Validity of self reported utilisation of primary health care services in an urban population in Spain

Juan Ángel Bellón, Pablo Lardelli, Juan de Dios Luna, Ana Delgado

Abstract

Study objective—To assess the validity and factors related with the validity of self reported numbers of visits to a primary health care centre, in comparison with the recorded number.

Design—Cross sectional study.

Setting—The urban area served by the Zaidín-Sur Primary Health Care Centre (Granada, Spain).

Participants—Two population samples (236 high users and 420 normal users) who were seen at the centre from 1985 to 1991 were interviewed in 1993.

Main results—A net tendency to over-report the actual number of visits was observed. Absolute concordance between self reported and recorded utilisation decreased as time interval lengthened, although this mainly reflected the increase in maximum variability both with time interval length and with the number of recorded visits. Corrected Spearman ρ coefficients obtained between the number of self reported and recorded visits ranged from 0.602 for the two weeks before the interview to 0.678 for the year before. Regression slopes of self reported utilisation upon recorded utilisation did not change between periods. In multiple regression analyses the actual number of visits was the main factor associated with both under-reporting and overreporting. Older age was also significantly associated with underreporting. Poor health status and high satisfaction with health care were significantly associated with overreporting.

Conclusions—There was a substantial degree of inaccuracy in self reported utilisation, with a net tendency to overreport the number of visits. In relative terms, however, accuracy of self reports did not seem to decrease appreciably as the recall time lengthened. To compare the accuracy of different measures, it is important to take into account the maximum variability of each one. Otherwise, contradictory results may be obtained.

(J Epidemiol Community Health 2000;54:544-551)

The utilisation of health services has been studied from different angles during recent decades,¹⁻¹⁰ and usage patterns have been widely used to plan health care services.¹¹⁻¹⁵ Utilisation of primary health care services can be measured from records, health diaries, and questionnaires,¹⁶⁻²¹ but questionnaires are the

most frequently used source of data.²²⁻³⁰ However, few studies have focused on the accuracy of self reported utilisation, and the factors that influence it. Both issues are central to the soundness of methods used to analyse the utilisation of primary health care services.³⁰⁻³¹ Therefore a number of questions need to be answered to establish the validity of self reported utilisation data. How accurate is self reported utilisation? Does accuracy of the information decrease as the recall interval increases? If the answer to this latter question is yes, when does the information cease to be valid? And finally, what variables influence the validity of self reported utilisation data?

Methods

This cross sectional study was originally designed to identify factors related to high usage of medical services in primary health care. A detailed description of the method has been published in a previous article.³²

SETTING

The research was done at the Centro de Salud Zaidín-Sur in Granada (a city of 300 000 inhabitants in south eastern Spain), which serves a population of approximately 24 000 inhabitants living in a defined area within the city limits. As part of the national health service, the centre provides free medical care to 100% of the population in the catchment area. More information about the Spanish national health system is available in Hart's 1990 article.³³

POPULATION AND SAMPLES

The eligible population consisted of all subjects aged 14 years or older who were seen at the health centre between 1 August 1985 and 1 March 1991. Five of the 10 family physicians of the centre participated. From their patient lists we selected a random sample of 2018 people, for whom we recorded use during the period from 1 March 1991 to 29 February 1992 from the centre's patient records. High use (HU) was defined as a number of visits to the health centre greater than the mean plus one standard deviation for the age and sex group. The HU group contained 285 subjects (14% of the sample). The remaining 1733 subjects were considered normal users (NU).

All high users and a random sample of 511 NU were asked to respond to a specially designed questionnaire. Forty nine patients in the HU group and 91 in the NU group could not be interviewed. Therefore the final sample

Unidad Docente de Medicina Familiar y Comunitaria de Granada, Centro de Salud Zaidín-Sur, Granada, Spain
J Á Bellón

Departamento de Medicina Preventiva y Salud Pública, Facultad de Farmacia, Universidad de Granada, 18071 Granada, Spain
P Lardelli

Departamento de Bioestadística, Facultad de Medicina, Universidad de Granada, 18071 Granada, Spain
J Luna

Unidad Docente de Medicina Familiar y Comunitaria, Hospital Clínico Universitario de Granada, Granada, Spain
A Delgado

Correspondence to:
Dr Lardelli
(lardelli@goliat.ugr.es)

Accepted for publication
15 February 2000

Table 1 Distribution of the sample according to the main variables for respondents and non-respondents

