

Screening for Asthma by Population Ranking: A Validation Study

SYBIL HIRSCH, PhD, TIMOTHY L. FRANK, MD, MICHELLE HAZELL, MBChB,
AND PETER I. FRANK, MD

PURPOSE: A neural network system was previously developed to rank a population in order of asthma probability based on responses to a postal questionnaire. Respondents ranked higher than a percentage point screening threshold are offered clinical review. The present study validates this system in a new population that had not been involved in system development.

METHODS: The system was used to rank respondents to a community survey and to predict positive predictive value (PPV) for percentage point thresholds between the top 1% and 10% of the ranking. Respondents in the top 10% were invited for clinical review. Review information and expert opinion was used to designate respondents as clinically "asthmatic" or "non-asthmatic." PPV prediction for each threshold was compared with clinical status of respondents.

RESULTS: As the threshold increased from 1% to 10%, the additional yield of clinical asthmatics decreased, indicating a ranking in asthma probability order (all 7 in the top 1% were clinical asthmatics compared with 91% of the top 5% and 83% of the top 10%). Percentages of clinical asthmatics were generally slightly higher than system PPV predictions.

CONCLUSIONS: The system ranked the population in asthma probability order and estimated PPV conservatively, enabling health-care providers to predict resource implications of a screening program.

Ann Epidemiol 2005;15:64–70. © 2004 Elsevier Inc. All rights reserved.

KEY WORDS: Asthma Screening, Expert Opinion, Neural Network, Questionnaire.

INTRODUCTION

Evidence exists that the prevalence of asthma is increasing and it is under-diagnosed and under-treated (1, 2). As a result, there may be many people in the community who need treatment but do not receive it because their condition has not been recognized. One reason may be that asthma is difficult to diagnose other than by a full clinical review, and limited health service resources mean that it is not possible to offer this to all individuals in a population. To allow health-care providers to target clinical resources to those most in need, a neural network system was previously developed to rank a population in order of probability of asthma based on their responses to a respiratory screening questionnaire (Fig. 1) (3). Using this ranking system, a percentage point in the ranked population list is chosen as a screening threshold. Respondents ranked above the threshold but with no recorded evidence of asthma diagnosis are invited for clinical assessment.

The neural network is a sophisticated statistical technique that models complex relationships between questions on the questionnaire to predict the required probability of asthma

for each respondent in the population. The probabilistic basis of the system allows positive predictive value (PPV) to be estimated in advance for any screening threshold in a given population (3, 4), enabling health-care providers to adjust the threshold to reflect resources available for diagnostic review and subsequent treatment. Clinical resources will be most effectively used if the chosen threshold has a high PPV, since this minimizes the resource wastage of reviewing those who are found not to have asthma.

The neural network ranking system was developed as part of the Wythenshawe Community Asthma Project (WYCAP), a long-term prospective investigation of asthma in two general practices (5). Four postal respiratory questionnaire surveys of all adults (16 years and older) registered to the two practices were conducted in 1993, 1995, 1999, and 2001. Questionnaires were derived from the European Community Respiratory Health Questionnaire (ECRHQ) (6) and included questions on age, gender, respiratory symptoms in the last 12 months, asthma attack in the last 12 months, current asthma medication, family history of asthma, associated conditions of hay fever or eczema, and smoking habits (7). Details of the ranking system development are described in a previous article (3). Briefly, neural network techniques (8) were used to model the relationship between questionnaire responses and probability of asthma using a subset of the 1995 survey data. Most of the response inputs were presented to the network as "0" for "No" and "1" for "Yes." There were two continuous variables, date of birth and smoking, which were

From the General Practice Research Unit, North West Lung Research Centre, Wythenshawe Hospital, Manchester, M23 9LT, U.K.

Address correspondence to: Dr. Sybil Hirsch, GP Research Unit, North West Lung Research Centre, Wythenshawe Hospital, Manchester M23 9LT, U.K. Tel.: +44-161-291-5044; Fax: +44-161-291-5047. E-mail: sybil@miga.demon.co.uk

Received August 7, 2003; accepted March 24, 2004.

