A parent-completed respiratory questionnaire for one-year olds: repeatability

Marie-Pierre F. Strippoli[1], Michael Silverman[2], Gisela Michel[1], Claudia E. Kuehni[1]

Institution and country of all co-authors

(1) Swiss Paediatric Respiratory Research Group, Dept. of Social and Preventive Medicine, University of Berne, Berne, Switzerland

(2) The Leicester Children's Asthma Centre, Division of Child Health, Department of Infection, Immunity & Inflammation, University of Leicester, Leicester, UK

Corresponding author

Dr. Claudia E. Kuehni, MD MSc, Dept. of Social and Preventive Medicine

Finkenhubelweg 11, CH-3012 Berne, Switzerland

Phone: +41 (0)31 631 35 07, Fax: +41 (0)31 631 35 20, e-mail: kuehni@ispm.unibe.ch

Keywords

Preschool; Asthma; Kappa; Repeatability; Questionnaire

Word count 2195

ABSTRACT

Background and aims: There are few standardised questionnaires for the assessment of respiratory symptoms in preschool children. We have developed and tested the short-term repeatability of a postal questionnaire on respiratory symptoms for one-year olds.

Methods: A newly developed postal questionnaire for the assessment of wheeze and other respiratory symptoms was sent to parents of a population-based random sample of 4300 children aged 12 to 24 months. After an interval of 3 months, a random sample of 800 respondents received the questionnaire a second time. The responses were compared using Cohen's kappa (κ) to assess agreement corrected for chance.

Results: The first questionnaire was returned by 3194 (74%) families, the second one by 460/800 (58%). Repeatability was excellent (κ 0.80-0.96) for questions on household characteristics, environmental exposures and family history, good (κ 0.61-0.80) for questions on prevalence, severity and treatment of wheeze, and moderate (κ 0.39-0.66) for chronic cough and upper respiratory symptoms.

Conclusions: This short postal questionnaire designed for use in population-based studies has excellent repeatability for family and household characteristics and good repeatability for questions on wheeze. Short-term changes in symptom status might be responsible for variable answers on recent chronic cough and upper respiratory symptoms. Overall, the questionnaire is a valuable instrument for community-based research on respiratory symptoms in one to two-year old children.

INTRODUCTION

In epidemiological surveys the prevalence of asthma and other wheezing disorders is usually assessed by questionnaires. Because objective tests are difficult to perform, especially in young children, and their sensitivity and specificity for the diagnosis of asthma are questionable,[1] the design of appropriate questionnaires is crucial. Basically, a new instrument needs to demonstrate repeatability and validity. This paper focuses on test-retest repeatability, a measure of the consistency of the performance of a questionnaire when used under similar circumstances.

Standard respiratory questionnaires for schoolchildren have shown good or satisfactory repeatability,[2-6] but may not be appropriate for infants and preschool children, where symptoms may differ due to developmental changes and rely exclusively on proxy reports. For instance, limitation of speech with severe wheeze and exercise-induced symptoms cannot be assessed in infants unable to talk and run. Moreover, different clinical phenotypes are thought to co-exist within the large group of preschool children suffering from wheeze.[7-9] To distinguish these phenotypes, symptoms must be assessed in detail. Therefore standard questionnaires developed for schoolchildren cannot be applied automatically for very young children. Few targeted questionnaires exist. Their repeatability has been tested in small numbers of children recruited in neonatal units or outpatient clinics [10] or excluding ethnic minority groups.[11] They are therefore not necessarily generalisable.

We developed a new postal questionnaire for use in a large population-based cohort study of respiratory symptoms in preschool children in Leicestershire.[12-16] This paper describes the short-term repeatability of this questionnaire when it was reapplied to parents of one-year old children after a three-month interval.

METHODS

Setting and study population

In 1998, we used the Leicestershire Child Health Database, which includes the birth notification with mother's self-reported ethnic origin, religion, country of birth and language, to select a random sample of 3500 white (mother self-identified as British Isles or European) and 800 south Asian (mother self-identified as Indian, Pakistani or Bangladeshi) children aged 1 year (i.e. in their second year of life at recruitment) with complete birth records. South Asians, the largest ethnic minority group, accounted for 14% of one-year old children resident in the county at that time.

