ORIGINAL RESEARCH

Do people self-reporting information about chronic respiratory disease have corroborative evidence in their general practice medical records? A study of intermethod reliability

*Lisa Iversen[a], Philip C Hannaford[a], David J Godden[b], David B Price[a]

[a] Dept of General Practice & Primary Care, University of Aberdeen, Foresterhill Health Centre, Westburn Road, Aberdeen, AB25 2AY, UK
[b] Centre for Rural Health, The Greenhouse, Beechwood Business Park, Inverness, IV2 3BL, UK

Abstract

Aims: To use intermethod reliability to compare self-reported data about chronic respiratory disease and health service utilisation with data contained in general practice medical records.

Methods: Self-reported postal questionnaire information from a small cohort of an age-sex stratified sample of 2318 patients was compared with information contained in their medical records. The agreement between the two sources of information was assessed.

Results: The case notes of 115/135 individuals from eight general practices were examined. For self-reported chest injury or operation (kappa, κ=-0.03), or chronic bronchitis (κ=0.10), agreement was poor. Agreement for self-reported pleurisy (κ=0.32), hay fever or rhinitis (κ=0.40), or eczema or dermatitis (κ=0.30) was fair; for chronic obstructive pulmonary disease (COPD) or emphysema (κ=0.56), or heart trouble (κ=0.54), agreement was moderate; for asthma (κ=0.78) or pneumonia (κ=0.62), agreement was good; and for pulmonary tuberculosis (κ=0.88), agreement was very good. The strength of agreement for information about health service utilisation for respiratory problems ranged from moderate to very good and was good for smoking status.

Conclusions: Although based on small numbers, our results suggest good or very good agreement between self-reported data and general practice medical records for the absence or presence of some respiratory conditions and some types of respiratory-related health care utilisation. Depending on the research question being examined self-reported information may be appropriate.

© 2007 General Practice Airways Group. All rights reserved.
L Iversen, *et al. Prim Care Resp J* 2007; **16**(3): 162-168.
doi:10.3132/pcrj.2007.00013

Keywords agreement, chronic respiratory disease, intermethod reliability, medical records, postal questionnaire, self-report

Introduction

Outcomes research and economic evaluations often depend on self-reported information about health service utilisation, with remarkably little consideration given to the quality of the data.[1] This information is often obtained via a postal questionnaire. In 2005, we published prevalence estimates of chronic respiratory disease among adults living in Scotland based on self-reported information collected using a postal questionnaire.[2] Currently, little is known about the relationship between asthma reported in response to a postal questionnaire and an actual clinical diagnosis of asthma.[3] Similar difficulties affect epidemiological studies of chronic obstructive pulmonary disease (COPD), especially since current clinical definitions of COPD and its severity rely heavily on spirometry – an impractical investigation for large scale epidemiological studies such as ours.[4] Ideally, self-reported information should be assessed against an error-free, gold standard. However, in the absence of such a gold standard – as is the case for the diagnosis of chronic respiratory disease – intermethod reliability can be measured. This approach

* Corresponding author: Tel: +44 (0)1224 554226   Fax: +44 (0)1224 550683   Email: l.iversen@abdn.ac.uk

compares measurements made during an epidemiological study (e.g. in response to a postal questionnaire) with more burdensome but possibly more accurate measurements (e.g. clinical diagnosis or features recorded in medical records).[5] We used intermethod reliability to compare self-reported data about chronic respiratory disease and health service utilisation supplied for our epidemiological study, with corresponding information contained in the general practice medical records of a sample of respondents.

Methods

Details of our study have been published elsewhere.[2] Briefly, during 2001-2002 an age-sex stratified sample of 4560 adults registered with 57 general practices located in Scotland were

sent a postal questionnaire. The questionnaire collected information about age, gender, socio-economic factors, smoking habits, selected illnesses (major respiratory and atopic diseases, and other major chronic conditions including heart disease), respiratory symptoms (cough, phlegm, breathlessness and wheeze), use of health services and medicines, and quality of life. After two mailings a corrected response rate – i.e. after individuals were removed from the denominator if they did not receive a questionnaire (for example, if it was returned undelivered) – of 60% (2603 of 4376 questionnaires) was achieved. Most of the respondents who returned a consent form with the questionnaire agreed that their medical records could be examined (2318/2545: 91%).