Selected Abbreviations and Acronyms

ECRHQ = European Community Respiratory Health Questionnaire
PPV = positive predictive value
WYCAP = Wythenshawe Community Asthma Project

converted from continuous to categorical values resulting in four age bands (16–34, 35–54, 55–74 and ≥ 75 years respectively) and the two categories “Smoking” and “Non-smoking.” In all, the 11 questions resulted in 16 network inputs. The final neural network system comprised a “committee” of 18 representative models, most of which had 1 or 2 nodes in 1 hidden layer, and whose outputs were averaged to give the predicted probability of asthma.

This model was then applied to the questionnaires from all respondents to the 1995 survey to produce a ranking order based on probability of asthma. PPVs for various screening thresholds were estimated by calculating the cumulative proportion of those ranked higher than the threshold who were expected to be diagnosed asthmatic. These PPV estimates for the 1995 survey were based on the sample stratification and the probability of asthma of

respondents in the subset used for system development (see [Appendix](#)). Results indicated that the neural network system was successfully targeting individuals who would benefit most from clinical examination. For example, when the screening threshold was set to capture the top 5% of the ranked 1995 population, 68 individuals above this threshold had no evidence of diagnosis and would need assessment. Of these, the system predicted that 88% would have the diagnosis of asthma confirmed.

The aim of the present study was to confirm the validity of system PPV predictions by using independent clinical assessment of a sample of respondents to a new survey (2001).

METHODS

The neural network system was applied to all fully completed adult questionnaires from one of the practices in the 2001 WYCAP survey. The system estimated the probability of asthma in each respondent and the population was ranked with those individuals with higher probabilities nearer the top of the list. PPVs (expected cumulative

Please tick the appropriate box.

1. What is your date of birth?
2. Are you FEMALE ☐ MALE ☐
3. Have you had wheezing or whistling in your chest NO ☐ YES ☐
at any time **in the last 12 months?**
IF ‘NO’ GO TO QUESTION 4,
IF ‘YES’:
 - 3.1 Have you been at all breathless when the wheezing noise was present? NO ☐ YES ☐
 - 3.2 Have you had this wheezing or whistling when you did not have a cold? NO ☐ YES ☐
4. Have you woken up with a feeling of tightness in your chest **in the last 12 months?** NO ☐ YES ☐
5. Have you been woken by an attack of shortness of breath at any time **in the last 12 months?** NO ☐ YES ☐
6. Have you been woken by an attack of coughing at any time **in the last 12 months?** NO ☐ YES ☐
7. Have you had an attack of asthma in the last 12 months? NO ☐ YES ☐
8. Are you currently taking any medicine for asthma? NO ☐ YES ☐
(including inhalers, aerosols or tablets)
9. Has any person in your family (parents, grandparents, sisters, brothers, or your children) had asthma? NO ☐ YES ☐
10. Have you ever had hayfever or eczema? NO ☐ YES ☐
11. How many cigarettes do you smoke each day? per day
12. How many other adults live in your house?
How many of these adults smoke?

THANK YOU FOR YOUR HELP
PLEASE RETURN THIS FORM TO US IN THE REPLY-PAID ENVELOPE

FIGURE 1. The screening questionnaire.

proportion of asthmatics above each percentage point) were estimated for percentage points from the top 1% to the top 10% of this ranking.

All individuals ranked in the top 10% of the list and a random sample (approximately 1 in 5) of the 200 respondents at the end of the ranked list were invited for full clinical review. The purpose of including the low probability respondents was simply to provide objective evidence of the diagnostic discrimination of the experts. This review was similar to the one previously used in the WYCAP study (7). It was performed by a research physician and a practice nurse and comprised a full history, physical examination, spirometry, a test for reversibility to β -2 agonists, bronchial challenge with histamine, electronic peak flow diaries, and skin prick testing to five common allergens.