Categories of variables	Proportion of respondents (n=589)	Proportion of non-respondents (n=67)	χ^2 test p value
Sex			0.315
Male	28.2	22.4	
Female	71.8	77.6	
Age group (y)			0.002
14–30	19.0	9.0	
31–45	25.8	22.4	
46–60	21.4	19.4	
61–75	24.8	25.4	
>75	9.0	23.9	
Marital status			0.009
Unmarried/married	82.0	68.7	
Widowed/divorced/separated	18.0	31.3	
Level of education			0.001
Primary-secondary	59.9	38.8	
No schooling	40.1	61.2	
Employment status			0.002
Working	21.7	7.5	
Unemployed	7.6	3.0	
Retired	27.2	34.3	
Works at home	37.9	55.2	
Other	5.6	0.0	
Number of persons in household			0.057
1–2	30.4	41.8	
>2	69.6	58.2	
Travelling time to health centre			0.212
<10 min	82.3	76.1	
>10 min	17.7	23.9	
Satisfaction with physician			0.425
Never-sometimes	23.9	28.4	
Almost always/always	76.1	71.6	
Seen by private physician			0.853
No	82.7	83.6	
Yes	17.3	16.4	
Seen by faith healer			0.665
No	95.9	97.0	
Yes	4.1	3.0	
Seen by alternative medicine practitioner			0.820
No	96.1	95.5	
Yes	3.9	4.5	
Level of use of health care services			0.034
Normal user	65.4	52.2	
High user	34.6	47.8	
Number of chronic illnesses			0.005
0–10	87.3	74.6	
>10	12.7	25.4	
Level of self perceived health			0.022
Excellent	7.3	1.5	
Acceptable	35.8	25.4	
Neither good nor bad	17.0	13.4	
Fair	30.2	47.8	
Poor	9.7	11.9	
Family APGAR			0.927
Functional	84.0	83.6	
Dysfunctional	16.0	16.4	
Score on Duke scale (dichotomised)			0.817
Normal social support	88.6	89.6	
Low social support	11.4	10.4	
GHQ score (dichotomised)			0.001
Probably healthy	52.0	31.3	
Probably ill	48.0	68.7	
Mean number of recorded visits	5.43	7.59	

Table 2 Descriptive statistics of self reported and recorded utilisation in the four periods

	Mean	SD	Minimum	Median	Maximum
Two weeks					
Number of recorded visits	0.24	0.52	0	0	3
Number of self reported visits	0.39	0.66	0	0	3
Three months					
Number of recorded visits	1.34	1.77	0	1	13
Number of self reported visits	1.53	1.96	0	1	15
Six months					
Number of recorded visits	2.57	3.09	0	2	24
Number of self reported visits	2.94	3.49	0	2	24
One year					
Number of recorded visits	5.43	5.89	0	4	45
Number of self reported visits	5.53	6.41	0	4	52

consisted of 236 HU and 420 NU. There were no significant differences in age or sex distribution between patients who were unavailable for interview and those who were interviewed.

MEASUREMENTS

All variables except recorded utilisation were obtained in the course of individual interviews with the patient in their home. The interview conditions were standardised, and all interviews were held between 1 January 1993 and 30 June 1993. The following information was obtained:

- Number of scheduled visits during the 2 weeks, 3 months, 6 months and 12 months before the interview.
- General variables: age, sex, marital status, educational level and employment status according to the classification of the Spanish Instituto Nacional de Estadística³⁴; socioeconomic level according to an adaptation of the Clasificación Nacional de Ocupaciones to social class³⁵; number of persons per household.
- Health status: self reported health and number and type of reported chronic illnesses according to the pertinent items in the Encuesta Nacional de Salud³⁶; mental health, measured using Goldberg's GHQ-28 questionnaire.^{37–38}
- Family function, measured with the family APGAR index.^{39–41}
- Social support, measured with the Duke-UNC-11 scale.^{42–43}

Information about scheduled recorded utilisation during the same periods for which the patients were questioned in the interview was obtained from the list of patients who had a scheduled appointment at the health centre.

DATA ANALYSIS

Descriptive statistics of self reported utilisation (SU), recorded utilisation (RU), and the difference between them (DSR = SU–RU) were estimated for each time interval. The logarithm of the SU/RU ratio was also calculated (replacing 0 values with 0.001) to obtain estimates of relative discordances weighted by the actual number of visits. The Wilcoxon rank test was used to find significant differences between SU and RU distributions. The correlation between the two measures was studied by estimating Spearman's ρ coefficient. Because maximum variability differed for each interval we calculated corrected coefficients (ρ_c) for maximum possible variability in each interval with the equation: $\rho_c = \rho/\rho_{\max}$, where $\rho_{\max}$ was obtained according to the expression proposed by Shih and Huang.⁴⁴ Simple linear regression of SU upon RU was done for each time interval, and DSR was plotted against RU for each period.

To identify factors related with validity, we constructed two multiple linear regression models for each time interval: one for over-reporting and one for under-reporting. In the first case, the dependent variable was DSR, ranging from 0 to its maximum value. In the second one we used absolute values of DSR, ranging from the minimum value to 0.

Finally, DSR values were grouped into three categories: underreporters (values lower than the mean minus one standard deviation (1SD)), accurate reporters (values within the mean ($\pm 1SD$)) and overreporters (values

Table 3 Comparison of the difference between self reported and recorded utilisation (DSR) and the logarithm of the ratio between self reported and recorded utilisation ($\ln$ SU/RU) in the four recall periods

	Mean	SD	Min	Median	Max	% (U) ^a	% (A) ^b	% (O) ^c
DSR								
2 weeks	0.144	0.601	-2	0	3	7.1	73.7	19.2
3 months	0.193	1.519	-6	0	8	23.3	44.1	32.6
6 months	0.370	2.583	-14	0	17	29.7	29.9	40.4
12 months	0.102	4.864	-28	0	28	37.9	17.5	44.6
	Mean	SD	Min	Median	Max	% (U) ^d	% (A) ^e	% (O) ^f
$\log$(SU/RU)								
2 weeks	0.678	3.140	-7.6	0	8.0	5.9	78.7	15.4
3 months	0.648	3.491	-8.0	0	8.3	10.0	70.1	19.9
6 months	0.722	3.519	-8.8	0	8.7	11.5	68.0	20.5
12 months	0.377	2.868	-9.3	0	9.4	11.2	73.0	15.8

^a: % (U): Proportion of underreporters: DSR < 0. ^b: % (A): Proportion of accurate reporters: DSR = 0. ^c: % (O): Proportion of overreporters: DSR > 0. ^d: % (U): Proportion of underreporters: $\ln$ (SU/RU) < -1. ^e: % (A): Proportion of accurate reporters: -1 < $\ln$ (SU/RU) < +1. ^f: % (O): Proportion of overreporters: $\ln$ (SU/RU) > +1.

greater than the mean +1SD). Nominal polychotomous logistic regression analysis was then done for each subperiod, using the category of accurate reporters as the reference. Two adjusted odds ratios (OR) were obtained for each independent variable to measure the strength of its association with overreporting and underreporting separately. The SPSS (version 6.0)⁴⁵ and the BMDP (version 7.0)⁴⁶ statistical packages were used for all analyses.