Figure 1 shows the study profile. Nine of the 57 general

Figure 1. Flowchart of study methods.

L Iversen *et al.*
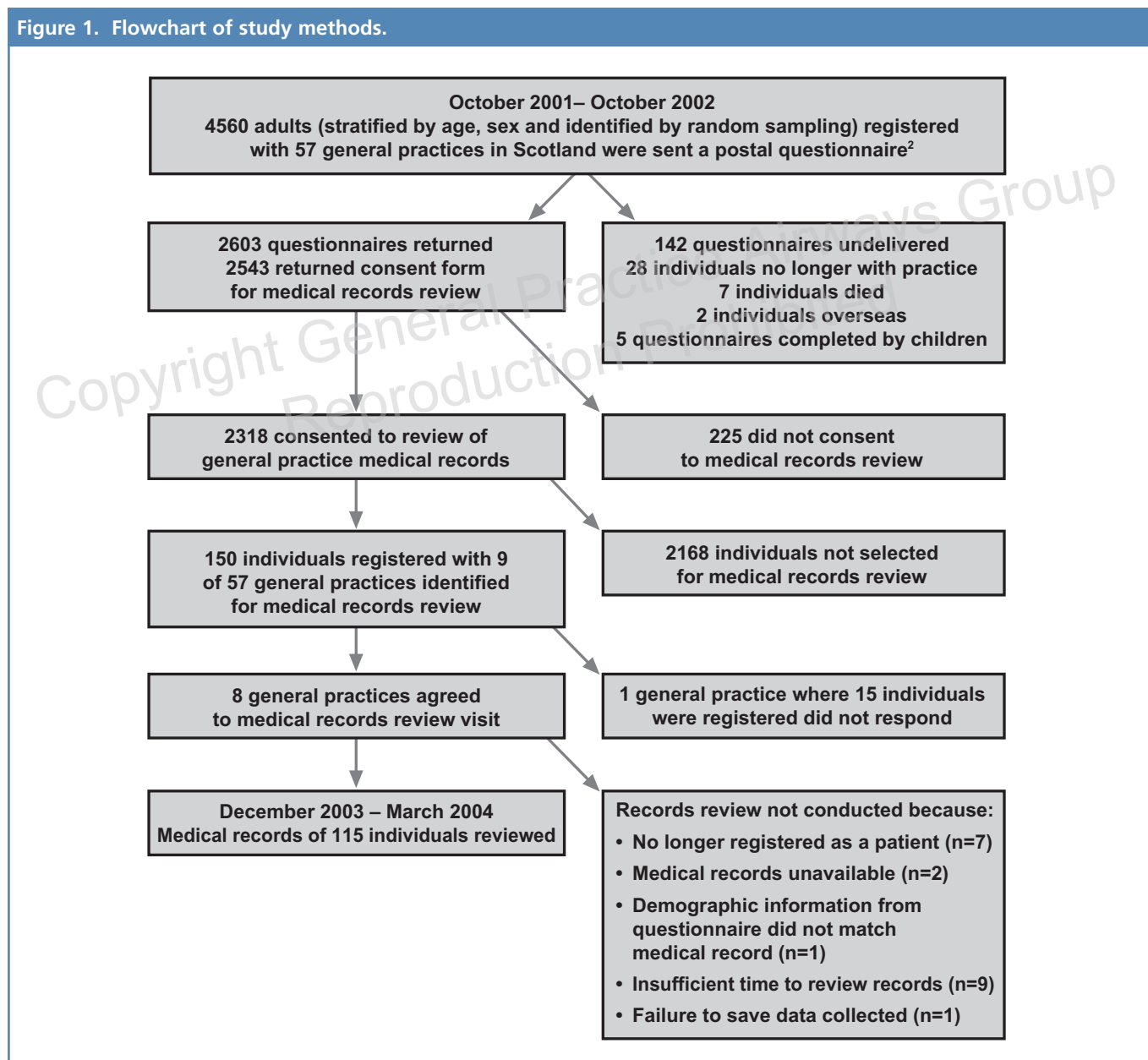
practices were approached to request permission to undertake a case notes review. Practices were selected on the basis that there had been a good response to the postal questionnaire and that there was a good geographical distribution which reflected the Scottish population – i.e. most were in the west and central areas. All selected practices used the General Practice Administration System for Scotland (GPASS) clinical computer system[6] and none were paperless. To identify respondents whose notes we wanted to review, we generated for each practice a list of the study identification numbers of everyone who had consented to having their medical records reviewed. We stratified respondents into those who did and those who did not report having at least one respiratory condition. Everyone who reported a respiratory condition was selected for case notes review, together with a random sample, minimum five, of those not reporting such problems. This procedure was followed until 150 individuals (100 with and 50 without any of the respiratory conditions) were chosen for review. No formal sample size calculation was done for this study since it was part of a doctoral research programme; the number of notes for review was based solely on feasibility and time constraints.

