Original

A sampling design for a Sentinel General Practitioner Network

Napoleón Pérez-Farinós *, Iñaki Galán, María Ordobás, Belén Zorrilla, José Luis Cantero y Rosa Ramírez

Servicio de Epidemiología, Instituto de Salud Pública de la Comunidad de Madrid, Madrid, España

ABSTRACT

Objective: To construct a design for probabilistic sampling of reporting physicians in sentinel networks.

Methods: We performed a multi-stage sample selection study. Data on primary care physicians and their patients were obtained from the Madrid Health Institute for 2005. The geographical unit of reference was the basic health area. A factorial analysis was performed on the basis of demographic, socio-cultural and socio-occupational variables. A cluster analysis was conducted to group the 247 basic health areas into homogeneous strata, which were then tested using a discriminant analysis. The general practitioners and pediatricians needed in each stratum were selected by simple random sampling. The representativeness of the population monitored by the selected physicians was studied with respect to the population of Madrid.

Results: Factorial analysis yielded five factors. Using these, 14 strata were obtained, which were shown to be homogeneous and mutually different by discriminant analysis. The minimum population that needed to be monitored consisted of 146,946 adults and 24,518 children, proportionally distributed among the respective strata. Eighty-eight general practitioners and 32 pediatricians were selected, who respectively covered populations of 154,610 and 31,336 persons representative of the general population.

Conclusions: Obtaining samples through suitable designs improves the accuracy of the information gathered by health sentinel networks in epidemiologic surveillance. Ensuring the representativeness of the study population vis-à-vis the general population is essential; cluster analysis and simple random sampling are methods that meet this need. Selecting physicians by means of probabilistic methods enables the accuracy of estimates to be ascertained.

© 2008 SESPAS. Published by Elsevier España, S.L. All rights reserved.

Diseño de selección de muestra para una red de medicos centinela

RESUMEN

Objetivo: Elaborar un diseño de muestreo probabilístico de los médicos notificadores en redes centinelas sanitarias.

Métodos: Estudio de muestreo multietápico. La información sobre los médicos de Atención Primaria y la población atendida procede del Instituto Madrileño de Salud para 2005. La unidad geográfica de referencia fue la Zona Básica de Salud. Partiendo de variables demográficas, socioculturales y sociolaborales, se realizó un análisis factorial. Se hizo un análisis de clusters para agrupar las 247 Zonas Básicas de Salud en estratos homogéneos. Como comprobación se realizó un análisis discriminante. Se seleccionaron los médicos necesarios en cada estrato mediante muestreo aleatorio simple. Se estudió la representatividad de la población vigilada con respecto a la población de Madrid.

Resultados: Se extrajeron 5 factores mediante análisis factorial. Se obtuvieron 14 estratos, que el análisis discriminante demostró homogéneos y distintos entre sí. Se determinó que era necesario vigilar una población mínima de 146.946 adultos y 24.518 niños, distribuidos proporcionalmente en cada estrato. Se seleccionaron 88 médicos de familia y 32 pediatras, que cubrían unas poblaciones de 154.610 y 31.336 personas, representativas de la población general.

Conclusiones: La obtención de muestras con un adecuado diseño asegura la validez de la información generada por redes centinelas sanitarias en la vigilancia epidemiológica. Debe buscarse la representatividad de la población estudiada con respecto a la población general. El análisis de *clusters* y el muestreo aleatorio simple son métodos que responden a esta necesidad. La selección de médicos mediante métodos probabilísticos permite conocer la precisión de las estimaciones.

© 2008 SESPAS. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Introduction

Sentinel networks are an increasingly useful tool in epidemiological surveillance and research. Since networks were first set up in Holland[1] and the United Kingdom[2] in the 1960s, their numbers have grown worldwide and they are used in fields as diverse as communicable[3–7] and non-communicable disease surveillance[8–11] and economic assessment studies[12].

In sentinel networks, a group of physicians reports cases of diseases or processes of interest, on a regular basis, generally weekly.

The Spanish National Health System offers universal health care coverage, and consequently the majority of the population tends to use public medical services. This provides a good opportunity to study population-based primary health care demand through a probabilistically obtained sample of physicians. The population attended by the sentinel physicians must be representative of the general population, if the information obtained by such a network is intended to be at all generalisable to the latter[13].