Results

CHARACTERISTICS OF THE SAMPLE AND NON-RESPONDENTS

Information about the number of recorded visits was not available for two subjects. Sixty five subjects (9.6%) did not answer the question about the previous number of visits for at least one of the recall periods considered. The proportion of non-responders increased as the recall period became longer: eight subjects (1.2%) for the past two weeks; 14 (2.1%) for the past three months; 38 (5.8%) for the past six months, and 63 (9.6%) for the past year. We present the results only for the 589 subjects for whom information about recorded and self reported visits was available for all periods.

Differences between responders and non-responders according to the main variables studied are presented in table 1. Non-responders were generally older than responders, and included a larger proportion of persons with a lower level of education, divorced or separated persons, widows and widowers, women who did not work outside the home,

and retired persons. Their perceived health was worse, the number of reported illnesses higher, and their GHQ score significantly higher than in responders. Non-responders were more frequently high users; accordingly, the mean number of recorded visits in the previous year was significantly higher than for responders.

VALIDITY ASSESSMENT

Table 2 shows descriptive statistics for SU and RU. Means, medians, standard deviations and maximum values of both SU and RU increased as the recall period lengthened. Mean values and standard deviations of SU were always higher than the corresponding RU values. Table 3 shows descriptive statistics for DSR. All mean values were positive numbers. Standard deviations and ranges both increased as the recall interval lengthened. Accordingly, the proportions of overreporters and underreporters also increased with length of the recall period, but the former were always higher. Table 3 also shows descriptive statistics for the logarithm of SU/RU. The proportion of accurate reporters (considered as those values between -1 and +1) did not differ between periods. Table 4 shows the measures of association between SU and RU. Significant differences between SU and RU distributions were observed for all time intervals except 12 months. The differences in the values of uncorrected Spearman correlation coefficients for each period (from 0.489 in the two week interval to 0.672 in the one year interval) were much lower when corrected coefficients were used (from 0.602 to 0.678 in the two week and one year periods respectively). The value of the intercept of the regression of SU upon RU increased as recall period lengthened, but the slopes did not change significantly across periods. Figures 1 to 4 show the plots of DSR against RU for each recall period. The plots for the three month, six month and one year periods were quite similar, and they showed that the magnitudes of both positive and negative DSR values increased with the recorded number of visits.

VARIABLES ASSOCIATED WITH ACCURACY

Table 5a shows the results of multiple linear regression analyses for the subsample of accurate or underreporting subjects (DSR ≤ 0). The number of recorded visits was directly associated with underreporting in all recall

Table 4 Measures of association between self reported (SU) and recorded utilisation (RU)

Wilcoxon rank test				
	2 weeks	3 months	6 months	1 year
p value	<0.001	0.028	0.004	0.510
Spearman correlation coefficient				
	2 weeks	3 months	6 months	1 year
Uncorrected	0.489	0.609	0.602	0.672
Corrected	0.602	0.646	0.620	0.678
Coefficients and 95% CI for the simple linear regression of SU upon RU:				
	2 weeks	3 months	6 months	1 year
Intercept	0.229 (0.178, 0.281)	0.531 (0.384, 0.639)	0.911 (0.647, 1.175)	1.455 (0.943, 1.966)
Slope	0.648 (0.559, 0.778)	0.748 (0.681, 0.814)	0.789 (0.724, 0.855)	0.751 (0.687, 0.815)

Figure 1 Plot of DSR against recorded use in two week recall period. DSR: difference between self reported and recorded use.

periods. The magnitude of the coefficients increased as recall period lengthened. Older age was associated with underreporting in the one year period. The number of chronic illnesses and no schooling showed a significant inverse association with underreporting in the three month and six month periods. In the model for overreporting ($DSR \geq 0$) (Table 5b), the number of recorded visits was directly and significantly associated with overreporting for all but the first period. The number of chronic illnesses was also associated with overreporting in two week and six month periods. Finally, a strong direct association between visits to faith healers and overreporting was seen in the

Figure 2 Plot of DSR against recorded use in three month recall period. DSR: difference between self reported and recorded use.

Figure 3 Plot of DSR against recorded use in six month recall period. DSR: difference between self reported and recorded use.

Figure 4 Plot of DSR against recorded use in one year recall period. DSR: difference between self reported and recorded use.

three and six month periods. In the same time intervals, the opposite effect appeared for visits to alternative medicine practitioners.

