

To evaluate the reproducibility of induced sputum analysis, and to estimate the sample size required to obtain reliable results, sputum was induced by hypertonic saline inhalation in 29 asthmatic subjects on two different days. The whole sample method was used for analysis, and inflammatory cells were counted on cytospin slides. Reproducibility, expressed by intra-class correlation coefficients, was good for macrophages (+0.80), neutrophils (+0.85), and eosinophils (+0.87), but not for lymphocytes (+0.15). Detectable differences were 5.5% for macrophages, 0.6% for lymphocytes, 5.2% for neutrophils, and 3.0% for eosinophils. We conclude that analysis of induced sputum is a reproducible method to study airway inflammation in asthma. Sample sizes greater than ours give little improvement in the detectable difference of eosinophil percentages.

Key words: Induced sputum, Bronchial asthma, Airway inflammation

Induced sputum is a reproducible method to assess airway inflammation in asthma

Elena Bacci^{CA}, Silvana Cianchetti, Stefano Carnevali, Maria Laura Bartoli, Federico L. Dente, Antonella Di Franco, Daniele Giannini, Barbara Vagaggini and Pier Luigi Paggiaro

Cardiothoracic Department, University of Pisa, Italy
U.O. Fisiopatologia Respiratoria, Ospedale di
Cisanello, via Paradisa 2, 56124 Pisa, Italy

^{CA}Corresponding author:
Tel: +39 50 995365
Fax: +39 50 580124
E-mail: ebacci@katamail.com

Introduction

Ever since it has been suggested that asthma could be an inflammatory disease, many efforts have been made to study airway inflammation through both direct and indirect methods. The use of direct methods, such as bronchoalveolar lavage and bronchial biopsy, has been limited mainly by the reluctance of patients to undergo such invasive procedures, which may not be ethically justified for clinical routine.

The analysis of induced sputum has been recently introduced to study airway inflammation in asthma.¹ This method is simple, well tolerated, and can thus be easily repeated over time. As with all new methods, reproducibility is of paramount importance. A key factor affecting reproducibility is saliva contamination, which is inevitably associated with sputum collection. Indeed, variable saliva contamination may make inflammatory cell profiles difficult to recognize and thus cause poor reliability of the results. Good reproducibility data have been obtained when mucus plugs were selected from the collected sample,^{2,3} whereas slightly worse data have been obtained when the whole sample was used for analysis.^{4,5}

In the present study, we wanted to extend previous observations on reproducibility of sputum induction by also assessing the importance of saliva contamination and of sample size.

Materials and methods

Subjects

We studied 29 asthmatic subjects in a stable phase of the disease. The diagnosis of asthma was made according to internationally accepted criteria⁶ after assessing reversible airway obstruction and/or non-specific bronchial hyperresponsiveness to methacholine. Seventeen of 29 subjects were regularly treated with inhaled long-acting β_2 -agonists and/or corticosteroids according to International Guidelines (Table 1).⁶

Protocol

On two different days, separated by 1 week, the subjects underwent sputum induction with hypertonic saline. Clinical conditions and treatment were the same at the time of each evaluation. Inhaled glucocorticoids and long-acting β_2 -agonists were withdrawn 48 h before each evaluation, and short-acting β_2 -agonists were withdrawn 8 h before each evaluation. The protocol was approved by the University Ethical Committee. All subjects gave informed consent to the study procedures.

Sputum induction and bronchial responsiveness to hypertonic saline

Sputum was induced according to the method of Pin *et al.*,¹ slightly modified.⁷ Because all subjects were in a