At present, sentinel networks are in place in most of Spain's Autonomous Regions and under different names share objectives and methodological aspects adapted to the Spanish National Health System[14,15]. Since first initiatives took with the sentinel networks in Castilla y León[16,17] and the Basque Country[18], the rest of Autonomous Regions have developed network for the surveillance of several different diseases, and they have tackled common projects[14].

The Madrid Autonomous Region Sentinel General Practitioner Network came into operation in 1991, and the reporters are volunteer physicians. Although the sampling in 1991 took into account geographic and demographic criteria, the final selection was done by requesting the collaboration from all physicians. Since its creation the network has suffered leavings and substitutions of reporting physicians; these factors, along demographic changes in population, make necessary a reviewing of the design and composition of the network.

The stated aim of this study was to develop a design for probabilistic sampling of reporting physicians in the Madrid Autonomous Region Sentinel General Practitioner Network (SGPN), to ensure the representativeness of the population covered by the network, in order to generalize the design for similar populations.

Methods

Data source: Information on general practitioners and paediatricians and their patients, broken down by age and sex, was obtained from the Madrid Health Institute's Health-Card Database for the 2005 period.

The designated geographic health care unit of reference was the Basic Health Zone (BHZ). There are 247 Basic Health Zones in the Madrid Autonomous Region, each of which represents the catchment area of a given Health Centre and has a median population size of 21,103 persons (range 2522–72,662).

Demographic, socio-cultural and socio-work variables were used to describe each BHZ (Table 1). Variables were obtained from the data base of the Madrid Statistics Institute for municipal register of 2005[19]. To reduce the number of variables without losing relevant information, a principal components analysis was performed[20–22], using the Varimax method. When extracting factors, account was taken of their eigenvalues and the values obtained in the matrix of rotated components.

Once the factors had been obtained, a cluster analysis was performed to group the 247 BHZs into homogeneous strata. Distance between units (BHZs) was measured in terms of the squared Euclidean distance, and the grouping method chosen was Ward's Method, because it is a hierarchical agglomerative method and it tends to make compact, similar sized clusters. The clustering history and dendrogram were then used to obtain the adequate number of clusters after different options were studied. The available statistical package used didn't offer other tools to determinate the number of clusters, but some other can be used, as pseudo-f, pseudo-t-square or cubic cluster criterion (CCC).

To test whether the strata were homogeneous and mutually different, we conducted a discriminant analysis, with the stratum or cluster of membership as the grouping variable, and the 5 factors yielded by the principal components analysis as the independent variables. Using variance analysis and Wilks' Lambda statistic, we evaluated whether the factors employed were different in the strata by means of the equality of means test. To assess the discriminant power of the discriminant functions obtained, we examined the eigenvalues and canonical correlation coefficients of such functions (which, when squared, express the percentage of variability explained by the differences between groups), and performed Wilks' Lambda test for each function. The latter can also be converted into a Chi-squared test, with its associated *p*-value of statistical significance, to verify the null hypothesis that there are no differences among strata by reference to the factors used. In addition, each BHZ's real stratum of membership was compared against its *a priori* probability of classification, through crossed validation, thereby enabling us to observe how many BHZs had been erroneously classified.

The number of persons needed to estimate the incidence rate of the most infrequent process of those monitored in Madrid in 2004 (varicella, herpes zoster, influenza and asthma) was ascertained by assuming that this disease followed a Poisson distribution and that the expected number of cases was greater than 10. Herpes zoster was the selected disease for this

Table 1
Variables used to describe Basic Health Zones. Contribution of each variable to the factors obtained in the principal components analysis. Matrix of rotated components