In the polychotomous regression model (table 6), the variable most strongly associated with underreporting was the number of

Table 5 Multiple linear regression analysis

5a Variables associated with underreporting ($DSR \leq 0$)

Variable	2 weeks Coeff (95% CI)	3 months Coeff (95% CI)	6 months Coeff (95% CI)	1 year Coeff (95% CI)
Male sex	-0.011 (-0.066, 0.044)	0.077 (-0.074, 0.228)	-0.152 (-0.422, 0.118)	-0.095 (-0.680, 0.490)
Age	-0.000 (-0.002, 0.002)	0.006 (-0.001, 0.011)	0.008 (-0.001, 0.018)	0.025 (0.005, 0.045)
Widowed/divorced/separated	0.001 (-0.071, 0.072)	0.019 (-0.179, 0.216)	0.062 (-0.294, 0.419)	0.241 (-0.467, 0.949)
No schooling	-0.036 (-0.095, 0.023)	-0.001 (-0.158, 0.157)	-0.345 (-0.628, -0.062)	-0.454 (-1.043, 0.136)
Retired/works at home	0.029 (-0.032, 0.090)	-0.162 (-0.330, 0.007)	0.041 (-0.260, 0.341)	-0.366 (-1.035, 0.302)
Number of persons in household	0.012 (-0.042, 0.067)	0.053 (-0.100, 0.207)	0.054 (-0.223, 0.332)	0.059 (-0.539, 0.658)
Travel time >10 min to health centre	0.020 (-0.043, 0.084)	0.166 (-0.003, 0.336)	0.152 (-0.152, 0.457)	-0.078 (-0.711, 0.555)
Number of recorded visits	<i>0.400 (0.356, 0.444)</i>	<i>0.438 (0.339, 0.476)</i>	<i>0.467 (0.422, 0.511)</i>	<i>0.573 (0.524, 0.621)</i>
High user	0.012 (-0.045, 0.069)	-0.020 (-0.184, 0.143)	-0.278 (-0.577, 0.021)	-0.574 (-1.178, 0.031)
Almost always/always satisfied	-0.018 (-0.072, 0.036)	-0.022 (-0.171, 0.127)	0.070 (-0.192, 0.333)	-0.043 (-0.612, 0.527)
Fair/poor self perceived health	0.013 (-0.012, 0.038)	0.056 (-0.053, 0.084)	0.011 (-0.110, 0.133)	-0.099 (-0.359, 0.160)
Number of chronic illnesses	-0.037 (-0.077, 0.003)	-0.133 (-0.242, -0.023)	-0.085 (-0.281, 0.112)	-0.302 (-0.724, 0.121)
Seen by private physician	-0.019 (-0.082, 0.045)	0.160 (-0.020, 0.341)	0.216 (-0.100, 0.532)	0.333 (-0.345, 1.011)
Seen by faith healer	0.060 (-0.084, 0.203)	0.100 (-0.285, 0.485)	0.205 (-0.510, 0.919)	-0.809 (-2.118, 0.500)
Seen by alternative medicine practitioner	-0.092 (-0.047, 0.232)	0.135 (-0.206, 0.477)	0.088 (-0.484, 0.660)	0.577 (-0.683, 1.837)
Family APGAR	-0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)	-0.000 (-0.000, 0.000)	-0.000 (-0.000, 0.000)
Score on DUKE scale	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)
GHQ score	0.008 (-0.015, 0.031)	-0.035 (-0.100, 0.029)	-0.089 (-0.202, 0.024)	-0.068 (-0.308, 0.172)

Statistically significant coefficients and confidence intervals are in *italics*.

5b Variables associated with overreporting ($DSR \geq 0$)

Variable	2 weeks Coeff (95% CI)	3 months Coeff (95% CI)	6 months Coeff (95% CI)	1 year Coeff (95% CI)
Age	-0.001 (-0.005, 0.003)	0.002 (-0.007, 0.011)	-0.000 (-0.017, 0.016)	-0.001 (-0.037, 0.024)
Widowed/divorced/separated	-0.032 (-0.160, 0.097)	-0.105 (-0.416, 0.206)	0.263 (-0.328, 0.854)	0.597 (-0.554, 1.748)
No schooling	-0.007 (-0.115, 0.101)	-0.039 (-0.299, 0.220)	-0.447 (-0.928, 0.033)	-0.500 (-1.405, 0.405)
Retired/works at home	0.101 (-0.014, 0.217)	0.068 (-0.210, 0.347)	0.413 (-0.091, 0.917)	0.543 (-0.388, 1.475)
Number of persons in household	-0.011 (-0.115, 0.093)	-0.067 (-0.323, 0.189)	0.171 (-0.308, 0.650)	0.408 (-0.471, 1.287)
Travel time >10 min to health centre	-0.024 (-0.140, 0.092)	0.133 (-0.152, 0.418)	-0.150 (-0.672, 0.371)	-0.075 (-1.087, 0.937)
Number of recorded visits	0.028 (-0.075, 0.131)	<i>0.291 (0.208, 0.374)</i>	<i>0.178 (0.101, 0.254)</i>	<i>0.222 (0.144, 0.300)</i>
High user	0.081 (-0.023, 0.185)	0.072 (-0.185, 0.328)	0.241 (-0.251, 0.733)	-0.089 (-1.109, 0.930)
Almost always/always satisfied	0.037 (-0.066, 0.140)	0.147 (-0.098, 0.392)	0.343 (-0.114, 0.800)	0.529 (-0.336, 1.394)
Fair/poor self perceived health	0.031 (-0.015, 0.077)	0.112 (-0.002, 0.226)	0.185 (-0.025, 0.394)	0.361 (-0.033, 0.754)
Number of chronic illnesses	<i>0.103 (0.029, 0.178)</i>	0.036 (-0.142, -0.214)	<i>0.339 (0.013, 0.666)</i>	0.256 (-0.355, 0.866)
Seen by private physician	-0.057 (-0.173, 0.059)	0.134 (-0.144, 0.413)	0.094 (-0.425, 0.613)	-0.112 (-1.088, 0.865)
Seen by faith healer	0.196 (-0.035, 0.426)	<i>0.911 (0.358, 1.464)</i>	<i>1.024 (0.032, 2.016)</i>	0.234 (-1.677, 2.145)
Seen by alternative medicine practitioner	-0.071 (-0.182, 0.324)	-0.881 (-1.514, -0.248)	-1.547 (-2.695, -0.399)	0.060 (-2.211, 2.332)
Family	-0.002 (-0.000, 0.000)	0.000 (-0.000, 0.000)	-0.000 (-0.000, 0.000)	-0.000 (-0.000, 0.000)
Score on DUKE scale	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)	0.000 (-0.000, 0.000)
GHQ score	0.008 (-0.044, 0.041)	-0.018 (-0.123, 0.087)	0.046 (-0.148, 0.238)	0.037 (-0.324, 0.398)