Questionnaire and mailings

We developed a short four-page questionnaire that could be posted to families with young children, for self-administration by parents (questionnaire available at: http://adc.bmjjournals.com/supplemental). It contained sections on upper and lower respiratory symptoms and diagnoses, healthcare utilisation and treatments for wheeze, environmental exposures including indoor air pollutants, breastfeeding, pets, nursery care, number of siblings and other household members, parental history of atopic diseases and ethnicity, language and social conditions. Some questions were derived from an earlier questionnaire used locally [11] or from the International Study of Asthma and Allergy in Childhood (ISAAC).[6] Others were newly developed. The sources of the different questions are summarised in the **online Table 1** (see: http://adc.bmjjournals.com/supplemental). All questionnaires were printed in English, with an accompanying letter translated into the four main local south Asian languages offering translation services. We sent this questionnaire to all 4300 families, with a reminder letter to non-responders six weeks later (first mailing in April, reminder in May, baseline survey). Within 3 months, 3194 (74%) questionnaires had been returned. To a random sample of responders stratified by ethnic group (600 whites and 200 south Asians) an identical questionnaire with an explanatory letter was then sent exactly three months later (in July, repeat survey).

Data analyses

Statistical analyses were performed with Stata, version 8.2 (Stata corporation, Austin, Texas). Symptom prevalence at baseline was calculated separately for participants and non-participants of the repeat survey, and comparison between these groups was assessed by Chi square tests (dichotomous variables) and tests for trend (categorical variables). To assess repeatability, Cohen's kappa coefficients were calculated.[17] Kappa compares the observed agreement between two assessments made on two different occasions, with the agreement that would be expected simply by chance. Because the kappa coefficient is sensitive to the population prevalence of responses and also to asymmetrical imbalance in marginal totals, we also present the percentage of observed total agreement (number of positive and negative answers to both questionnaires divided by the total), and the separate proportions of positive and negative agreement (number of answers in positive agreement divided by the average number of positive answers; number of answers in negative agreement divided by the average number of negative answers). These values help in understanding individual results.[18, 19] For ordinal and quantitative variables, agreement was assessed as Intraclass Correlation Coefficients (ICC). Landis and Koch [20] have suggested that kappa coefficients ≤0.4 indicate poor agreement, values of 0.41-0.60 moderate agreement, 0.61-0.80 good agreement and >0.8 excellent agreement. To examine which factors might influence agreement, we stratified analyses by household language (English, other), ethnicity (white, south Asian), the Townsend local area deprivation score (low, medium, high), interval between baseline and repeat survey (<3 months, ≥3 months), domicile (urban, rural), and respondent (same respondent, different respondent to repeat questionnaire, and mother vs. father). The equality of subgroup agreement measurements was tested using the methods of Donner *et al*.[21, 22] Missing values were coded as "no" because sensitivity analyses showed no difference in results.

Ethic committee approval

The Leicestershire Health Authority Research Ethics Committee approved the study.

RESULTS

Response rates

The response rates were 74% (3194/4300) in the baseline survey and 58% (460/800) in the repeat survey, and higher for white (n=368/600, 61%) than for south Asian children (n=92/200, 46%, p<0.001). Most baseline questionnaires were completed in April or May (n=2691/3194, 84%), repeat questionnaires were mostly completed in July and August (n=451/460, 98%). Median (interquartile range) age of the children was 17.7 (14.8-20.7) months at baseline survey and 20.5 (17.7-23.2) months at the repeat survey. For test-retest analyses we included the 413 children with the same respondent (mother or father) in both occasions, as conventional.[2] However, results were very similar in a sensitivity analysis including all 460 questionnaires.

Prevalence of respiratory symptoms

The 413 study participants who replied to the repeat questionnaire had a lower prevalence of wheeze, shortness of breath and cough in the baseline survey than the children who failed to respond (**Table 1**). Among 413 participants in the repeat survey, the overall prevalence of respiratory symptoms was similar on the two occasions.

Table 1 - Prevalence of respiratory symptoms, in respondents and non-respondents to the repeatability study.