	Factor				
	1	2	3	4	5
Demographic variables					
Percentage of women	0.732	0.486	0.130	0.037	−0.042
Proportion of subjects aged under 15 years	−0.725	0.100	−0.122	0.120	0.613
Proportion of subjects aged over 64 years	0.981	−0.019	0.071	−0.061	−0.046
Proportion of subjects aged over 84 years	0.829	0.232	0.231	−0.015	−0.066
Independency index (quotient between population aged under 15 years and over 64 years, and population aged 15 to 64 years)	0.858	0.033	−0.011	−0.004	0.375
Replacement index (quotient between population aged 15 to 39 years, and population aged 40 to 64 years)	−0.325	−0.296	0.050	0.782	0.299
Ratio of progressiveness (quotient between population aged 0 to 4 years, and population aged 5 to 9 years (x 100))	0.027	0.199	0.055	0.892	−0.140
Percentage of foreigners	0.265	0.076	0.921	0.116	−0.072
Percentage of exclusively foreign homes	0.068	0.007	0.943	−0.004	−0.090
Percentage of mixed homes	−0.001	0.564	0.671	−0.061	0.305
Socio-cultural variables					
Percentage of separated or divorced subjects	0.529	0.367	0.478	0.117	0.162
Percentage of illiterate subjects	0.152	−0.796	−0.088	−0.129	0.387
Percentage of subjects with elementary or primary education	−0.047	−0.959	−0.118	0.013	−0.141
Percentage of subjects with secondary education (school leaving certificate)	−0.481	0.698	−0.175	0.240	−0.146
Percentage of subjects with university education (diplomas, degrees, doctorates)	0.120	0.930	0.140	−0.080	0.132
Socio-work variables					
Percentage of unemployed persons	−0.173	−0.721	0.074	0.286	0.133
Percentage of pensioners	0.970	−0.108	0.103	−0.046	−0.066
Percentage of gainfully employed subjects	−0.906	0.161	−0.047	0.235	0.156
Percentage of housewives	−0.074	−0.811	−0.281	−0.259	−0.155

calculation, as it was the most infrequent among the monitored processes. Assuming all these factors, the 95% confidence interval of the needed number of cases could it be ascertained by the next approximation:

$$A \pm Z_{1-\alpha/2} \bullet \sqrt{a}$$

The accuracy of this estimation is the relative accuracy (p) that we want to estimate the rates, and multiplied by the needed number of cases:

$$Z_{1-\alpha/2} \bullet \sqrt{a} = p \bullet a$$

So, the needed number of cases to estimate a rate with a determined accuracy will be:

$$a = \frac{Z_{1-\alpha/2}^2}{p^2}$$

We calculated the number of cases needed to estimate a rate with an accuracy of 10%, and an α of 0.05, taking into account a design effect of 1.5[23]. Once the necessary number of cases had been established, we then obtained the population required for monitoring purposes based on the herpes zoster incidence rate in Madrid in the preceding year. The resulting population was allocated to the strata in proportion to their respective populations.

In the Spanish National Health System, Primary Health Care is given by general practitioners or paediatricians, depending on the age of patients, so it is necessary that sentinel network includes both of them. As the SGPN is made up of Primary Care general practitioners and paediatricians, the population required for monitoring purposes was divided into adult and childhood populations, according to the proportion of population aged under 15 years drawn from the 2005 Madrid Regional population register.

In each stratum, general practitioners and paediatricians were consecutively selected by simple random sampling, until the pre-established required minimum populations were reached.

The representativeness of the population monitored by the selected physicians was studied, by age group and sex, with respect to the population of the Madrid Autonomous Region, in line with the 2005 population register. To this end, the proportion of the population in each sex and age group over the total sample selected was compared to the same proportion in the Madrid Autonomous Region.

The software used was the statistical package SPSS® 13.0.

Results

Factorial analysis

Five factors were extracted using the principal-components-analysis method. The first 4 had eigenvalues of over 1 and accounted for 81.5% of the variance. The fifth factor had an eigenvalue very close to 1. As these 5 factors would together account for 86.6% of the variance, we decided to include the fifth factor (Table 2).

The contribution of the variables to the 5 factors can be seen in the matrix of rotated components (Table 1).

Factor 1 (eigenvalue: 6.793) was related to demographic variables, registering high values in the most elderly populations, characterised by a high percentage of pensioners and a low percentage of gainfully employed subjects.