Statistically significant coefficients and confidence intervals are in *italics*.

Table 6 Variables associated with reliability (DSR). Polychotomous regression analysis

Variable	2 weeks		3 months		6 months		1 year	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
<i>Model for underreporters</i>								
Male sex	0.94	0.27, 3.3	1.3	0.49, 3.5	0.75	0.30, 1.8	0.83	0.35, 2.0
Age	0.98	0.94, 1.0	1.0	1.0, 1.1	1.0	1.0, 1.1	1.0	1.0, 1.1
Widowed/divorced/separated	0.99	0.25, 3.9	0.96	0.36, 2.6	0.96	0.38, 2.4	1.5	0.62, 3.4
No schooling	0.48	0.15, 1.6	0.92	0.39, 2.1	0.55	0.24, 1.2	0.62	0.28, 1.4
Unemployed	0.15	0.00, 3.3	0.70	0.13, 3.7	0.99	0.18, 5.4	1.2	0.22, 6.7
Retired	4.2	0.82, 2.2	0.66	0.18, 2.4	1.1	0.31, 3.8	1.5	0.43, 5.2
Works at home	3.3	0.81, 1.4	0.81	0.25, 2.6	1.5	0.49, 4.7	1.3	0.41, 4.0
Other	3.1	0.44, 2.2	1.9	0.36, 1.0	3.5	0.83, 1.5	2.0	0.33, 1.2
Number of persons in household	1.4	0.42, 4.4	2.2	0.94, 5.1	1.1	0.49, 2.5	1.2	0.53, 2.5
Travel time >10 min to health centre	2.1	0.61, 7.0	2.3	0.92, 5.8	1.1	0.43, 2.7	0.99	0.43, 2.3
Number of recorded visits	35	14, 88	2.6	2.1, 3.2	1.5	1.3, 1.7	1.3	1.2, 1.3
High user	2.6	0.83, 8.4	1.3	0.56, 3.1	1.1	0.48, 2.4	2.1	0.94, 4.7
Almost always/always satisfied	0.45	0.16, 1.3	0.58	0.26, 1.3	0.98	0.45, 2.2	0.61	0.28, 1.3
Fair/poor self perceived health	1.5	0.88, 2.4	0.94	0.65, 1.4	0.97	0.68, 1.4	0.88	0.63, 1.2
Number of chronic illnesses	0.92	0.79, 1.1	0.98	0.87, 1.1	0.99	0.88, 1.1	0.95	0.86, 1.1
Seen by private physician	0.64	0.16, 2.6	1.6	0.59, 4.2	0.80	0.30, 2.1	1.2	0.49, 2.7
Seen by faith healer	1.4	0.15, 1.3	0.25	0.02, 2.7	1.0	0.16, 6.3	0.33	0.05, 2.4
Seen by alternative medicine practitioner	6.5	0.92, 4.6	4.4	0.91, 2.1	2.5	0.56, 1.1	3.9	0.87, 1.7
Family APGAR	0.96	0.76, 1.2	0.89	0.74, 1.1	0.90	0.76, 1.1	0.94	0.80, 1.1
Score on DUKE scale	1.0	0.95, 1.1	1.0	0.99, 1.1	1.0	0.97, 1.1	1.0	0.96, 1.0
GHQ scale	0.68	0.40, 1.2	0.77	0.54, 1.1	0.89	0.63, 1.3	0.88	0.63, 1.2
<i>Model for overreporters</i>								
Male sex	1.0	0.55, 1.9	1.3	0.69, 2.6	1.1	0.57, 2.3	1.1	0.49, 2.5
Age	0.99	0.97, 1.0	1.0	0.98, 1.0	0.99	0.97, 1.0	1.0	0.97, 1.0
Widowed/divorced/separated	0.91	0.48, 1.7	0.70	0.33, 1.5	0.86	0.41, 1.8	1.0	0.41, 2.5
No schooling	0.95	0.54, 1.6	0.91	0.49, 1.7	0.89	0.48, 1.7	0.72	0.34, 1.5
Unemployed	1.6	0.57, 4.5	0.89	0.29, 2.8	2.1	0.72, 6.2	2.7	0.84, 8.8
Retired	3.9	1.6, 9.4	1.4	0.54, 3.4	2.7	1.0, 7.3	1.9	0.61, 6.1
Works at home	2.2	0.99, 4.7	0.75	0.35, 1.6	1.5	0.62, 3.4	1.2	0.43, 3.1
Other	0.73	0.15, 3.6	0.24	0.03, 2.0	0.35	0.04, 3.0	0.47	0.05, 4.3
Number of persons in household	1.1	0.60, 1.9	1.3	0.69, 2.5	1.5	0.82, 2.9	2.6	1.2, 5.5
Travel time >10 min to health centre	0.79	0.43, 1.4	1.7	0.89, 3.1	0.96	0.49, 1.9	0.68	0.29, 1.6
Number of recorded visits	1.4	0.84, 2.3	1.3	1.1, 1.6	1.1	1.0, 1.3	1.2	1.1, 1.3
High user	1.3	0.76, 2.2	1.2	0.65, 2.2	1.2	0.62, 2.2	0.78	0.36, 1.7
Almost always/always satisfied	1.2	0.66, 2.0	1.9	0.98, 3.7	2.1	1.0, 4.1	2.3	1.0, 5.4
Fair/poor self perceived health	1.2	0.94, 1.5	1.3	0.97, 1.7	1.3	1.0, 1.8	1.6	1.2, 2.3
Number of chronic illnesses	1.1	1.0, 1.2	1.1	0.98, 1.1	1.0	0.96, 1.1	1.0	0.92, 1.1
Seen by private physician	0.80	0.43, 1.5	1.9	0.97, 3.6	1.3	0.67, 2.6	1.4	0.62, 3.1
Seen by faith healer	2.8	1.0, 7.9	4.2	1.3, 14	3.3	0.99, 1.1	3.0	0.79, 1.1
Seen by alternative medicine practitioner	1.2	0.37, 4.0	0.21	0.03, 1.3	0.25	0.04, 1.5	1.2	0.26, 5.7
Family APGAR	0.99	0.87, 1.1	0.99	0.86, 1.1	0.96	0.83, 1.1	0.97	0.81, 1.2
Score on DUKE scale	1.0	1.0, 1.1	1.0	0.99, 1.1	1.0	0.98, 1.0	1.0	0.98, 1.1
GHQ scale	1.1	0.85, 1.3	0.93	0.72, 1.2	1.0	0.79, 1.3	1.0	0.73, 1.4