	Non-respondents (n=340)			Respondents (n=413)						
	Baseline survey			Baseline survey			Repeat survey			p
	n	%	(95 CI)	n	%	(95 CI)	n	%	(95 CI)	
Wheeze										
Wheeze ever	138	40.6	(35.3-45.8)	116	28.1	(23.7-32.4)	120	29.1	(24.7-33.5)	<0.001
Wheeze last 12 months	121	35.6	(30.5-40.7)	110	26.6	(22.3-30.9)	105	25.4	(21.2-29.6)	0.008
Wheeze without colds	42	12.4	(8.8-15.9)	34	8.2	(5.6-10.9)	35	8.5	(5.8-11.2)	0.062
Shortness of breath	73	21.5	(17.1-25.9)	57	13.8	(10.5-17.1)	60	14.5	(11.1-18.0)	0.006
Cough										
Cough without colds	150	44.1	(38.8-49.4)	130	31.5	(27.0-36.0)	138	33.4	(28.8-38.0)	<0.001
Cough at night	83	24.4	(19.8-29.0)	86	20.8	(16.9-24.8)	81	19.6	(15.8-23.5)	0.240
Diagnoses										
Diagnosis of asthma	43	12.6	(9.1-16.2)	39	9.4	(6.6-12.3)	42	10.2	(7.2-13.1)	0.160
Eczema last 12 months	122	35.9	(30.8-41.0)	148	35.8	(31.2-40.5)	153	37.1	(32.4-41.7)	0.989
Ears, nose, throat										
Chronic rhinitis	103	30.3	(25.4-35.2)	124	30.0	(25.6-34.5)	105	25.4	(21.2-29.6)	0.936
Snoring	179	52.6	(47.3-58.0)	230	55.7	(50.9-60.5)	215	52.1	(47.2-56.9)	0.404
Chronic otitis	146	42.9	(37.7-48.2)	157	38.0	(33.3-42.7)	160	38.7	(34.0-43.5)	0.170

p: significance of difference in baseline symptom prevalence between respondents to the repeat survey and non-respondents; for the respondents there were no significant differences in symptom prevalence between the baseline and repeat surveys.

Repeatability

The repeatability of the most important symptoms are shown in **Table 2** (see also **Online Tables 2 and 3** on ADC website: http://adc.bmjjournals.com/supplemental which give detailed results for all symptoms, family history and environmental exposures). In general, agreement was excellent (kappa >0.8) for family history of atopic diseases and environmental exposures (smoking, breastfeeding, pets, cooking fuel). Questions on the frequency and severity of wheeze, asthma diagnosis, inhaler treatment and healthcare utilisation had good repeatability (kappa 0.61 - 0.80). Repeatability was moderate for questions on triggers of wheeze and cough (kappa 0.49 - 0.72). In contrast, agreement for questions on chronic cough, upper respiratory symptoms and skin problems were (with the exception of a diagnosis of eczema) only moderate (kappa 0.39 - 0.68). Stratified analyses did not show any significant differences in agreement between subgroups, with one exception: repeatability of several questions was poorer when both questionnaires had been completed by the fathers (N=39), compared to mothers (N=374). This was significant for questions on night cough (with a kappa of 0.10 vs. 0.44 for fathers and mothers respectively, p=0.003) and cough apart from colds (kappa 0.28 vs. 0.56, p=0.09), snoring (kappa 0.51 vs. 0.67, p=0.08), eczema (kappa 0.22 vs. 0.70,

p=0.003), paternal smoking (kappa 0.53 vs. 0.88, p=0.001), number of physician visits (ICC 0.55 vs. 0.76, p=0.02), posseting (ICC 0.36 vs. 0.70, p=0.002) and duration of breastfeeding (ICC 0.85 vs. 0.96, p<0.001).

Table 2 - Agreement between answers to repeat questionnaires at 3-month intervals by parents of 413 one-year olds.