Factor 2 (eigenvalue: 5.033) was mainly related to educational variables, assuming high values in the indicators of intermediate and higher education, and low values in the percentages of:

Table 2
Eigenvalues of factors extracted and percentage of variance explained

Factor	Eigenvalues	Percentage of variance explained	Cumulative percentage
1	6.793	35.8	35.8
2	5.033	26.5	62.2
3	2.212	11.6	73.9
4	1.441	7.6	81.5
5	0.977	5.1	86.6

Table 3
Discriminant analysis to test strata homogeneity: Wilks' Lambda coefficient, ANOVA and p-value of statistical significance

Factor	Wilks' Lambda	F[a]	p value
1	0.235	58.5	<0.001
2	0.205	69.4	<0.001
3	0.354	32.7	<0.001
4	0.504	17.7	<0.001
5	0.284	45.2	<0.001

[a] Snedecor's F tests for ANOVA.

persons with primary education and without any formal education; unemployed persons; and housewives.

Factor 3 (eigenvalue: 2.212) was related to variables of immigration, assuming high values in the populations with most immigrants.

The main contribution to Factor 4 (eigenvalue: 1.441) was made by the replacement and progressiveness indices, with high values in the youngest populations.

Factor 5 (eigenvalue: 0.977) was related to the proportion of subjects under the age of 15 years.

Cluster analysis

On completion of the cluster analysis, 14 strata were obtained. The number of BHZs in these strata ranged from 4 to 34, with a median of 18.

Discriminant analysis

Following the discriminant analysis, the Wilks' Lambda coefficient displayed low values in the equality of means test, reflecting differences among the strata for each factor. Moreover, the F statistics for each of the resulting factors in the analysis of variance (ANOVA) registered associated p-values that were statistically significant (Table 3).

A total of 5 discriminant functions were obtained: their eigenvalues ranged from 0.90 to 4.06, and their canonical correlation coefficients registered high values, with the lowest being 0.69 and the highest 0.90. As a result, Wilks' Lambda test yielded values very close to 0 for all of the first 4 discriminant functions (Table 4). Reference to the classification matrix showed that, according to the discriminant functions, 86.6% of the health zones were correctly classified.

Sample selection

The number of cases required to estimate an incidence rate with a relative accuracy of 10% was 384. As the herpes zoster incidence rate in Madrid in 2004 was 0.00335932 year^{-1}, a population of 114,309 was thus needed. Applying a design effect

Table 4
Discriminant analysis to test strata homogeneity. Canonical discriminant functions, Wilks' Lambda coefficient, Chi-squared test and p-value of statistical significance

Function	Eigenvalue	Canonical correlation coefficient	Wilks' Lambda	Chi-squared	p
1	4.06	0.90	0.002	1443.4	<0.001
2	3.82	0.89	0.011	1059.8	<0.001
3	2.56	0.85	0.055	687.9	<0.001
4	1.70	0.79	0.194	387.3	<0.001
5	0.90	0.69	0.526	152.1	<0.001

Table 5
Population required for surveillance purposes in each stratum obtained. Total, adult and childhood populations