The clinical review results were sent to three respiratory experts who had no knowledge of the subjects' postal questionnaire responses. The experts were asked to give an opinion as to each subject's probability of having obstructive airways disease. For those subjects rated 50% probability or higher, the experts were asked to distinguish whether the individual had asthma, chronic obstructive airways disease (COPD) or mixed disease (both asthma and COPD). Those subjects for whom at least 2 experts diagnosed asthma or mixed disease were designated asthmatic. For each percentage point screening threshold from 1% to 10% at the top of the ranking, the PPV estimated in advance by the system was compared with the cumulative proportion of individuals designated asthmatic according to the experts' majority verdict.

In addition to these cumulative percentages of asthmatics relevant to the screening situation, observed and predicted percentages of asthmatics were calculated for individual percentage bands of width 1%.

To assess any potential response bias, the age, gender, and ranking position of those attending for review was compared with those invited but not attending and reasons for non-attendance were recorded.

RESULTS

In the 2001 survey of the study practice, 5309 individuals were sent a postal respiratory questionnaire and 3440 responded (65%). Of these, 2832 (82%) were fully completed questionnaires and used in this study. All 283 individuals from the top 10% of the ranked list were invited for review, and 118 attended. Table 1 shows the cumulative number and percentage of individuals above each of the 10 screening thresholds who were designated asthmatic according to the consensus majority verdict of the experts (observed PPV). As the screening threshold increased from 1% to 10%, the cumulative percentage of clinical asthmatics above that threshold decreased, indicating that the system was ranking in order of probability of asthma. In addition, Table 1 shows that for seven of the screening thresholds investigated, the observed percentages of asthmatic subjects were higher than those predicted in advance by the system.

In addition, 39 individuals randomly selected from a pool of 200 individuals at the end of the ranked list (low probability) were also invited for review. Nineteen attended, and of these 18 were diagnosed non-asthmatic according to the majority verdict of the experts.

Table 2 shows the observed and predicted percentages of asthmatics associated with percentage bands of width 1%. For example, for respondents ranked lower than 4% but not lower than 5%, the clinically observed percentage of asthmatics was 85% as compared with 89% predicted. In general, there was poorer correspondence between observed and predicted for these individual percentage bands than

TABLE 1. The cumulative percentage of asthmatics associated with screening thresholds from 1% to 10% and end 200 respondents in the ranked population

Top 10%	Percentage point screening threshold	Cumulative number eligible for review	Cumulative number reviewed	*Observed PPV [%] (n)	Predicted PPV [%]
	1%	28	7	100 (7)	90.65
	2%	57	21	95.24 (20)	90.57
	3%	85	30	96.67 (29)	90.28
	4%	113	43	93.02 (40)	90.22
	5%	142	56	91.07 (51)	89.99
	6%	170	69	88.41 (61)	89.47
	7%	198	80	87.50 (70)	88.91
	8%	226	92	82.61 (76)	85.74
	9%	255	110	83.64 (92)	83.51
	10%	283	118	83.05 (98)	81.75
End 200	93%–100%	200	19	5.2 (1)	0.42

*Respondents designated asthmatic according to the consensus majority verdict of the three experts.

TABLE 2. Asthmatics associated with percentage bands of width 1% in the top 10% of the ranked population

Percentage band position in ranking	Observed percentage of asthmatics (based on experts)	Predicted percentage of asthmatics (based on neural network)
≥ 0% and ≤ 1%	100	91
> 1% and ≤ 2%	93	91
> 2% and ≤ 3%	100	90
> 3% and ≤ 4%	85	90
> 4% and ≤ 5%	85	89
> 5% and ≤ 6%	77	87
> 6% and ≤ 7%	82	85
> 7% and ≤ 8%	50	64
> 8% and ≤ 9%	89	66
> 9% and ≤ 10%	75	66

the correspondence for the cumulative percentages shown in Table 1.

The frequencies of the expert diagnoses are shown in Table 3. Agreement between the experts was good, with pair-wise kappa statistics ranging from 0.63 to 0.77. Complete diagnostic agreement occurred for 88 (74.6%) of the 118 respondents in the top 10% of the ranked list. Combining the physicians' opinions into a consensus majority verdict resulted in 98 individuals (83.1%) designated asthmatic (to include mixed asthma and COPD) and 20 (16.9%) non-asthmatic.