Once the list of individuals selected for review was compiled, details of respiratory status were removed so that the review was blinded. This meant that data extraction from the respondents' medical records was not influenced by their self-report (in the questionnaire) of the presence or absence of a history of respiratory disease. For feasibility reasons, the number of notes reviewed at each practice was usually limited to that which could be completed in one day. Each practice was contacted with the study identification numbers for selected respondents so that it could identify the relevant people from its master study list. The practice then retrieved the medical notes of identified participants for the review, which was carried out at the practice. There was a time lag of two to three months between compiling the lists and the review of medical records. Apart from the first practice visited where two researchers (LI and PCH) reviewed the notes and inter-rater reliability was assessed (by visual inspection alone, with very good agreement found), all other records were reviewed by just one of the authors (LI).

Although the review of medical records occurred some time after the return of the postal questionnaires, searching was done in relation to the date the questionnaire was originally sent to each participant. The medical records were examined systematically, first for a record (ever) of the terms: asthma; injury or operation to the chest; chronic bronchitis; COPD; emphysema; pleurisy; pneumonia; pulmonary tuberculosis; hay fever or rhinitis; eczema or dermatitis; and heart disease (i.e. myocardial infarction, left ventricular dysfunction, angina, atrial fibrillation, aortic valve disease or

cardiac failure – hypertension was not included as evidence of heart disease). The notes were then similarly examined for: any courses of antibiotics or steroids prescribed in the 12 months prior to the postal survey; current receipt of respiratory medicines; and any documented health service use (i.e. an appointment with a general practitioner (GP) or practice nurse, a home visit, use of the out of hours medical services, referral to hospital, attendance at/admission to hospital, or use of an accident and emergency department for respiratory problems) in the previous year. The notes were also searched for any lifetime admissions to hospital for respiratory problems, and the most recently recorded smoking status of the individual. The practice computer system was then searched for details of any diagnoses or current respiratory prescriptions not recorded in the paper records. The medication items would normally be repeat prescriptions, often issued without the patient seeing their GP.

Data were entered directly into a database held on a laptop computer. Following each practice visit, the data were transferred into SPSS for Windows (version 11.5). For each outcome of interest, the proportion of agreement between self-reported data and medical records was tabulated. Kappa (κ) statistics and their corresponding 95% confidence interval (computed using the Confidence Interval Analysis package version 2.0.0[7]) were calculated, to assess agreement between self-reported data and the medical records. Although κ values vary between 0.00 and 1.00, it is common to group them to assist interpretation: poor (<0.20); fair (0.21 to 0.40); moderate (0.41 to 0.60); good (0.61 to 0.80) and very good (0.81 to 1.00).[8] Negative values indicate agreement that is less than that which would be expected by chance.

