

Clinical diagnosis of pneumonia, typical of experts

Olli S. Miettinen MD PhD,¹ Kenneth M. Flegel MD MSc² and Johann Steurer MD³

¹Professor, Department of Epidemiology, Biostatistics, and Occupational Health, Faculty of Medicine, McGill University, Montreal, QC, Department of Medicine, Faculty of Medicine, McGill University, Montreal, QC, Canada, and Department of Medicine, Weill Medical College, Cornell University, New York, NY, USA

²Professor, Department of Medicine, Faculty of Medicine, McGill University, Montreal, QC, Canada

³Professor, Horten Centre for Practice-oriented Research and Knowledge Transfer, Department of Internal Medicine, University of Zurich, Zurich, Switzerland

Keywords

diagnosis, expert, pneumonia, probability, software

Correspondence

Olli S. Miettinen
McGill University
1020 Pine Avenue West
Montreal, QC
Canada H3A 1A2
E-mail: olli.miettinen@mcgill.ca

Accepted for publication: 27 February 2007

doi:10.1111/j.1365-2753.2007.00873.x

Abstract

Background Clinical diagnosis of pneumonia is a concern when a patient presents with recent cough – new or worsened – together with fever as the chief complaint. Given this presentation, the doctor would benefit from having access to software that specifies, first, what diagnostic indicators experts typically use in that diagnosis; then, upon entry of those facts, what experts' typical probability of pneumonia is in such a case; and finally, how much this probability might change upon adding the facts from chest radiography.

Methods We specified a set of 36 hypothetical presentations of this type by patients 20–70 years of age, involving a comprehensive set of clinical-diagnostic indicators. Members of three separate expert panels independently set the probability of pneumonia in each of these cases, and also the range of possible post-radiography probabilities. A logistic function of the diagnostic indicators was fitted to the medians of the probabilities.

Results The median probability of pneumonia was a joint function of the patient's age and current rate of cigarette smoking; history as to the cough's duration, the fever's maximum, dyspnea (including whether on effort only) and rigors; and physical examination as to temperature, signs of upper respiratory infection, prolongation of expiration, dullness on percussion and some auscultation findings. Non-contributory were history of wheezing, pain on inspiration, type of sputum and signs of cold or influenza. This probability function, and the post-radiography functions based on the same indicators, are accessible at <http://www.evimed.ch/pneumonia>.

Interpretation The expert inputs to clinical diagnosis that were derived and made readily accessible provide for expertly clinical diagnosis of pneumonia, relevant for decisions about radiography and treatment without it.

Introduction

When a patient presents with a complaint of recent cough – new or worsened – together with fever, the doctor needs to know, first, what set of clinical facts – from history and physical examination – is to be ascertained before turning to diagnosis proper. Then, with these facts at hand, the doctor presumably first considers the possibility that at issue may be a case of pneumonia.

As the clinical set of diagnosis-relevant facts – the clinical-diagnostic profile – generally is incompletely discriminating between pneumonia's presence and absence, clinical diagnosis of this disease generally can represent only uncertain knowing about its presence/absence. The correct level of certainty about the presence of pneumonia – the *correct diagnosis* of pneumonia, that is – would coincide (numerically) with the proportion of instances of the diagnostic profile in general such that pneumonia actually is

present, the general prevalence of pneumonia conditional on the profile.

While this prevalence-determined correct diagnosis of pneumonia in the context of whatever clinical-diagnostic profile remains unknown, and while even experts' diagnoses are quite divergent (in their probabilities), any doctor pursuing clinical diagnosis of pneumonia would do well substituting *typical expert* diagnosis for what otherwise would be prone to be an excessively subjective probability.

We studied how experts' typical diagnosis of pneumonia – the probability characterizing this – now is a joint function of (a subset of) the full set of clinical-diagnostic indicators that reasonably could be considered; and we made diagnosis based on this function accessible to doctors at large via their personal computers.

Whereas the diagnostician in the face of experts' typical diagnosis of pneumonia in the case at hand may consider invoking chest

radiography, we supplemented that study by its counterparts for experts' conceptions of the corresponding maxima and minima of the post-radiography probabilities, based on the same inputs as the clinical diagnoses.