OR: adjusted odds ratio; statistically significant OR and confidence intervals are in *italics*.

recorded visits, followed by older age. The number of recorded visits was also associated with overreporting, although the magnitude of the association was smaller than for underreporting. Overreporting was also associated with worse subjective perception of health, a greater number of chronic illnesses, visits to faith healers, satisfaction with the health care system, and retirement.

Discussion

OVERALL ASSESSMENT OF VALIDITY

Our results seem to indicate that the accuracy of information provided in user surveys of the number of visits is relatively low. Other authors have found low validity for self reported utilisation in comparison with recorded utilisation. Roberts *et al*³⁰ obtained a Spearman correlation coefficient of 0.58, a value similar to ours and much higher than the 0.37 reported by Reynoso *et al*.⁴⁷ Glandon *et al*⁴⁸ also reported relatively low validity in a study that used a six month recall period before interview. The proportion of concordant results they found (28.2%) was very close to the 29.9% obtained in this study.

In comparison with some earlier findings,^{30 48} overreporting of the number of visits was more

frequent in our sample. None the less, underreporting by our subjects seemed to be linked to the longer duration of the recall period, as found by Roberts *et al*.³⁰ One possible explanation for the excess number of recalled visits is that although the participants were instructed not to count unscheduled visits, they did not distinguish between visits that were being investigated and unplanned visits.

CHANGE IN ACCURACY WITH INCREASING RECALL PERIOD

Accuracy, expressed in absolute terms as a proportion of concordant results or as the reciprocal of the standard deviation of the DSR, clearly decreased as the recall period lengthened. However, neither of these measures is entirely suitable to compare accuracy across periods, as they do not take the differences in maximum variability between periods into account. The longer the period, the greater the total number of visits, and hence the greater margin of potential variability between the number of recorded visits and the number of recalled visits. Moreover, the increase in maximum variability as recall period lengthened probably accounts for the

higher uncorrected correlation coefficients obtained for increasing recall periods. However, our analysis suggests that in relative terms (when different maximum variabilities in each period are taken into account), the validity of self reported utilisation does not decrease with time, at least for periods shorter than one year. All periods yielded similar corrected Spearman correlation coefficients, and there were no significant differences in the slopes of the regression of SU upon RU. Moreover, the one year period was the only one for which the Wilcoxon rank test did not detect a significant difference between SU and RU distributions. Finally, when the distribution of the logarithm of SU/RU was compared between periods, the proportion of accurate responses was practically equal in all of them.

FACTORS ASSOCIATED WITH VALIDITY

The number of recorded visits was strongly associated with validity of the number of visits recalled: the more visits, the greater the bias toward both underreporting and overreporting, although the magnitude of the former association was clearly higher. This finding suggests, in agreement with other studies,^{30 31 48} that the more recorded visits, the greater the likelihood that some visits will be forgotten. The association between increasing age and greater frequency of underreporting was expected, in view of the greater likelihood that older subjects would be more forgetful. Other studies also found this association,^{30 31 48} although the relation between the magnitude of underreporting and duration of the recall period was not investigated.