For asthma, chronic bronchitis, COPD or emphysema, pleurisy and pneumonia, instances of disagreement (i.e. no corroborative evidence found in the case notes) were investigated to see if there was evidence of alternative respiratory diagnoses in the medical records. Disagreement about chest injury or operation, eczema or dermatitis, hay fever or rhinitis, heart troubles or pulmonary tuberculosis, were not explored further due to a lack of possible alternative diagnoses.

Results

Eight of the nine practices approached gave permission for the case notes review. The remaining practice did not respond to our request to visit and therefore the target number of records to review reduced from 150 to 135. Of the 135 sets of medical records requested, 115 sets (85%) were reviewed including the records of 22 individuals who did not report having a history of any respiratory condition in the postal questionnaire. The remaining 20 could not be reviewed because of the following reasons: respondents had left the practice (n=7);

Table 1. Comparison of self-reported conditions with general practice medical records.

Self-reported condition		Evidence from medical records				Total*	κ (95% CI)
		Yes		No			
		n	%	n	%		
Asthma	Yes	48	(85.7)	8	(14.3)	56	0.78 (0.66-0.90)
	No	3	(6.8)	41	(93.2)	44	
Chest injury/operation	Yes	0	(0.0)	4	(100.0)	4	-0.03 (-0.83-0.77)
	No	2	(2.2)	91	(97.8)	93	
Chronic bronchitis	Yes	3	(14.3)	18	(85.7)	21	0.10 (-0.22-0.42)
	No	5	(6.4)	73	(93.6)	78	
COPD or emphysema	Yes	7	(87.5)	1	(12.5)	8	0.56 (0.29-0.83)
	No	8	(9.2)	79	(90.8)	87	
Eczema or dermatitis	Yes	9	(42.9)	12	(57.1)	21	0.30 (0.05-0.56)
	No	10	(13.5)	64	(86.5)	74	
Hay fever or rhinitis	Yes	19	(55.9)	15	(44.1)	34	0.40 (0.20-0.60)
	No	11	(17.2)	53	(82.8)	64	
Heart trouble	Yes	13	(72.2)	5	(27.8)	18	0.54 (0.33-0.76)
	No	10	(12.2)	72	(87.8)	82	
Pleurisy	Yes	3	(21.4)	11	(78.6)	14	0.32 (-0.06-0.70)
	No	0	(0.0)	84	(100.0)	84	
Pneumonia	Yes	8	(61.5)	5	(38.5)	13	0.62 (0.37-0.87)
	No	3	(3.5)	82	(96.5)	85	
Pulmonary tuberculosis	Yes	4	(80.0)	1	(20.0)	5	0.88 (0.66-1.11)
	No	0	(0.0)	88	(100.0)	88	

* Row totals and percentages presented. The number of individuals in each comparison varies because some respondents did not answer a particular item on the questionnaire. Individuals could report more than one condition. κ = kappa statistic

respondents were still registered with the practice but their notes were unavailable (n=2); the demographic information provided by one respondent in the questionnaire did not match that in their record; there was insufficient time to review the records (n=9); and failure to save data (n=1). The mean age (at the time of the postal survey) of the respondents whose records were reviewed was 53.0 years (standard deviation=16.7); 67 (58.3%) were women and 48 (41.7%) were men.

There was poor agreement for information about self-reported chest injury or operation, or chronic bronchitis (see Table 1). Agreement for self-reported pleurisy, hay fever or rhinitis, or eczema or dermatitis was fair; for COPD or emphysema, or heart trouble, agreement was moderate; for asthma or pneumonia, agreement was good; and for pulmonary tuberculosis, agreement was very good.

Of the eight respondents who reported having asthma for whom no corroboration could be found in the medical records, one had evidence of chronic bronchitis and two COPD or emphysema. If these three respondents were re-classified into the group of respondents with corroborative evidence of asthma, the κ statistic improved to 0.84 – very good agreement.

There were 18 instances where chronic bronchitis was reported on the questionnaire without corroborative evidence in the notes; ten of these respondents had asthma in their notes and three COPD or emphysema. If these 13 respondents who had evidence of a different respiratory disease were re-classified into the group of respondents with corroborative evidence of chronic bronchitis, then the κ statistic improved to 0.70 (good agreement). One respondent reported COPD or emphysema in the postal survey without corroborative evidence or evidence of an alternative respiratory condition either.