| Stratum	Number of physicians	Reference population	Percentage of population	Total sample size	Adult sample size	Childhood sample size
1	275	402 459	6.75	11 568	9914	1654
2	248	403 352	6.76	11 594	9936	1658
3	303	458 948	7.69	13 192	11 306	1886
4	273	486 402	8.15	13 981	11 982	1999
5	525	782 290	13.11	22 486	19 271	3215
6	477	801 647	13.44	23 042	19 747	3295
7	252	392 893	6.59	11 293	9678	1615
8	98	162 858	2.73	4681	4012	669
9	407	618 929	10.38	17 790	15 246	2544
10	304	469 124	7.86	13 484	11 556	1928
11	51	67 286	1.13	1934	1657	277
12	479	728 355	12.21	20 935	17 941	2994
13	47	68 664	1.15	1974	1692	282
14	74	122 112	2.05	3510	3008	502
Total	3813	5 965 319		171 464	146 946	24 518
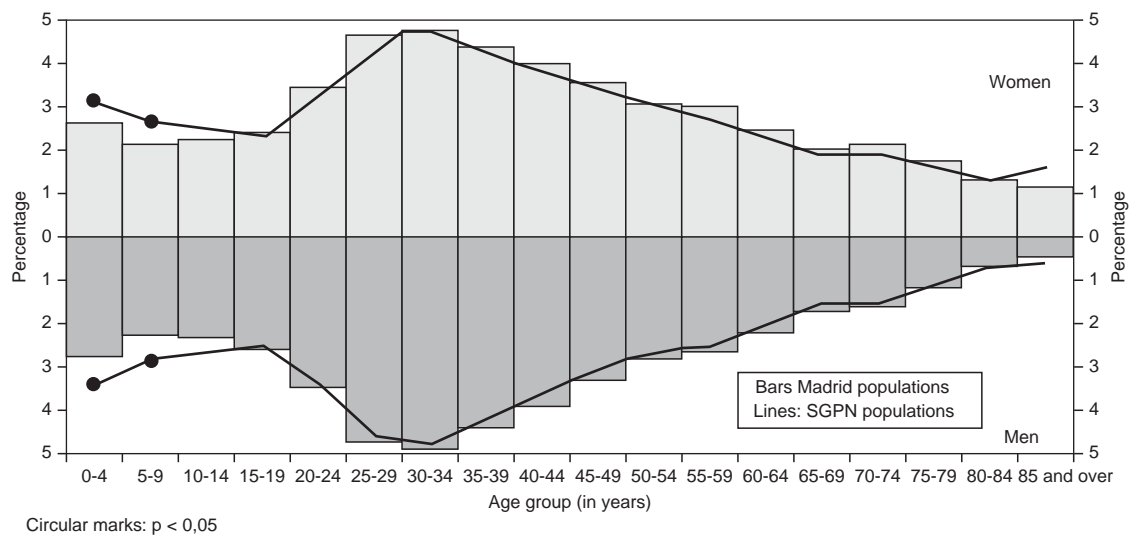
Fig. 1. Comparison between proportions of population monitored by the selected physicians and the population of Madrid. Percentages of populations of each age group and both sexes are represented; Madrid populations are represented by bars and SGPN populations by lines. When a circular marks appears it means that the difference between both proportions is statistically significant.

of 1.5, the necessary population was 171,464 persons, representing 3% of the Madrid Autonomous Region's population. In view of the fact that 14.3% of the Madrid population in 2003 was under the age of 15 years, the required minimum population was 146,946 adults and 24,518 children. These populations were distributed among the 14 strata in proportion to the total populations of each stratum (Table 5).

In each stratum, general practitioners and paediatricians were consecutively selected by simple random sampling, until the pre-established required minimum populations were reached. We consecutively selected 88 general practitioners and 32 paediatricians, until the pre-established required minimum populations were reached; finally general practitioners and paediatricians selected covered populations of 154,610 and 31,336 subjects, respectively.

In the representativeness study, no statistically significant differences ($p > 0.05$) were observed between the sample and the Madrid Autonomous Region in terms of the proportions of the population in each age group and sex, except for the groups aged 0 to 4 and 5 to 9 years in both sexes (Fig. 1).

Discussion

One of the main difficulties of sentinel networks lies in selecting the reporting physicians in such a way that the monitored population is representative of the general population. Most sentinel networks use design or quality control tools to enhance such representativeness as far as possible[3,15].

Indeed, these types of designs, which at times encounter difficulties posed by the great extent of territory to be covered or the dispersion of the population to be monitored, are not always a simple task[24,25]. The reporting physicians in many highly important, traditional networks are volunteers[26–28].

The Madrid Autonomous Region has a considerable population that is not homogeneous, with a great degree of demographic, cultural and economic diversity. Simple random sampling is an effective technique but does not ensure representativeness. However, by dividing the population into strata by means of variables which inform on characteristics that differ between some population groups and others, and selecting subsamples in each stratum, the representation of all in the total sample is guaranteed[15]. Furthermore, if the strata can be made uniform, the accuracy of the estimates obtained will be greater[17,23]. The choice of a probabilistic method is always desiderable, but sometimes the method of selection is limited by the available resources, the health system structure, and the Primary Health Care Services features. In the method presented in this article, physicians are the selected units, since it is not feasible to select a random sample of people attended in Primary Health Care Services; the probabilistic method is not applied to persons but physicians; so every citizen have their own likelihoods to be selected depending on their assigned physicians.

In the Madrid Autonomous Region, Basic Health Zones are units of sufficiently small size to ensure that the demographic, socio-economic and socio-cultural variables within them are homogeneous. Hence, a technique such as cluster analysis can efficiently yield strata composed of various BHZs having similar characteristics. Random selection of reporting physicians in each stratum renders this design capable of producing a sample that is representative and estimators that are unbiased and accurate.

