- Kuriyama S, Kusaka Y, Fujimura M, Wakai K, Tamakoshi A, Hashimoto S, Tsuji I, Inaba Y, Yoshimoto T. Prevalence and clinicoepidemiological features of moyamoya disease in Japan: findings from a nationwide epidemiological survey. *Stroke* 2008;39:42-7.
- Migita K, Uehara R, Nakamura Y, Yasunami M, Tsuchiya-Suzuki A, Yazaki M, Nakamura A, Masumoto J, Yachie A, Furukawa H, Ishibashi H, Ida H, Yamazaki K, Kawakami A, Agematsu K. Familial Mediterranean fever in Japan. *Medicine (Baltimore)* 2012;91:337-43.
- Sada KE, Amano K, Uehara R, Yamamura M, Arimura Y, Nakamura Y, Makino H; Research Committee on Intractable Vasculitides, the Ministry of Health, Labour, Welfare of Japan. A nationwide survey on the epidemiology and clinical features of eosinophilic granulomatosis with polyangiitis (Churg-Strauss) in Japan. *Mod Rheumatol* 2014;24:640-4.
- Hashimoto S, Fukutomi K, Shimizu H, Sasaki R, Ohno Y, Nagai M, Yanagawa H. A study on methodological issues in nationwide epidemiological surveys of intractable diseases. *Bull Inst Public Health* 1993;42:219-28. (in Japanese).
- Ohno Y, Suzuki S, Tamakoshi A, Sasaki R, Yanagawa H, Shibasaki H. Nationwide epidemiological survey for estimating annual number of patients treated for intractable diseases in Japan: its significance and methodology. In: *Recent Progress of Epidemiologic Study of Intractable Diseases in Japan*. In: Yanagawa H, Sasaki R, Nagai M, Ohno Y, Hirohata T, Hashimoto T, Inaba Y, Nakamura K., eds. *The Epidemiology of Intractable Diseases Research Committee* 1992:9-14.
- Hashimoto S, Fukutomi K, Nagai M, Nakamura Y, Yanagawa H, Sasaki R, Ohno Y. A note on methods for estimating the number of patients in the nationwide epidemiological survey on intractable diseases. *Nihon Koshu Eisei Zasshi* 1990;37:768-74. (in Japanese).
- Hashimoto S, Fukutomi K, Nagai M, Nakamura Y, Yanagawa H, Sasaki R, Ohno Y, Kubo N, Aoki K. A method of interval estimation for number of patients in the nationwide epidemiological survey on intractable diseases. *Nihon Koshu Eisei Zasshi* 1991;38:880-3. (in Japanese).
- Hashimoto S, Fukutomi K, Nagai M, Nakamura Y, Yanagawa H, Sasaki R, Ohno Y, Aoki K. Response bias in the nationwide epidemiological survey of an intractable disease in Japan. *J Epidemiol* 1991;1:27-30.
- Rönmark EP, Ekerljung L, Lötvall J, Torén K, Rönmark E, Lundbäck B. Large scale questionnaire survey on respiratory health in Sweden: effects of late- and

non-response. *Respir Med* 2009;103:1807-15.

- Nakai S, Hashimoto S, Murakami Y, Hayashi M, Manabe K, Noda H. Response rates and non-response bias in a health-related mailed survey. *Nihon Koshu Eisei Zasshi* 1997;44:184-91. (in Japanese).
- Breslow NE, Day NE. *Statistical Methods in Cancer Research Volume I - The analysis of case-control studies*. Lyon: International Agency for Research on Cancer (IARC Scientific Publications No. 32); 1980.

Appendix

Method of estimating the ratio of response rate between hospitals with and without patients with a disease

Let p_1 and p_2 be the underlying response rates, for hospitals with affected patients, in the initial and second surveys in a nationwide epidemiological survey of intractable diseases, respectively. Assume that the numbers of hospitals with affected patients that responded to the initial and second surveys are independent draws from binominal distributions with parameters p_1 and p_2 , respectively.

We assume that the ratio of the response rate of hospitals without affected patients to that of those with affected patients in the initial survey is equal to the equivalent ratio in the second survey (denote this ϕ). The response rates among the hospitals without affected patients in the initial and second surveys are then ϕ^*p_1 and ϕ^*p_2 , respectively. Assume further that the numbers of hospitals without affected patients that responded to the initial and second surveys are also independent draws from binominal distributions, but with parameters ϕ^*p_1 and ϕ^*p_2 , respectively. Under the model described above, ϕ can be estimated by the method of moments, using the number of hospitals with and without affected patients that responded to the initial and second surveys. Under the assumption of no response bias, $\phi=1$.

Method of estimating the total number of patients with a disease

Let m and α be the total numbers of all hospitals and all patients with the disease, respectively. Let N and X be the number of hospitals that responded to the survey and the number of affected patients reported in the survey, respectively. The estimate of α under the assumption of no response bias is then $\alpha^{\wedge}=m*X/N$. i.e., the number of patients with the disease is estimated by multiplying the total number of hospitals (m) by the mean number of patients with the disease reported by hospitals that responded to the questionnaire (X/N).

The estimate of α based on the estimate of $\phi(\phi^{\wedge})$ is given to be $\alpha_{\phi}^{\wedge}=m*X/(N_a+N_b/\phi^{\wedge})$, where N_a and N_b are the number of hospitals with and without affected patients that responded, respectively. For $\phi^{\wedge}=1$, $\alpha_{\phi}^{\wedge}=\alpha^{\wedge}$.

Consider that the number of patients in some strata, such as department type and the number of beds, are estimated using either of the above two methods. Let k be the number of strata, and $\alpha_1^{\wedge}, \alpha_2^{\wedge}, \dots, \alpha_k^{\wedge}$ be the estimated numbers of patients in the strata. The total numbers of patients in the strata are estimated to be $\alpha_1^{\wedge}+\alpha_2^{\wedge}+\dots+\alpha_k^{\wedge}$.

Copyright©2016 Shuji Hashimoto, PhD et al. 

This is an Open access article distributed under the Terms of Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.