	++	+-	-+	--	P_O	P_{pos}	P_{neg}	κ	(95CI)
Wheeze									
Wheeze ever*	95	21	25	272	88.9	80.5	92.2	0.73	(0.65-0.80)
Wheeze last 12 months*	83	27	22	281	88.1	77.2	92.0	0.69	(0.61-0.77)
Wheeze without cold	25	9	10	369	95.4	72.5	97.5	0.70	(0.57-0.83)
Shortness of breath	39	18	21	335	90.6	66.7	94.5	0.61	(0.50-0.72)
Cough									
Cough without cold	92	38	46	237	79.7	68.7	84.9	0.54	(0.45-0.62)
Cough at night last 12 months*	43	43	38	289	80.4	51.5	87.7	0.39	(0.28-0.50)
Diagnoses									
Diagnosis of asthma	31	8	11	363	95.4	76.5	97.4	0.74	(0.63-0.85)
Eczema last 12 months*	120	28	33	232	85.2	79.7	88.4	0.68	(0.61-0.76)
Ears, nose and throat									
Chronic rhinitis*	66	58	39	250	76.5	57.6	83.8	0.42	(0.32-0.51)
Snoring	176	54	39	144	77.5	79.1	75.6	0.55	(0.47-0.63)
Chronic otitis	125	32	35	221	83.8	78.9	86.8	0.66	(0.58-0.73)
Treatment									
Inhaled bronchodilators	43	8	15	347	94.4	78.9	96.8	0.76	(0.66-0.85)
Inhaled corticosteroids	19	1	8	385	97.8	80.9	98.8	0.80	(0.67-0.93)
Family history									
Paternal wheeze/asthma	69	12	6	326	95.6	88.5	97.3	0.86	(0.79-0.92)
Maternal wheeze/asthma	61	13	9	330	94.7	84.7	96.8	0.82	(0.74-0.89)
Household and environment									
Maternal smoking	79	5	6	323	97.3	93.5	98.3	0.92	(0.87-0.97)
Other smoking	79	5	16	313	94.9	88.3	96.8	0.85	(0.79-0.91)
Maternal smoking during pregnancy	61	2	4	346	98.5	95.3	99.1	0.95	(0.90-0.99)
Cooking with gas	299	7	8	99	96.4	97.6	93.0	0.91	(0.86-0.95)
Central heating	388	2	8	15	97.6	98.7	75.0	0.74	(0.58-0.89)
Pets	169	7	7	230	96.6	96.0	97.0	0.93	(0.90-0.97)
Attended nursery	95	20	20	278	90.3	82.6	93.3	0.76	(0.69-0.83)
Breastfed	264	3	4	142	98.3	98.7	97.6	0.96	(0.94-0.99)

++ positive on baseline and repeat questionnaires, +- positive on baseline questionnaire only, -+ positive on repeat questionnaire only, -- negative on baseline and repeat questionnaires.
P_O proportion of observed total agreement; P_{pos} proportion of observed positive agreement; P_{neg} proportion of observed negative agreement; κ kappa; 95CI 95% confidence interval of kappa.
Level of agreement indicated by kappa: ≤0.4 poor, 0.41-0.60 moderate, 0.61-0.80 good, >0.8 excellent.
* Questions from ISAAC core questionnaire [6]

DISCUSSION

This paper presents the repeatability of a short respiratory questionnaire designed for self-completion by parents of one-year old children. Repeatability was excellent for sections on family history and environmental exposures, good for questions on wheeze, asthma, treatment and healthcare utilisation over the past 12 months, and moderate for upper respiratory symptoms and cough.

One strength of our study compared to most others is the relatively large sample size, providing increased precision (reasonable confidence intervals around kappa coefficients). We had a well defined population-based study group, including British south Asians, the main local ethnic minority group in a proportion similar to that of the general population (14%). The only other published repeatability studies for preschool questionnaires of which we are aware included few and/or selected children: 72 hospital-based children [10] and <100 (exact number not quoted) children of white ethnicity.[11] The response rate in the repeat survey (58%) compares well with response rates of 47-50% in other repeatability studies (**Table 3**).

Table 3 – Published studies on repeatability (κ) of a self-reported questionnaire in preschool and schoolchildren.

Authors	Strippoli	Powell[10]	Luyt[11]	Haby[24]	Brunekreef[2]	Clifford[5]	Salome[4]
N (response rate)	413 (58%)	114 (47%)	100? (?)*	104 (50%)	410 (87%)	200? (<50%)*	111 (83%)
Age	1 yr	6-35 m	1-5 yr	3-5 yr	6-12 yr	7-11 yr	8-11 yr
Interval	3 m	2 wk	6 m	2 m	1 m	4 m	1 m
Wheeze ever	0.73	0.68	0.88	0.84	0.76		0.80
Wheeze last 12 months	0.69		0.79	0.60	0.78	0.78	
Shortness of breath	0.61	0.48			0.71	0.50	
Diagnosis of asthma	0.74		0.82	1.00	0.76		0.77
Chronic rhinitis	0.42				0.57		
Cough without a cold	0.54	0.58	0.19				
Cough at night	0.39	0.56			0.49	0.60	0.51

yr: year ; m: month; wk: week

Level of agreement indicated by kappa: ≤0.4 poor, 0.41-0.60 moderate, 0.61-0.80 good, >0.8 excellent.