A number of physicians was selected in each stratum in such a way as to guarantee that the monitored population exceeded the pre-established minimum required.

The study's representativeness highlighted the fact that the population covered by the selected sample of physicians was representative of the population of the Madrid Autonomous Region in terms of age and sex, though in the 0–4 and 5–9 age groups, the proportion of the monitored population was significantly higher than the proportion of the population of the Madrid Autonomous Region. This was due to the fact that in the smallest strata, the minimum paediatric population required was likewise very small, and that the population monitored in these strata by the first paediatrician selected was far higher. Nevertheless, the magnitudes of the differences observed were small, and there is no reason why this should in any way compromise the validity of the results obtained by the network. The good representativeness of the selected sample according to the general population of Madrid is a positive point in the sample selection process.

The data of population assigned to each physician come from Madrid Health Institute's Health-Card Database, which includes all population; however there are a percentage of people that use private health services. It is important to take into account that this fact could to influence the sampling design by underestimating the incidence rates of the monitored diseases. This bias can be partially reduced if we monitor a larger population than theoretically defined.

One of the limitations of this design might be the fact that geographic criteria were not taken into account. The Madrid Autonomous Region is a small territory, with little geographic spread, so that any possible social, economic or cultural differences are geographically close. That is the reason of include social and demographic variables, but in case of a selection of a sample in territories with different features is necessary to consider other data as population density or geographic dispersion. Accordingly, we chose to take cluster analysis into account as a technique for obtaining homogeneous units, even though the ensuing strata may not be territories that are geographically united. The choice of a mainly adult disease as herpes zoster to pre-determinate the sample size does not affect the validity neither of the design nor the obtained results, as childhood diseases currently monitored are more frequent. Although a specific process as herpes zoster has supported it, we must emphasize that the aim of this work is to set the frame of a design applicable to future studies.

Although a design effect of 1.5 was applied, when a multistage method is designed, a higher design effect should be used if the design is very complex.

Another limitation of the study could be the fact of a high percentage of selected physicians refused to collaborate, as a voluntariness bias could be introduced and the representativeness of sample would be affected. Nevertheless, if a physician refused to participate, another one would be selected by simple random sampling in the same stratus. It is not foreseeable a high number of rejections, as Primary Health Care physicians keep a strong commitment with their labour, and the burden derived from the work in the sentinel network is not expected to be excessive.

In view of the fact that sentinel networks are being increasingly used in epidemiological surveillance, suitable designs are required to ensure attainment of their designated goals.

In order to be able to generalise the results obtained, a feature to be sought is the representativeness of the network vis-à-vis the general population. Cluster analysis is a method that answers this need. To this end, it is necessary to have sampling units that are as unaggregated as possible, together with information on demographic, social, economic or cultural factors of interest.

Furthermore, selection of physicians by probabilistic methods, as in this study, affords the chance of ascertaining the degree of accuracy of the estimates obtained, thereby adding value to the data collected by sentinel networks.

Acknowledgements

References

1. Bartelds AIM. Continuous morbidity registration sentinel at Dutch sentinel stations, 2003. Utrecht: NIVEL; 2004.
2. Thomas P, Griffiths F, Kai J, et al. Networks for research in primary health care. BMJ. 2001;322:588–90.
3. Deckers JG, Paget WJ, Schellevis FG, et al. European primary care surveillance networks: their structure and operation. Fam Pract. 2006;23:151–8.
4. Ciofi Degli Atti ML, Salmaso S, Bella A, et al. Pediatric sentinel surveillance of vaccine-preventable diseases in Italy. Pediatr Infect Dis J. 2002;21:763–8.
5. Dedman DJ, Zambon M, Buynder PV, et al. Influenza surveillance in England and Wales: October 1997 to June 1998. Commun Dis Public Health. 1998;1:244–51.
6. Massari V, Maison P, Desenclos JC, et al. Six years of sentinel surveillance of hepatitis B in general practice in France. Eur J Epidemiol. 1998;14:765–7.
7. Ordobas MA, Zorrilla B, Arias P. Influenza in Madrid, Spain, 1991–92: validity of the sentinel network. J Epidemiol Community Health. 1995;49(Suppl 1): 14–6.
8. Arteagoitia JM, Larranaga MI, Rodriguez JL, et al. Incidence, prevalence and coronary heart disease risk level in known type 2 diabetes: a sentinel

practice network study in the Basque Country, Spain. Diabetologia. 2003;46: 899–909.