Eleven respondents reported having had pleurisy for which there was no corroborative evidence in their medical records; one respondent had a diagnosis of pneumonia in the case notes. Re-classifying this respondent would improve the κ statistic to 0.41 (moderate agreement). None of the five respondents who reported pneumonia without corroborative evidence had pleurisy noted in their medical records.

There was moderate agreement for information about attending a GP in the previous 12 months for a respiratory problem, or ever being admitted to hospital; good agreement for receipt of a current prescription and for having had at least one course of antibiotics in the previous 12 months; and very

L Iversen *et al*.

Table 2. Comparison of self-reported information about health service utilisation for respiratory problems with general practice medical records.

Self-reported health service use for respiratory problems		Evidence from medical records					
		Yes		**No**			
		n	%	n	%	Total*	κ (95% CI)
Attended GP in last 12 months	Yes	32	(68.1)	15	(31.9)	47	0.54 (0.38 to 0.70)
	No	10	(14.9)	57	(85.1)	67	
Ever hospital admission	Yes	20	(69.0)	9	(31.0)	29	0.60 (0.43 to 0.78)
	No	8	(9.4)	77	(90.6)	85	
Receipt of a current prescription	Yes	38	(90.5)	4	(9.5)	42	0.67 (0.51 to 0.83)
	No	10	(23.3)	33	(70.7)	43	
Had a course of steroids in the last year	Yes	13	(100.0)	0	(0.0)	13	0.92 (0.81 to 1.03)
	No	2	(2.1)	94	(97.9)	96	
Had a course of antibiotics in the last year	Yes	23	(76.7)	7	(23.3)	30	0.69 (0.54 to 0.85)
	No	6	(8.0)	69	(92.0)	75	

*Row totals and percentages are presented.

Table 3. Comparison of self-reported smoking status with smoking status as recorded in general practice medical records.

Self-reported smoking status	Smoking status from general practice medical records						
	Never		**Former**		**Current**		
	n	(%)	n	(%)	n	(%)	Total*
Never	39	(84.8)	7	(15.2)	0	(0.0)	46
Former	3	(8.6)	25	(71.4)	7	(20.0)	35
Current	0	(0.0)	8	(32.0)	17	(68.0)	25

*Row totals and percentages are presented.

good agreement for having had a course of steroids during the previous 12 months (Table 2). Information about smoking status was recorded in the majority of the case notes examined (106/115: 92%) and there was good agreement (κ =0.64, 95% CI 0.51 to 0.76) for this information (Table 3).

Discussion

We have found that if respondents denied having a respiratory condition, only a small proportion (less than 10%) had contradictory evidence in their records. Respondents were more likely to report a medical condition for which there was no corroborative evidence in the notes, especially for conditions where "self-diagnosis" could occur – for example, hay fever. The strength of agreement for information about health service utilisation for respiratory problems ranged from moderate to very good. The strength of agreement for smoking status was good.

Although no formal statistical testing was carried out, good inter-rater reliability was achieved when two researchers

reviewed the medical records, suggesting that the data extraction was accurate and consistent across different case notes. In addition to being blinded to the respiratory status of respondents at the time of review, the time lag (2-3 months) between compiling the lists of individuals and the case notes review makes information bias from prior knowledge of disease status unlikely. Standardised data abstraction procedures helped ensure that there was no difference between groups (e.g. those with respiratory and non-respiratory conditions) in the quality of information gathering.

A criticism of self-reported data is that it might be affected by recall bias. As a source of corroboration for self-reported diagnoses, general practice medical records avoid the problems of patient recall bias. However, they cannot be considered a gold standard since the practice records may be incomplete or inaccurate. For example, information about use of health care services external to the general practice might be incomplete or missing. A study comparing patient-recalled health care utilisation with that detailed in primary care

records found an apparent under-recording in the medical records of consultations for chronic conditions, and of multiple problems presented at one consultation.[9] General practice records are also an unsuitable sampling frame for the study of respiratory symptoms because the majority of symptoms are managed in the community without seeking health care – the 'symptom iceberg' phenomenon.[10,11] Nevertheless, despite the difficulties, general practice records are still the most likely source of corroborative evidence for self-reported information about diagnoses and health care utilisation,[12] at least in health care systems where GPs provide most primary care and remain gatekeepers to secondary care.