Over half (54%) of those on the invitation list but not attending for review were because the general practitioner was not able to contact the patients to invite them. Sixteen percent of the non-attendees were unfit to attend and the remaining 30% refused for other reasons. There was no significant difference in ranking position or gender proportions between attendees (38.4% males) and non-attendees (40.1% males). However, those attending had a higher median age (50 years) than those not attending (32 years)

DISCUSSION

This study has demonstrated the value of the neural network ranking system as a screening tool. For example, examining the top 10% of a ranked community will result in the discovery of individuals with asthma. This is useful in determining individuals requiring treatment who may previously have been undiagnosed, or in identifying subjects potentially suitable for clinical trials. Alternatively, using the system to rank all individuals in a community with a recorded diagnosis of asthma and examining those ranked low probability will give information on overdiagnosis.

Traditional screening methods commonly partition populations into two groups—positives requiring full diagnostic examination and negatives assumed to be healthy. Earlier WYCAP investigations indicated that this may

TABLE 3. Classification of reviewed individuals by three experts

	*Expert classifications			
	X Y Z	Number of individuals	Percentage of individuals	Consensus opinion
Top 10%	N A A	6	5.1	Asthmatic
	A N A	2	1.7	Asthmatic
	A A N	1	0.8	Asthmatic
	A A A	77	65.3	Asthmatic
	A A C	2	1.7	Asthmatic
	A A N	1	0.8	Asthmatic
	A C A	2	1.7	Asthmatic
	C A A	7	5.9	Asthmatic
	N N N	2	1.7	Non-asthmatic
	N N A	2	1.7	Non-asthmatic
	N A N	1	0.8	Non-asthmatic
	A C N	1	0.8	Non-asthmatic
	C N N	3	1.7	Non-asthmatic
	C A C	2	1.7	Non-asthmatic
	C C A	1	0.8	Non-asthmatic
	C C C	9	7.6	Non-asthmatic
	Total from top 10%	118	100	
End 200 respondents	NNN	17	89.4	Non-asthmatic
	CNN	1	5.3	Non-asthmatic
	NAA	1	5.3	Asthmatic
	Total from end 200	19	100	

*The three experts are denoted X, Y, and Z respectively.

A: asthma or mixed asthma and COPD.

C: COPD alone.

N: no respiratory disease.

result in more individuals screening positive than the medical services can cope with in a reasonable period of time, necessitating a more sophisticated approach (7, 9, 10). To address this issue, the neural network system was developed which allowed the population to be ranked in order of probability. The threshold above which clinical review takes place can be adjusted taking into account both probability of the disease and availability of resources. In addition, once a community has been surveyed, the PPV for any proposed threshold can be predicted in advance. This allows health-care providers to estimate the clinical resource implications associated with any particular threshold and may influence the choice of threshold. The importance of this study is that it tested the accuracy of these theoretical PPV predictions in a clinical setting.

PPV of a test in a population is dependent on prevalence in that population of the condition being tested and cannot be inferred directly from a subsample unless population prevalence is known. In many conditions such as asthma there may be uncertainty associated with diagnosis and consequently there may be no reliable estimates of population prevalence. This is a particular difficulty in relation to a screening test where the ultimate aim is to find undiagnosed individuals who would clearly not be reliably represented in any prevalence estimation method. For this reason, clinical review in this study was offered to every respondent in the top 10% of the neural network ranking, avoiding the problem of extrapolating PPV from subsample to population.

It was also important that the testing was conducted in respondents from an independent survey. Commonly, the cost of collecting data prevents independent testing of data intensive modeling methods such as neural networks. The value of this study is that the test population had not been involved in any way with system development.