Table 1. Clinical characteristics of asthmatic subjects

Patient	Sex	Age (years)	Asthma degree	Treatment*	FEV ₁ **	
					Test 1	Test 2
1	Female	16	Mild persistent	1000 + β 2	103	100
2	Female	44	Moderate persistent	1500 + β 2	90	80
3	Male	22	Intermittent	No	83	96
4	Female	20	Moderate persistent	500	98	104
5	Male	17	Moderate persistent	500	113	118
6	Male	35	Moderate persistent	1000 + β 2	92	83
7	Male	27	Intermittent	No	95	93
8	Female	31	Intermittent	No	87	90
9	Female	54	Intermittent	No	92	79
10	Female	40	Intermittent	No	83	80
11	Female	67	Mild persistent	500	107	105
12	Male	20	Intermittent	No	83	87
13	Female	57	Mild persistent	No	85	77
14	Male	23	Intermittent	No	109	109
15	Female	42	Intermittent	No	90	89
16	Female	31	Intermittent	No	103	106
17	Male	51	Moderate persistent	1000 + β 2	97	100
18	Male	68	Moderate persistent	1000	84	86
19	Male	45	Mild persistent	500	78	73
20	Male	31	Moderate persistent	No	85	88
21	Female	58	Severe	1500 + β 2	70	66
22	Male	64	Severe	1000 + β 2 + T	57	57
23	Female	29	Moderate persistent	1500 + β 2 + T	69	80
24	Male	32	Moderate persistent	1500 + β 2	72	81
25	Female	64	Mild persistent	No	91	82
26	Male	42	Moderate persistent	2000 + β 2 + T	86	86
27	Female	43	Moderate persistent	2000 + β 2 + T	86	113
28	Female	59	Severe	1750 + β 2 + T	48	58
29	Female	32	Moderate persistent	1000 + T	91	88

* Daily beclomethasone dipropionate dose (μ g). β 2, long-acting β 2-agonist; T: theophylline.

** FEV₁ is expressed as the percent of the predicted value.

stable phase of the disease, the β 2-agonist was administered as a pre-treatment only in subjects with Forced Expiratory Volume in 1 second (70% of predicted ($n = 4$)). Hypertonic saline solution was nebulized with an ultrasonic nebulizer (Sirius; Technomed, Florence, Italy) with a 2.8 ml/min output, and was inhaled for 5 min periods for up to 30 min. NaCl concentration was increased at intervals of 10 min from 3% to 4% to 5%. Every 5 min after the start of nebulization, subjects were asked to rinse their mouth and throat carefully, to discard saliva, and to try to cough sputum into a container; FEV₁ was then measured. Nebulization was stopped after 30 min or when the FEV₁ fell by 20% or more from baseline.

In subjects not pre-treated with β 2-agonist, bronchial reactivity to hypertonic saline inhalation was evaluated as the slope of the dose-response curve (i.e. the ratio between the maximum FEV₁ fall and the dose of delivered hypertonic saline). The dose of delivered hypertonic saline was calculated: dose = sum of (NaCl concentrations \times minutes of saline delivery).

Sputum processing

Sputum samples were processed according to the method of Fahy *et al.*,⁸ slightly modified. After

assessing the sputum volume, sputum samples from all 29 subjects were diluted with an equal volume of 0.1% dithiothreitol in phosphate-buffered saline (Sputasol; Unipath Ltd, Basingstoke, UK). Samples were incubated in a shaking bath at 37°C for 20 min, then aspirated in and out of a pipette to further dissolve mucus plugs. An aliquot (150 μ l) of the sputum sample was cytocentrifuged (Cytospin; Shandon Scientific, Sewickley, PA, USA) and stained with Diff-Quik (Baxter Scientific Products, Miami, FL, USA). Two investigators, blinded to the subject's code, each first counted at least 500 cells on each sputum slide so as to obtain the squamous cell percentage as an indicator of saliva contamination. At least 300 non-squamous cells were then counted on satisfactory slides. Cytospin slides with an amount of squamous cells such that 300 non-squamous cells could not be counted were considered unsatisfactory and discarded. Slides were then given a score ranging from 0 (cells so disrupted that cannot be recognized) to 3 (cells well recognizable, cell membranes intact, no cell clusters). A slide score ≥ 1 was considered satisfactory. All cell percentages were averaged to give the final values reported here. Macrophage, lymphocyte, neutrophil, eosinophil percentages were thus expressed as a percent of the total inflammatory cells,