Hay fever or rhinitis, and eczema or dermatitis, are conditions that many people may self-diagnose and manage without ever presenting to general practice. It is perhaps not surprising therefore, that agreement between self-reported information about these conditions and the case notes was only fair. There might also be under-reporting of some conditions in relation to the medical records due to patients misunderstanding or forgetting diagnoses. For example, it is possible that patients with atrial fibrillation might not necessarily report a history of heart troubles. This would reduce the strength of agreement between the two sources of information. The case notes review concentrated on respiratory disease, hay fever/rhinitis, eczema/dermatitis and heart troubles. Our paper, therefore, provides no information about the strength of agreement between self-report and general practice medical records for other conditions.

For practical reasons we were only able to review the case notes of a small proportion (115/2318: 4.9%) of respondents who were willing for this procedure to be conducted. Reviewing a relatively small number of notes meant that the results were relatively imprecise, resulting in wide confidence intervals around the κ estimates. It also meant that there was little data about respiratory conditions which would be reported infrequently. The small numbers in such cases could produce misleading perceptions about the level of agreement between the two sources of information since κ values depend on prevalence.[8] However, we have followed convention by presenting the proportions in each category as well as κ statistics. The estimates of prevalence of each condition vary between the self-reported data and that in the medical records, and are likely to vary among different individuals. However, our small sample size studied prevents calculation of age- and sex-adjusted estimates of prevalence.

In many instances, evidence of other respiratory diagnoses was found when disagreement between questionnaire responses and medical records occurred. Although the strength of agreement improved when these alternative diagnoses were included, such actions might not be appropriate – for example, when considering an entry of chronic bronchitis in the medical records as corroborative evidence of self-reported asthma. Evidence of other respiratory conditions, however, may illustrate the fact that chronic respiratory disease can be given different labels by doctors and/or patients. For example, the nomenclature for COPD has changed over time, frequently being referred to by doctors in the UK in the 1950s as "chronic bronchitis," and as "emphysema" by doctors in the United States.[13]

Few studies in the literature have compared respiratory information collected by postal questionnaire with medical records and it is difficult to compare studies since they are based on different populations. One study found that nearly 14% of respondents (n=1112) self-reported symptoms of asthma-like illness, with only 52.4% having corroborative evidence of asthma in their medical records.[10] This proportion was lower than in our study and might reflect the afore-mentioned 'symptom iceberg' phenomenon.

Every two years since 1988, the Nurses' Health Study has used a postal questionnaire to ask about physician-diagnosed COPD and asthma.[14] Between 1988 and 1996, 2790 nurses met the study criteria for self-reported COPD, derived from responses to the biennial questionnaires. In a 10% random sample of these women, corroborative evidence was found in 78% of medical records reviewed, leading the authors to conclude that self-reported COPD is a valid marker of medical record evidence of the condition. In our study, the level of corroborative evidence for self-reported COPD was higher (87.5%) but this figure was based on a very small number of individuals.

Mohangoo and colleagues recently compared prevalence estimates of asthma and COPD obtained by self-report during an interviewer-administered questionnaire with estimates from the Dutch National Survey of General Practice.[15] Analyses of data on 9411 individuals who consulted their GP during the 12-month period found moderate agreement between self-reported asthma or COPD and GP-diagnosed asthma or COPD, similar to the strength of agreement we found for asthma (good) and COPD or emphysema (moderate).

Self-reported data collected through telephone interviews in Ontario about use of health services by respiratory patients was compared with the provincial health service administrative database.[16] Agreement was higher for health services that were rare and prominent from the patient's perspective, such as hospital admissions in the previous 12 months (κ=1.00) and visits to respiratory specialists (κ=0.74). The study found fair agreement for GP respiratory visits in the previous six months – in our study, agreement was good for lifetime-ever hospital admission, and moderate for GP respiratory visits in the previous 12 months – and the authors concluded that in economic evaluations, self-reports of health services utilisation should be combined with other sources of information.