* The questionnaire was mailed to this number of patients. No detailed informations about the response rate.

As in other surveys,[2] respondents in the repeatability study were less symptomatic than non-respondents. Kappa depends on the marginal observed prevalence.[18] With a fixed agreement rate, kappa is maximal for a prevalence of 0.5 and decreases if the prevalence approaches 0 or 1.[23] In our study, the prevalence of wheeze in the last 12 months in the participants of the repeatability study was 25%, with a resulting kappa of 0.69 and an agreement rate (Po) of 88%. For a prevalence of 35%, as in the total study population, kappa would have been 0.74 (assuming the same agreement rate). Thus we may have underestimated kappa values. On the other hand, respondents may be more reliable than non-respondents, leading to an overestimation of kappa. Overall, the resulting bias is likely to be small. Other characteristics of the study, such as the age of the children, time period over which the questionnaire extends, respondent to the questionnaire and the interval between repeat surveys will have a bigger impact on kappa values.[23] These differences in methodology have to be taken into account when comparing different studies (**Table 3**). For instance short intervals between the two measurements [4, 10, 24] give the parents less time to forget previous answers and the children less opportunity to change their true symptom status. Also, we need to be cautious when extrapolating our results from one-year old children to older toddlers. For instance, repeatability of the question on duration of breastfeeding is likely to be poorer if the question is asked at an older age.

It is reassuring that repeatability did not differ much between subgroups defined by language, ethnicity or social class, although power for these comparisons was limited. Where fathers responded, significantly lower repeatability was found for questions on current symptoms and infant care confirming clinical observations. In most families, mothers spend more time at home with young children than fathers and are more likely to take time off work when the children are ill. Therefore, mothers usually provide more accurate reports on children's health status. The relatively poor repeatability we found for the questionnaires completed by fathers might, however, not have a large impact for the interpretation of questionnaire surveys, because most questionnaires are usually completed by mothers.

Although the absence of objective measurements of atopy and environmental exposures such as parental smoking in epidemiological studies is often criticised, the high repeatability of parental answers gives confidence in these questionnaires. When interpreting the results for symptoms over the past 12 months, it has to be kept in mind that perfect repeatability can never be obtained because the time windows do not completely overlap. We sent the second questionnaire 3 months after the first one, so that some of the children will have developed new symptoms and others remitted.

This is not the only cause for imperfect agreement, as shown for the symptom 'wheeze ever' which, illogically, was reported in the first but not the repeat questionnaire by about 5% of parents (as also in [2]) (**Table 2**). Poorer repeatability for cough and upper respiratory symptoms compared to wheeze has also been noted in other studies. It may be partly explained by high short-term variability in these symptoms, making it likely that symptoms are reported if they have occurred recently, but not if the last episode happened many months previously. As all questionnaires were sent at the same time of the year, we could not compare repeatability between different seasons of the year. In contrast, wheeze, shortness and breath and inhaler use might be recalled more consistently because parents are more concerned. For skin problems, repeatability was significantly better for the question on diagnosis ('eczema') than for ISAAC questions on symptoms ('itchy rash') (**Table 2** and **Online Table 2**).

Finally, we want to stress that having shown a good repeatability does not allow to conclude that our questionnaire has also a good validity. While repeatability refers to the reproducibility of a measurement, validity refers to whether the questionnaire measures what it intends to measure. The best way to assess validity is to compare answers to the questionnaire with objective measurements. These might include repeated respiratory sound recordings to validate reports of wheeze and cough, comparison with hospitalisation or GP records to validate health care utilisation and inhaler use, or measurement of urinary cotinine to validate exposure to environmental tobacco smoke. We did not have the possibility to do this.

In conclusion, this short postal respiratory questionnaire developed for cross-sectional and longitudinal studies of preschool wheeze and other respiratory symptoms has a repeatability, in one-year old children, similar to standard respiratory questionnaires for older children, and could therefore be recommended for further use in community-based studies in this age-group.

ACKNOWLEDGMENTS

We thank the parents of Leicestershire children for completing the questionnaires, and Tony Davis, Business Manager, Children's Directorate, Leicester City West Primary Care Trust, for his assistance.

COMPETING INTERESTS

There are no competing interests.

FUNDING

The work presented in this paper was funded by the Swiss National Science Foundation (PROSPER grant 3233-069348 and 3200-069349, and SNF grant 823B - 046481) and the UK Department of Health (grant # 0020014). Initial data collection was supported by a research grant from Trent NHS Executive (Trent Research Scheme, RBF # 98XX3).