In our study the main factors directly associated with overreporting (poor health status, retirement, or high level of satisfaction with health care), or inversely associated with underreporting (low educational level), are those usually described as related with a higher use of primary health care services.^{36 49-51} Although we included in the multiple regression models the actual number of visits (RU) and HU compared with NU status, the confounding effect of these variables might not have been completely removed. However, the opposite could also be true: as in the study by Cleary and Jette,³¹ our findings indirectly suggest that inaccuracy of self recorded utilisation by some subjects may distort the analysis of other factors related with utilisation. The association of poor health status with overreporting was also found by Cleary and Jette³¹ and by Glandon *et al*⁴⁸ in a study of elderly users. The tendency of persons who visit healers to overestimate the number of visits may be partially explained by confusion regarding the type of visit for which information was solicited.

METHODOLOGICAL CONSIDERATIONS

To control for potential confusion introduced by the original stratification of the sample in high and normal users, all results of the regression analyses were adjusted to take the level of use into account.

KEY POINTS

- Self reported utilisation has a considerable degree of inaccuracy, with a net tendency to overestimate the actual number of visits. It seems advisable to assess the validity of studies based on this measure of utilisation carefully.
- Validity expressed in relative terms does not seem to decrease appreciably as the recall period lengthens. It therefore seems advisable to use a recall period of at least one year to compensate for biases present in user responses for short recall periods.
- Because estimates that do not take into account the maximum variability of each variable can yield contradictory results, they should not be used to compare the accuracy of different measures.
- The number of recorded visits is the variable that most strongly influences the validity of self reported use: the greater the use (number of visits), the greater the tendency to underreport and, to a lesser degree, to overreport the number of visits. Other variables that influence the validity of self reported number of visits are older age (associated with underreporting) and poor perceived health (associated with overreporting).

In comparison with responders, most of the subjects who did not respond to the item about the number of visits made shared characteristics generally associated with lower accuracy: older age, worse perceived general health, and greater number of recorded visits. This means that if we had been able to obtain information for the entire sample, the validity of self reported visits may have been even lower. We found no published studies that examined this source of bias in the estimation of the validity of self reported utilisation.

It should be noted that our validity analysis was limited to scheduled visits. This approach may differ from that used in studies of other types of contact with different health care facilities. For example, Roberts *et al*⁵⁰ found a high validity for self reported days spent in the hospital, even for recall periods as long as one year.

We thank Karen Shashok for translating substantial parts of the original manuscript into English.

Funding: none.

Conflicts of interest: none.

- 1 Hulka BS, Wheat JR. Patterns of utilization. The patient perspective. *Med Care* 1985;23:438-59.
- 2 Rekola KE, Keinänen-Kiukaanniemi S, Takala J. Use of primary health services in sparsely populated country districts by patients with musculoskeletal symptoms: consultations with a physician. *J Epidemiol Community Health* 1993;47:153-7.
- 3 Hurwicz ML, Berkanovic E. Care seeking for musculoskeletal and respiratory episodes in a Medicare population. *Med Care* 1991;29:1130-45.
- 4 Barsky AJ, Ahern DK, Bailey D, *et al*. Predictors of persistent palpitations and continued medical utilisation. *J Fam Pract* 1996;42:465-72.
- 5 Mechanic D, Cleary PD, Greenley JR. Distress syndromes, illness behavior, access to care and medical utilisation in a defined population. *Med Care* 1982;20:361-72.