Table 2. Reproducibility of inflammatory cell percentages in induced sputum of asthmatic subjects

	Macrophages	Lymphocytes	Neutrophils	Eosinophils	Squamous
Day 1 (%)	42.7 (10.3–86.7)	0.3 (0–3.4)	33.7 (11.9–87.6)	4.5 (0–66.6)	22.4 (4–79)
Day 2 (%)	41.2 (8.1–94.9)	0.6 (0–3.9)	35.3 (3.4–91.9)	6.9 (0–58)	20.9 (1–79.3)
<i>p</i>	0.79	0.60	0.56	0.11	0.87
RI	+ 0.80	+ 0.15	+ 0.85	+ 0.87	+ 0.05

Inflammatory cells are expressed as the percentage of non-squamous cells. RI, intra-class correlation coefficient.

excluding squamous cells. The remainder of the sputum sample was centrifuged at 450 × *g* for 10 min. The cell pellets were resuspended in 1 ml of phosphate-buffered saline for total cell counts with the Türk staining and cell viability assessment by Trypan blue exclusion in a hemocytometer. Samples with cell viability <70% were discarded.

Statistical analysis

Cell percentages, sputum volumes, and the slope of the dose-response curve to hypertonic saline inhalation are expressed as the median and range. FEV₁ is expressed as the mean ± standard deviation. Wilcoxon's signed rank test was used to compare sputum cell percentages and the slope of the dose-response curve to hypertonic saline inhalation obtained in the two different sputum inductions. The paired *t*-test

was used to compare the FEV₁ values measured before each sputum induction. Intra-class correlation coefficients (RI) were calculated to evaluate the concordance of sputum cell percentages and of the slope of the dose-response curve to hypertonic saline inhalation obtained in the two different sputum inductions, and RI values ≥ +0.70 were considered satisfactory.⁹ To test whether saliva contamination affects reproducibility, RI values were calculated for sub-groups selected on the basis of the highest squamous cell percentage counted on each pair of slides. Variability between two observations was expressed by means of a plot of the differences between the values of each pair of observations against the mean value of the same pair of observations.¹⁰ Sample sizes and detectable differences were evaluated according to the method used by Ward *et al.*¹¹