L Iversen *et al.*

Although based on small numbers, our results suggest that this conclusion may be unduly pessimistic. Our findings suggest good or very good agreement between self-reported data and medical records, for the absence or presence of particular respiratory disorders (e.g. asthma), and some types of respiratory-related health care utilisation (e.g. use of steroids). Depending on the research question asked, self-reported data may be fit for purpose. For studies of the "true" prevalence of different respiratory conditions, professional or lay information may be unreliable. In this case, objective measures of disease using internationally accepted definitions would be preferred, although the cost of such studies can be considerable, frequently making large-scale studies impractical. However, for studies of the health care utilisation associated with respiratory disease, or community-based geographical comparisons of broad categories of respiratory disease such as ours,[2] self-reported information appears to be sufficient.

Acknowledgements

Conflict of interest declaration

Professor David B Price is a member of the international editorial board of the *Primary Care Respiratory Journal*. Otherwise the authors have no conflicts of interest.

References

1. Ungar WJ, Coyte PC. Health services utilization reporting in respiratory patients. Pharmacy Medication Monitoring Program Advisory Board. *J Clin Epidemiol* 1998;**51**:1335-42.

2. Iversen L, Hannaford PC, Price DB, Godden DJ. Is living in a rural area good for your respiratory health? *Chest* 2005;**128**:2059-67.

3. de Marco R, Cerveri I, Bugiani M, Ferrari M, Verlato G. An undetected burden of asthma in Italy: the relationship between clinical and epidemiological diagnosis of asthma. *Eur Respir J* 1998;**11**:599-605.

4. Global Initiative for Chronic Obstructive Lung Disease. *Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease*. Available at: http://www.goldcopd.com/. Accessed 18 August 2005.

5. Armstrong BK, White E, Saracci R. Validity and reliability studies. In: *Principles of exposure measurement in epidemiology*. New York: Oxford University Press; 1992. p. 78-79.

6. General Practice Administration System for Scotland. Available at: http://www.show.scot.nhs.uk/gpass. Accessed 15 February 2005.

7. Altman DG, Machin D, Bryant TN, Gardner MJ editors. *Statistics with confidence*. 2nd edition. Bristol: BMJ Books; 2000.

8. Altman DG. Measuring agreement. In: *Practical statistics for medical research*. London: Chapman & Hall; 1991. p. 403-07.

9. Jordan K, Jinks C, Croft P. Health care utilization: measurement using primary care records and patient recall both showed bias. *J Clin Epidemiol* 2006:**59**: 791-7.

10. Frank P, Ferry S, Moorhead T, Hannaford P. Use of a postal questionnaire to estimate the likely under-diagnosis of asthma-like illness in adults. *Br J Gen Pract* 1996;**46**:295-7.

11. Verbrugge LM Ascione FJ. Exploring the iceberg. Common symptoms and how people care for them. *Med Care* 1987;**25**:539-9.

12. Pekkanen J, Pearce N. Defining asthma in epidemiological studies. *Eur Respir J* 1999;**14**:951-7.

13. Fishman AP. One hundred years of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005;**171**:941-8.

14. Barr RG, Herbstman J, Speizer FE, Camargo CA,Jr. Validation of self-reported chronic obstructive pulmonary disease in a cohort study of nurses. *Am J Epidemiol* 2002;**155**:965-71.

15. Mohangoo AD, van der Linden MW, Schellevis FG, Raat H. Prevalence estimates of asthma or COPD from a health interview survey and from general practitioner registration: what's the difference? *Eur J Public Health* 2006;**16**: 101-05.

16. Ungar WJ, Coyte PC. Health services utilization reporting in respiratory patients. Pharmacy Medication Monitoring Program Advisory Board. *J Clin Epidemiol* 1998;**51**:1335-42.

Available online at **http://www.thepcrj.org**