There is a trade off between the accuracy and refinement of the neural network model and the interpretability of simpler methods. For example, using the subset of data from the 1995 survey, the relationship between questionnaire responses and clinical asthma status was re-modeled using linear logistic regression. Model coefficients were rounded to the nearest whole number to assign a weighting to each question reflecting its relative importance in determining probability of asthma. Weights ranged from – 1 to 3. For each respondent, weights were summed according to questionnaire responses to produce an overall questionnaire score reflecting probability of asthma. Scores ranged from –1 to 13. Such a system of discrete score levels is not easily comparable with the individual probability ranking of the neural network but to compare PPV of integer score thresholds with PPV achieved in the neural network method, scores were calculated for all respondents to the 2001 survey. All respondents scoring 9 or more appeared

somewhere in the top 10% of the neural network ranking and had therefore been invited for clinical review. Seventy-two such respondents scoring 9 or more underwent clinical review with a resulting PPV of 87.5%. Thirty-two respondents scored 10 or more but PPV remained at 87.5%. Table 1 indicates that examinations of approximately equivalent numbers of respondents from the neural network ranking resulted in PPV of 88.4% and 96.7%, respectively. However, the weighted scoring method provides quantitative information as to the relative influence of each question in estimating asthma probability. For example, the question, “Have you had wheezing or whistling in your chest at any time in the last 12 months?” has the highest weighting of 3 and the weighting applicable to age reduced steadily from 3 to 0 as age increased. The neural network method, on the other hand, provides no such simple explanation for its predictions.

Each increase in the neural network screening threshold from 1% to 10% resulted in a smaller additional yield of clinically diagnosed asthmatics, confirming that the system was ranking in order of probability of asthma. In addition, for most of the screening thresholds examined, the predicted PPVs were slightly lower than those observed in the clinical reviews. This is a direct result of the probabilistic nature of the neural network model in that every asthmatic individual will contribute one to the clinical count, but the model will predict a probability which is slightly less than one. If the main concern of the health-care provider is to minimize the wasted examination of false positives, the predicted PPVs represent a conservative estimate of performance. However, when using the predicted PPVs to anticipate future resource allocation, allowance must be made for the small number of additional individuals expected to be diagnosed asthmatic and requiring treatment. The intended use of the neural network ranking system is as a first level screening tool based on responses to a simple postal questionnaire and a screening threshold cut-off value. As with any such tool, and as confirmed in Table 2, it is too blunt an instrument to associate a respondent with a particular percentage band ranking.

Overall, there was no significant difference found between attendees and non-attendees with respect to gender or ranking position, but the median age of those attending was significantly higher than those invited but not attending (median ages 50 and 32, respectively). One possible reason for this is that older people may have less work commitments. Since age is believed to be inversely associated with probability of asthma (10), this indicates that the percentages of clinical asthmatics reported in the results would be slightly lower than would be expected if the whole of the invited group had attended.

In a successful screening program it is important not only to minimize the number of false positives, but also to have adequate clinical resources for reviewing those screening

TABLE 4. Information for calculating PPV for top 1% of ranked population list

Stratum (<i>i</i>) (number of symptoms/ risk factors)	Interval (<i>j</i>) (network output)	Number of respondents	ω_{ij}	Π'_{ij}	$\omega_{ij} \times \Pi'_{ij}$
6	0.9–1.0	4	0.142857	0.9885	0.141214
6	0.8–0.9	6	0.214286	0.9885	0.211821
5	0.8–0.9	11	0.392857	0.8569	0.336639
4	0.8–0.9	5	0.178571	0.8927	0.159411
1–3	0.8–0.9	2	0.071429	0.8041	0.057436

Estimated PPV for top 1% of 2001 ranked population list: $\sum_{i=1}^s \sum_{j=1}^r \omega_{ij} \Pi'_{ij} = 0.906521$

positive and treating those ultimately diagnosed with the disease. In terms of these issues, the neural network model gives a greater degree of flexibility than conventional screening methods. By testing the neural network system on an independent data set, this study has confirmed that the system ranks a population in order of probability of asthma, the number of false positives is low for screening thresholds up to 10%, and the estimates of PPVs in this range are reliable and can be used to set the screening threshold and anticipate demand on clinical resources.