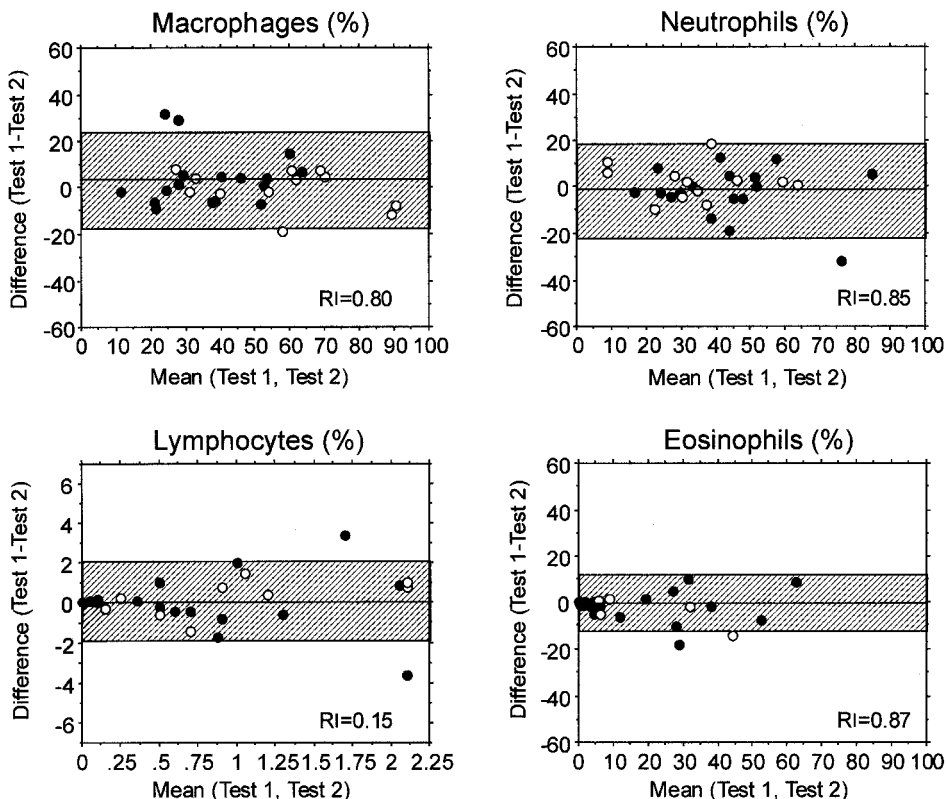


FIG. 1. Plots of differences between sputum induction procedures (test 1 – test 2) versus mean values for sputum cell percentages. Open circles, untreated patients; closed circles, treated patients. The hatched area includes the mean of differences between test 1 and test 2 ± two standard deviations of the mean.

Table 3. Intra-class correlation coefficients for sputum inflammatory cells, calculated for sub-groups selected on the basis of different levels of squamous cell contamination

Squamous cells (%)	n	Intra-class correlation coefficients			
		Macrophages	Lymphocytes	Neutrophils	Eosinophils
<80	29	+ 0.80	+ 0.15	+ 0.85	+ 0.87
<70	27	+ 0.73	+ 0.07	+ 0.83	+ 0.87
<60	26	+ 0.70	+ 0.05	+ 0.76	+ 0.87
<50	26	+ 0.70	+ 0.05	+ 0.76	+ 0.87
<40	22	+ 0.69	- 0.04	+ 0.64	+ 0.85
<30	16	+ 0.73	+ 0.05	+ 0.76	+ 0.86
<20	9	+ 0.91	- 0.05	+ 0.67	+ 0.69

n, number of patients.

Results

Cell counts

All slides were such that at least 300 cells could be counted. Slide scores were ≥ 1 and viability was $\geq 70\%$ for all samples. There were no significant differences in cell percentages from sputum samples obtained on the two different days. RI values were satisfactory for macrophages, neutrophils, and eosinophils, but not for lymphocytes (Table 2). RI values for total cell counts were low. The plots showing variability of differential cell counts are reported in Fig. 1.

Saliva contamination

Saliva contamination was variable, with squamous cell percentages reaching up to almost 80% (Table 2). However, when slides with progressively lower squamous cell contamination were selected, RI values for inflammatory cells did not increase except for macrophages, which showed an isolated higher RI value when only samples with $<20\%$ squamous cells were considered (Table 3).

Bronchial responsiveness to hypertonic saline

There was no significant difference in the slope of the dose-response curve between the two different

hypertonic saline inhalations [0.232 (0-3.133) versus 0.253 (0.025-2.733)], as well as in baseline FEV₁ ($88.6 \pm 2.5\%$ versus $88.2 \pm 2.6\%$). The RI value for the slope of the dose-response curve was satisfactory (+0.81). The FEV₁ decrease after hypertonic saline was 20% (0-47) on day 1 and 20% (2-41) on day 2. No relationship was observed between baseline FEV₁ and the FEV₁ decrease after hypertonic saline. Bronchoconstriction was promptly relieved by inhalation of a β_2 -agonist.

Sample size calculations

An estimate of the sample sizes required for a range of specified detectable differences in sputum cell percentages for paired observations was calculated. The standard deviation of the differences, required to calculate the sample sizes, was 10.6% for macrophages, 1.2% for lymphocytes, 9.9% for neutrophils, and 5.7% for eosinophils. Thus, at $p = 0.05$ and 80% power, a sample size of 29 subjects yields a detectable difference of 5.5% for macrophages, 0.6% for lymphocytes, 5.2% for neutrophils, and 3.0% for eosinophils. Figure 2 shows the plot of an estimate of the sample sizes calculated for eosinophils.