1. **“What is already known on this topic”**

   Standardised questionnaires on asthma and wheeze exist for schoolchildren.
   Similar instruments for preschool children are scarce despite significant age-related differences in the features of many respiratory symptoms.

2. **“What this study adds”**

   We developed a parent-completed questionnaire for assessment of respiratory symptoms in one-year old children which showed good to excellent repeatability.
   This instrument can be used in other community-based surveys of this age-group.

REFERENCES

1 Pekkanen J, Pearce N. Defining asthma in epidemiological studies. *Eur Respir J* 1999;**14**:951-57.

2 Brunekreef B, Groot B, Rijcken B*, et al*. Reproducibility of childhood respiratory symptom questions. *Eur Respir J* 1992;**5**:930-35.

3 Usherwood TP, Scrimgeour A, Barber JH. Questionnaire to measure perceived symptoms and disability in asthma. *Arch Dis Child* 1990;**65**:779-81.

4 Salome CM, Peat JK, Britton WJ*, et al*. Bronchial hyperresponsiveness in two populations of Australian schoolchildren. I. Relation to respiratory symptoms and diagnosed asthma. *Clin Allergy* 1987;**17**:271-81.

5 Clifford RD, Radford M, Howell JB*, et al*. Prevalence of respiratory symptoms among 7 and 11 year old schoolchildren and association with asthma. *Arch Dis Child* 1989;**64**:1118-25.

6 Asher MI, Keil U, Anderson HR*, et al*. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995;**8**:483-91.

7 Martinez FD, Helms PJ. Types of asthma and wheezing. *Eur Respir J* 1998;**27**:3s-8s.

8 Silverman M, Grigg J, Mc Kean M. Virus-induced wheeze in young children - A separate disease? In: Johnston S, Papadopoulos N, eds. *Respiratory infections in allergy and asthma.* New York: Marcel Dekker 2002:427-71.

9 Silverman M, Wilson N. Asthma--time for a change of name? *Arch Dis Child* 1997;**77**:62-64.

10 Powell CV, McNamara P, Solis A*, et al*. A parent completed questionnaire to describe the patterns of wheezing and other respiratory symptoms in infants and preschool children. *Arch Dis Child* 2002;**87**:376-79.

11 Luyt DK, Burton PR, Simpson H. Epidemiological study of wheeze, doctor diagnosed asthma, and cough in preschool children in Leicestershire. *BMJ* 1993;**306**:1386-90.

12 Kuehni CE, Davis A, Brooke AM*, et al*. Are all wheezing disorders in very young (preschool) children increasing in prevalence? *Lancet* 2001;**357**:1821-25.

13 Chauliac ES, Silverman M, Zwahlen M*, et al*. The therapy of pre-school wheeze: appropriate and fair? *Pediatr Pulmonol* 2006;**41**:829-38.

14 Kuehni CE, Strippoli M-PF, Silverman M. Food intolerance and wheezing in south Asian and white children: prevalence and clinical significance. *J Allergy Clin Immunol* 2006;**118**:528-30.

15 Pierse N, Rushton L, Harris RS*, et al*. Locally-generated particulate pollution and respiratory symptoms in young children. *Thorax* 2006;**61**:216-20.

16 Kuehni CE, Brooke AM, Strippoli M-PF*, et al*. Cohort profile: The Leicester Respiratory Cohorts. *Int J Epidemiol* 2007;**in press**:

17 Fleiss JL. In 2nd edn. *Statistical methods for rates and proportions.* New-York: John Wiley & Sons 1981.

18 Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990;**43**:543-49.

19 Feinstein AR, Cicchetti DV. High agreement but low Kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;**43**:551-58.

20 Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159-74.

21 Donner A, Eliasziw M, Klar N. Testing the homogeneity of kappa statistics. *Biometrics* 1996;**52**:176-83.

22 Donner A. Sample size requirements for the comparison of two or more coefficients of inter-observer agreement. *Stat Med* 1998;**17**:1157-68.

23 Vach W. The dependence of Cohen's kappa on the prevalence does not matter. *J Clin Epidemiol* 2005;**58**:655-61.

24 Haby MM, Peat JK, Marks GB*, et al*. Asthma in preschool children: prevalence and risk factors. *Thorax* 2001;**56**:589-95.