The authors thank doctors, staff, and patients at the two practices. Dr. Roseanne McNamee for help with selection of the 1995 stratified random sample. Clinicians and nurses for assistance with the clinical examinations. The expert physicians who classified the patients into diagnostic categories. Colleagues for critically reading the manuscript.

APPENDIX

Estimating PPV for Screening Thresholds 1% to 10% in the 2001 Population

PPV is the probability that the individual has the disorder given a positive screening result (ranked higher than the chosen threshold).

Using standard 2-phase survey techniques (4), in a stratified population, the population asthma prevalence (Π) is given by

$$\Pi = \sum_{i=1}^s \omega_i \Pi_i \quad (1)$$

where s is the number of population strata, ω_i is the proportion of population postal questionnaires in stratum i , and Π_i is the population prevalence of asthma in stratum i .

Considering each population stratum separately, the neural network outputs were divided into intervals of width

approximately 0.1 to give the more refined prevalence estimate

$$\Pi = \sum_{i=1}^s \sum_{j=1}^r \omega_{ij} \Pi_{ij} \quad (2)$$

where s is the number of strata, r is the number of network intervals in stratum i , ω_{ij} is the proportion of postal questionnaires in interval j in stratum i , and Π_{ij} is the prevalence of asthma in interval j of stratum i

Each Π_{ij} was estimated by Π'_{ij} , the prevalence of asthma in the corresponding interval in the 1995 stratified random sample, inferred from the known probability of asthma of the sample subjects.

Hence, PPV for any screening threshold was estimated as

$$\sum_{i=1}^s \sum_{j=1}^r \omega_{ij} \Pi'_{ij} \quad (3)$$

for all respondents ranked higher than the threshold.

Example: Estimating PPV for top 1% of the 2001 ranked population list

Number of respondents in top 1%: 28

As shown in Table 4, 4 respondents had 6 “yes” answers to the key questions and network output interval 0.9 to 1.0, 6 respondents had 6 “yes” answers to the key questions and network output interval 0.8 to 0.9, and so on.

Hence, estimated PPV for top 1% of 2001 ranked population list is

$$\sum_{i=1}^s \sum_{j=1}^r \omega_{ij} \Pi'_{ij} = 0.906521$$

REFERENCES

1. Burney PGJ. Epidemiological trends. In: Barnes PJ, Grunstein MM, Leff AR, Woolcock AJ, eds. Asthma. Lippincott Raven: Philadelphia; 1997:35–47.

2. Barbee R, Murphy S. The natural history of asthma. *J Allergy Clin Immunol*. 1998;158:352–357.
3. Hirsch S, Shapiro JL, Turega MA, Frank TL, Niven RM, Frank PI. Using a neural network to screen a population for asthma. *Ann Epidemiol*. 2001; 11(8):369–376.
4. Pickles A, Dunn G. Screening for stratification in two-phase (“two stage”) epidemiological surveys. *Stat Methods Med Res*. 1995;4:73–89.
5. Frank PI, 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;46:295–297.
6. Burney PGI, Luczynska C, Chinn S, Jarvis D. The European community respiratory health survey. *Eur Resp J*. 1994;7:954–960.
7. Frank TL, Frank PI, Cropper JA, Hazell ML, Hannaford PC, McNamee RR, et al. Identification of adults with symptoms suggestive of obstructive airway disease: Validation of a postal respiratory questionnaire. *BMC Fam Pract*. 2003;4(1):5.
8. Bishop CM. The Multi-layer Perception. In: *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press; 1995:116–163.
9. Frank TL, Frank PI, McNamee R, Wright T, Hannaford P, Morrison J, et al. Assessment of a simple scoring system applied to a screening questionnaire for asthma in children aged 5–15 years. *Eur Respir J*. 1999;14(5):1190–1197.
10. Hirsch S. A weighted scoring system to screen for asthma. *Proceedings of International Primary Care Respiratory Group (IPCRG) Amsterdam, Netherlands, June 7–10, 2002*.