Discussion

The present study confirms that the analysis of induced sputum is a reproducible means to evaluate airway inflammation in asthma. Reproducibility was good for most cell types, with the exception of lymphocytes. Also, the inhalation of hypertonic saline to induce sputum production is a reproducible method to assess non-specific bronchial hyperresponsiveness.

The recent introduction of the analysis of induced sputum in the evaluation of airway inflammation has raised great interest because it is simple and well tolerated. However, there are some controversies about the method to be used for sputum analysis, and about its reproducibility. Several papers dealing with reproducibility have recently been published,

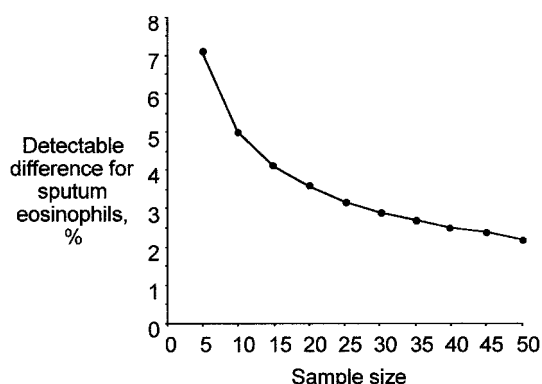


FIG. 2. Plot of detectable differences versus sample size required, based on 80% power and $p = 0.05$, for sputum eosinophil percentages.

reporting data obtained both in children¹² and in adults.²⁻⁴ The study in children dealt with reproducibility of the 'whole sample' method, and showed good sputum eosinophil reproducibility.¹² Among the studies in adults, two articles dealt with intra-subject reproducibility of the 'plugs' method^{2,3} and one article dealt with intra-subject reproducibility of the 'whole sample' method.⁴ For the 'plugs' method, both papers found good intra-class correlation coefficients for macrophage, neutrophil, and eosinophil percentages, whereas the paper on the 'whole sample' method found good results for eosinophil and lymphocyte percentages.⁴ Using the 'whole sample' method, we found good intra-class correlation coefficients for macrophage, neutrophil, and eosinophil percentages. These differences might be explained by the different amounts of saliva contamination. Our data are expressed as the median and range, but when the mean \pm standard deviation were calculated the values for squamous cell percentages in the present study (day 1, 25.1 \pm 16.2%; day 2, 23.7 \pm 16.9%) were lower than those reported by the study on the 'whole sample' method.⁴ Despite the expression of results in terms of percentages of inflammatory cells eliminating the variability due to sample dilution by saliva, the presence of a high number of squamous cells may make inflammatory cells hard to recognize on the slide. This fact confirms that low saliva contamination is essential to obtain reliable results in sputum analysis. Ward *et al.*¹³ have recently shown that excess squamous cell contamination negatively affects the accuracy of sputum differential cell counts. In the present study, the step-by-step elimination of sputum samples with progressively higher squamous cell contamination does not increase RI values for most cell types. This may be due to the fact that only seven out of 29 subjects had sputum samples with squamous cells \geq 40%, and thus their relative weight on reproducibility was low. Only macrophage reproducibility increased when samples with very low (<20%) squamous cell contamination were considered. This finding supports the hypothesis that excess squamous cell contamination makes macrophages hard to distinguish.¹³ Conversely, the low RI values obtained for the other inflammatory cells when squamous cells were <20% may be explained by the small number of samples ($n = 9$) included in statistical analysis. The good RI values obtained for inflammatory cells at the different levels of saliva contamination may also be explained by the fulfillment of criteria used for the selection of satisfactory slides.