- 6 Penning MJ. Health, social support, and the utilisation of health services among older adults. *J Gerontol* 1995;**50B**: S330-9.
- 7 Groenewegen PP, Hutten JBF. The influence of supply-related characteristics on general practitioners' workload. *Soc Sci Med* 1995;**40**:349-58.
- 8 Hurley J, Woodward C, Brown. Changing patterns of physician services utilisation in Ontario, Canada, and their relation to physician, practice, and market-area characteristics. *Med Care Res Rev* 1996;**53**:179-206.
- 9 Wagner PJ, Phillips W, Radford M, *et al.* Frequent use of medical services. Patients reports of intentions to seek care. *Arch Fam Med* 1995;**4**:594-9.
- 10 Potvin L, Camirand J, Beland F. Patterns of health services utilisation and mammography use among women aged 50 to 59 years in the Québec Medicare System. *Med Care* 1995;**33**:515-30.
- 11 Payne SM, Ash A, Restuccia JD. The role of feedback in reducing medically unnecessary hospital use. *Med Care* 1991;**29** (suppl):AS91-106.
- 12 Treviño FM, Moyer E, Valdez RB, *et al.* Health insurance coverage and utilisation of health services by Mexican Americans, Mainland Puerto Ricans, and Cuban Americans. *JAMA* 1991;**265**:233-7.
- 13 Rosenberg MW, Hanlon NT. Access and utilisation: a continuum of health service environments. *Soc Sci Med* 1996;**43**: 975-83.
- 14 Zook CJ, Moore FD. High-cost users of medical care. *N Engl J Med* 1980;**302**:996-1002.
- 15 Lloyd P, Lupton D, Donaldson C. Consumerism in the health care setting: an exploratory study of factors underlying the selection and evaluation of primary medical services. *Aust J Public Health* 1991;**15**:194-201.
- 16 Starfield B, Hankin J, Steinwachs D, *et al.* Utilization and morbidity: random or tandem? *Pediatrics* 1985;**75**:241-7.
- 17 Freeborn DK, Pope CR, Mullooly JF, *et al.* Consistently high users of medical care among the elderly. *Med Care* 1990;**28**:527-40.
- 18 Broadhead WE, Gehlbach SH, deGruy FV, *et al.* Functional versus structural social support and health care utilization in a family medicine outpatient practice. *Med Care* 1989;**27**:221-33.
- 19 Riley AW, Finney JW, Mellits D, *et al.* Determinants of children's health care use: an investigation of psychosocial factors. *Med Care* 1993;**31**:767-83.
- 20 Stoller EP, Forster LE, Portugal S. Self-care responses to symptoms by older people. A health diary study of illness behavior. *Med Care* 1993;**31**:24-42.
- 21 Osman LM, Dunt D. Factors influencing mothers' decisions to consult a general practitioner about their children's illnesses. *Br J Gen Pract* 1995;**45**:310-12.
- 22 Browne GB, Humphrey B, Pallister R, *et al.* Prevalence and characteristics of frequent attenders in a prepaid Canadian family practice. *J Fam Pract* 1982;**14**:63-71.
- 23 Rosner TT, Namazi KH, Winkle ML. Physician use among the old-old. Factors affecting variability. *Med Care* 1988;**26**: 982-91.
- 24 Counte MA, Glandon GL. A panel study of life stress, social support, and the health services utilisation of older persons. *Med Care* 1991;**29**:348-61.
- 25 Klerman GL, Weissman MM, Oullette R, *et al.* Panic attack in the community: Social morbidity and health care utilization. *JAMA* 1991;**265**:742-6.
- 26 Wetzler HP, Cruess DF. Self reported physical health practices and health care utilization: Findings from the National Health Interview Survey. *Am J Public Health* 1985;**5**:1329-30.
- 27 Grimsmo A, Siem H. Factors affecting primary health care utilisation. *Fam Pract* 1984;**3**:155-61.
- 28 Andersen AS, Laake P. A model for physician utilisation within 2 weeks. Analysis of Norwegian data. *Med Care* 1987;**25**:300-10.
- 29 Newacheck PW. Characteristics of children with high and low usage of physician services. *Med Care* 1992;**30**:30-42.
- 30 Roberts RO, Bergstralh EJ, Schmidt L, *et al.* Comparison of self-reported and medical record health care utilization measures. *J Clin Epidemiol* 1996;**9**:989-95.
- 31 Cleary PD, Jette AM. The validity of self-reported physician utilization measures. *Med Care* 1984;**22**:796-803.
- 32 Bellón JA, Delgado A, Luna JD, *et al.* Psychosocial and health belief variables associated with frequent attendance in primary care. *Psychol Med* 1999;**29**:1347-57.
- 33 Hart JT. Primary medical care in Spain. *Br J Gen Pract* 1990;**40**:255-8.
- 34 Instituto Nacional de Estadística. *Censo poblacional*. Madrid: Instituto Nacional de Estadística: 1991.
- 35 Domingo A, Marcos J. Proposal of an indicator of social class based on the occupation. *Gac Sanit* 1989;**10**:320-6.
- 36 Ministerio de Sanidad y Consumo. *Encuesta Nacional de Salud*. Madrid: Servicio de Publicaciones del Ministerio De Sanidad y Consumo, 1989.
- 37 Goldberg DP, Hiller VF. A scaled version of the General Health Questionnaire. *Psychol Med* 1979;**9**:139-48.
- 38 Lobo A, Pérez-Echevarría MJ, Artal J. Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. *Psychol Med* 1986;**16**:135-40.
- 39 Smilkstein G. The family APGAR: a proposal of a family function test and its use by physicians. *J Fam Pract* 1978;**6**: 1231-41.
- 40 Smilkstein G, Ashworth C, Montano D. Validity and reliability of the family APGAR as a test of family function. *J Fam Pract* 1982;**15**:303-11.
- 41 Bellón JA, Delgado A, Luna JD, *et al.* Validity and reliability of the APGAR-family questionnaire on family function. *Aten Primaria* 1996;**18**:289-96.
- 42 Broadhead WE, Gehlbach SH, Degruy FV, *et al.* The DUKE-UNC functional social support questionnaire: measurement of social support in family medicine patients. *Med Care* 1988;**26**:709-23.
- 43 Bellón JA, Delgado A, Luna JD, *et al.* Validity and reliability of the questionnaire on functional social support DUKE-UNC-11. *Aten Primaria* 1996;**18**:153-63.
- 44 Shih WJ, Huang WM. Evaluating correlation with proper bounds. *Biometrics* 1992;**48**:1207-13.
- 45 Norusis MJ. *SPSS for Windows. Base system and advanced user's guides. Release 6.0*. Chicago: SPSS Inc, 1993.
- 46 Dixon WJ. *BMDP statistical software manual. Version 7.0*. Princeton: University of California Press, 1992.
- 47 Reynoso L, Cronan T, Foster M. Are self-reports of health care utilization valid? *Ann Behav Med* 1995;**17** (suppl): S197-211.
- 48 Glandon GL, Counte MA, Tancredi D. An analysis of physician utilization by elderly persons: systematic differences between self-report and archival information. *J Gerontol* 1992;**47**:S245-52.
- 49 Fylkesnes K, Forde OH. Determinants and dimensions involved in self-evaluation of health. *Soc Sci Med* 1992;**35**: 271-9.
- 50 Cook DG, Morris JK, Walker M, *et al.* Consultation rates among middle aged men in general practice over three years. *BMJ* 1990;**301**:647-50.
- 51 Hojtdahl P, Laerum E. Continuity of care in general practice: effect on patient satisfaction. *BMJ* 1992;**304**: 1287-90.