9. Fleming DM, Schellevis FG, Van Casteren V. The prevalence of known diabetes in eight European countries. Eur J Public Health. 2004;14:10–4.

10. Devroey D, Van CV, De Lepeleire J. Revealing regional differences in the institutionalization of adult patients in homes for the elderly and nursing homes: results of the Belgian network of sentinel GPs. Fam Pract. 2001;18: 39–41.

11. Le Pont F, Letrilliart L, Massari V, et al. Suicide and attempted suicide in France: results of a general practice sentinel network, 1999–2001. Br J Gen Pract. 2004;54:282–4.

12. Fornaro P, Gandini F, Marin M, et al. Epidemiology and cost analysis of varicella in Italy: results of a sentinel study in the pediatric practice. Italian Sentinel Group on Pediatric Infectious Diseases. Pediatr Infect Dis J. 1999;18:414–9.

13. Schlaud M, Brenner MH, Hoopmann M, et al. Approaches to the denominator in practice-based epidemiology: a critical overview. J Epidemiol Community Health. 1998;52(Suppl 1):13S–9S.

14. Proyecto RECENT. Redes Centinelas Sanitarias. 2008. Available from: http://www.redescentinelas.com/.

15. Vega Alonso AT, Zurriaga LO, Galmes TA, et al. Guía de principios y métodos de las redes centinelas sanitarias en España. Gac Sanit. 2006;20(Suppl 3):52–60.

16. Consejería de Sanidad de Castilla y León. Red Centinela Sanitaria de Castilla y León. 2008. Available from: http://www.sanidad.jcyl.es/sanidad/cm/profesionales/tkContent?idContent=420&locale=es_ES&textOnly=false.

17. Vega Alonso AT, Gil CM, Ruiz CC, et al. La Red de Médicos Centinelas de Castilla y León: aplicación del análisis de conglomerados para la obtención de una población representativa. Gac Sanit. 1990;4:184–8.

18. Médicos vigía: un instrumento para la vigilancia e investigación epidemiológica. Red de médicos vigía del País Vasco 1990–2002. Vitoria-Gasteiz: Departamento de Sanidad, Gobierno Vasco; 2002.

19. Instituto de Estadística de la Comunidad de Madrid. Padrón continuo 2005. Zonificación de Salud. 2008. Available from: http://www.madrid.org/iestadis/fijas/estructu/demograficas/padron/pc05i5.htm.

20. Kleinbaum DG, Kupper LL, Muller KE. Applied regression analysis and other multivariable methods. 2nd ed. Belmont, CA: Duxbury Press; 1988.

21. Raykov T, Marcoulides GA. An Introduction to Applied Multivariate Analysis. New York: Taylor & Francis; 2008.

22. Anderson TW. An introduction to Multivariate Statistical Analysis. 2nd ed. New York: Wiley; 1984.

23. Silva Ayçaguer LC. Diseño razonado de muestras y captación de datos para la investigación sanitaria. Madrid: Díaz de Santos; 2000.

24. Wilson I. Australian Sentinel Practice Research Network. Commun Dis Intell. 2003;27:123–4.

25. Tilyard M, Dovey S. The role of sentinel networks in general practice research in New Zealand and internationally. NZFP. 1999;26.

26. Van Casteren V. Belgian Sentinel general practitioners. Belgium Scientific Institute of Public health. 2006. Available from: http://www.iph.fgov.be/epidemio/epien/PROG10.HTM.

27. Green LA, Miller RS, Reed FM, et al. How representative of typical practice are practice-based research networks? A report from the Ambulatory Sentinel Practice Network Inc (ASPN). Arch Fam Med. 1993;2:939–49.

28. Lobet MP, Stroobant A, Mertens R, et al. Tool for validation of the network of sentinel general practitioners in the Belgian health care system. Int J Epidemiol. 1987;16:612–8.