Three more articles dealt with intrasubject sputum reproducibility using the 'whole sample' method. The paper by Gershman *et al.*¹⁴ expressed reproducibility results in terms of the variation coefficient, and cannot therefore be compared with the present results. Fahy *et al.* found sputum eosinophil reproducibility slightly below acceptable values.⁵ However,

this was a multicenter study, which might partly explain the results. For the same reason, it is difficult to compare that study with our present one. Thomas *et al.*¹⁵ found very poor sputum cell count reproducibility, since the only RI value they provide is the highest that they obtained (0.44 for lymphocytes). Thus, the only paper dealing with 'whole sample' reproducibility that can be compared with the present one is that by in't Veen *et al.*,⁴ who found slightly worse results than we did, possibly because of the greater squamous cell contamination. Since there are discrepancies on results of sputum reproducibility studies, especially for the whole sample method, ranging from very poor¹⁵ to very good, we thought that adding one further report on this issue would help assess the reliability and usefulness of induced sputum analysis.

A crucial issue for reproducibility is the time interval between inductions. It has been shown that the induction procedure itself changes the sputum composition, detectable within 24 h.^{16,17} Thus, such a short time interval between two inductions is not recommended. On the contrary, a long time interval may affect reproducibility because asthma is variable *per se*. This might partly explain the poor reproducibility found by Thomas *et al.*, who repeated sputum inductions 2 weeks apart.¹⁵ Although there is no clear demonstration of the best time interval between inductions, we thought that 1 week was good enough to reduce factors negatively affecting sputum reproducibility.

The reproducibility of total cell counts was poor in all groups, as already shown by the studies on reproducibility of the 'plugs' method. This may be due to the fact that total cell counts are expressed per volume unit. Sputum volume may vary over a wide range because of variable saliva contamination, and this is especially true for the 'whole sample' method.

All patients were studied in a stable phase of the disease, but since it has been shown that the degree of the FEV₁ decrease after hypertonic saline is significantly correlated with the baseline FEV₁ value,¹⁸ patients with FEV₁ <70% were pre-treated with inhaled β_2 -agonist. However, hypertonic saline did not cause any major adverse event in subjects who experienced bronchoconstriction. On average, the FEV₁ decrease from baseline was of a mild degree, and this was due, at least in part, to the close monitoring of pulmonary function during sputum induction and to the prompt cessation of hypertonic saline inhalation as soon as the FEV₁ value fell below acceptable limits.

The calculations of sample sizes show that sample sizes greater than ours in the present study give little improvement in the detectable differences for all cell percentages. For example, by doubling the sample size, the detectable difference would be

3.9% for macrophages, 0.4% for lymphocytes, 3.7% for neutrophils, and 2.1% for eosinophils. When planning a study design, we must keep in mind what detectable difference is reasonable for the specific results one is looking for. Since in most cases the differences observed for eosinophils (e.g. eosinophil decrease after steroid treatment, or eosinophil increase during asthma exacerbation)^{19,20} are much greater than 3%, which is the detectable difference we found in the present study, there is no need to study a larger number of subjects to obtain reliable results.

The analysis of broncho-alveolar lavage (BAL) fluid has provided important results in the study of airway inflammation in asthma, and there are currently no doubts about the reliability of such results. Data on sputum reliability are similar to data on BAL reliability. In a paper by Ward *et al.*,¹¹ the detectable difference for eosinophils in BAL fluid was about 0.5%, with a mean value of about 2% and a standard deviation of 2–2.5%, and this was considered as acceptable. Again, our data are expressed as the median and range, but when the mean \pm standard deviation were calculated the values for eosinophil percentages in the present study were $14.6 \pm 17.9\%$. That is, our detectable difference is about one-fifth of the observed mean value, whereas the detectable difference for eosinophils in BAL fluid was about one-quarter of the mean value obtained by Ward *et al.*

Although studies on reproducibility of induced sputum analysis have already been published, we believe that confirming the validity of this method is useful. In conclusion, the present study confirms that the analysis of induced sputum is a reproducible method to study airway inflammatory cells when an adequate sample size is considered. In particular, when saliva contamination is kept reasonably low, reproducibility of the 'whole sample' method is similar to that of the 'plugs' method.

Acknowledgment

This study was supported by the Italian Ministry of Research.

References

1. Pin I, Gibson PG, Kolendowicz R, *et al.* Use of induced sputum cell counts to investigate airway inflammation in asthma. *Thorax* 1992; **47**: 25–29.
2. Pizzichini E, Pizzichini MMM, Efthimiadis A, *et al.* Indices of airway inflammation in induced sputum: reproducibility and validity of cell and fluid-phase measurements. *Am J Respir Crit Care Med* 1996; **154**: 308–317.
3. Spavevello A, Migliori GB, Sharara AM, *et al.* Induced sputum to assess airway inflammation: a study of reproducibility. *Clin Exp Allergy* 1997; **27**: 1138–1144.
4. in't Veen JCCM, de Gouw HWFM, Smits HH, Sont JK, Hiemstra PS, Sterk PJ, Bel EH. Repeatability of cellular and soluble markers of inflammation in induced sputum from patients with asthma. *Eur Respir J* 1996; **9**: 2441–2447.
5. Fahy JV, Boushey HA, Lazarus SC, *et al.* Safety and reproducibility of sputum induction in asthmatic subjects in a multicenter study. *Am J Respir Crit Care Med* 2001; **163**: 1470–1475.
6. National heart, lung and blood institute. Global initiative for asthma. *National Institutes of Health publication No. 95-3659*. January 1995.
7. Bacci E, Cianchetti S, Paggiaro PL, *et al.* Comparison between hypertonic and isotonic saline-induced sputum in the evaluation of airway inflammation in subjects with moderate asthma. *Clin Exper Allergy* 1997; **26**: 1395–1400.
8. Fahy JV, Liu J, Wong H, Boushey HA. Cellular and biochemical analysis of induced sputum from asthmatic and from healthy subjects. *Am Rev Respir Dis* 1993; **147**: 1126–1131.
9. Chinn S. Repeatability and method comparison. *Thorax* 1991; **46**: 454–456.
10. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* 1986; **i**: 307–310.
11. Ward C, Gardiner PV, Booth H, Walters EH. Intrasubject variability in airway inflammation sampled by bronchoalveolar lavage in stable asthmatics. *Eur Respir J* 1995; **8**: 1866–1871.
12. Wilson NM, Bridge P, Spavevello A, Silverman M. Induced sputum in children: feasibility, repeatability, and relation of findings to asthma severity. *Thorax* 2000; **55**: 768–774.
13. Ward R, Woltmann G, Wardlaw AJ, Pavord ID. Between-observer repeatability of sputum differential cell counts. Influence of cell viability and squamous cell contamination. *Clin Exp Allergy* 1999; **29**: 248–252.
14. Gershman NH, Wong HH, Liu JT, Mahlmeister MJ, Fahy JV. Comparison of two methods of collecting induced sputum in asthmatic subjects. *Eur Respir J* 1996; **9**: 2448–2453.
15. Thomas PS, Yates DH, Barnes PJ. Sputum induction as a method of analyzing pulmonary cells: reproducibility and acceptability. *J Asthma* 1999; **36**: 335–341.
16. Holz O, Richter K, Jörres RA, Speckin P, Mücke M, Magnussen H. Changes in sputum composition between two inductions performed on consecutive days. *Thorax* 1998; **53**: 83–86.
17. Nightingale JA, Rogers DF, Barnes PJ. Effect of repeated sputum induction on cell counts in normal volunteers. *Thorax* 1998; **53**: 87–90.
18. Wong HH, Fahy JV. Safety of one method of sputum induction in asthmatic subjects. *Am J Respir Crit Care Med* 1997; **156**: 299–303.
19. Claman DM, Boushey HA, Liu J, Wong H, Fahy JV. Analysis of induced sputum to examine the effects of prednisone on airway inflammation in asthmatic subjects. *J Allergy Clin Immunol* 1994; **94**: 861–869.
20. Pizzichini MMM, Pizzichini E, Clelland L, *et al.* Sputum in severe exacerbations of asthma. Kinetics of inflammatory indices after prednisone treatment. *Am J Respir Crit Care Med* 1997; **155**: 1501–1508.

Received 2 May 2002

Accepted 2 July 2002



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