Methods

For the pursuit of a diagnosis of pneumonia, we took the prompting complaint to be that of recent – new or worsened – cough together with fever, both still present, and this presentation we considered specifically in respect to persons 20–70 years of age.

The development of the questionnaire that the software system on this prompting might present began with the two senior internists among us (KF and JS) independently, though in consultation with their respective local colleagues, coming up with their suggestions for the complete set of possible diagnostic indicators and their scales. These the third one of us (OM) translated into a first draft of the questionnaire, which the senior internists critically examined, again in consultation with their colleagues. A couple of iterations led to the questionnaire's final form. It implied 25 statistical variates for full description of the clinical-diagnostic profile of any given case. These variates are specified in Table 1.

The number of elements in the various possible profiles that can be specified by filling out the questionnaire (25, Table 1), meant that scores of hypothetical profiles ideally would have been specified for expert diagnoses. Concerned to keep the number to the bare minimum necessary, we specified only 36 cases, four duplicates furtively included. Three considerations governed the case specifications, though quite informally. One of these was the concern to cover all possible cases by the resulting clinical probability function but with accent on low-probability cases, so as to serve particularly well the aim of providing for practical rule-out diagnoses (of pneumonia) without radiography. Another, competing one was maximization of the efficiency of learning by means of maximal variability of any given one of the diagnostic indicators. The third consideration, also efficiency-oriented, was the concern to minimize collinearity among (the statistical variates representing) the different diagnostic indicators in the database. The hypothetical cases that thus were specified are documented in

Appendix 1 and examples of the narrative counterparts of these specifications in Appendix 2.

The narratives of the 36 hypothetical case profiles were presented to the members of three expert panels on pulmonary/thoracic medicine: a European panel with 12 respondents, a US panel with six respondents, and a Canadian one with four respondents. The formation and the responding members of these panels are documented in Appendix 3. The main task of each of the panel members was to set, independently in each of the 36 cases, the diagnostic probability for pneumonia. An added task was to specify the minimum and maximum of the possible probabilities if findings from chest radiography were to be added to the diagnostic profile.

The case-specific medians of the expert probabilities were used to derive a logistic function for the clinical probability of pneumonia being present, applying a General Linear Model to the logit of that probability. The independent variates were the 25 specified in Table 1 together with variates representing the logit's quadratic relations to X_1 , X_2 , X_3 and X_{16} ; the square variate corresponding to X_{14} was unuseable on account of its inadvertent, complete collinearity with X_7 . Given the paucity of data points (36 cases, inclusive of four duplicates) in relation to the number of independent variates, the 'full' model (merely additive) could not be fitted.

The post-radiography minima and maxima were addressed in the same way, with the same determinants as in the final, reduced clinical function.

Results

The medians of the responding experts' probabilities for pneumonia for each of the 36 hypothetical cases are presented in Table 2. The US median probabilities were, on the whole, somewhat lower than the European ones, by 5.0% on the average. For this mean of those case-specific differences the standard error was 1.8%. The mean of the Canadian medians was 1.7% lower than the mean of the case-specific weighted means of the European and US median probabilities (12:6 weighting), and for this mean difference the standard error was 1.5%. The case-specific weighted means combining the European and US median probabilities also are shown

Table 1 Statistical variates for comprehensive description of the clinical-diagnostic profile

Variate	Specification	Variate	Specification
1	Age (years)	14	Smoking (cigarettes day ⁻¹)
2	New/worsened cough's duration (days)	15	History of pneumonia*
3	Chronic cough*	16	Current temperature (°C)
4	Daily fever*	17	Signs of URI*
5	Maximum temperature (°C)	18	Respiratory rate (no. minute ⁻¹)
6	Dyspea*	19	Prolonged expiration*
7	Dyspea at effort only	20	Percussion dullness*
8	Wheezing*	21	Auscultation friction rub*
9	Pain on inspiration*	22	Auscultation diminished inspiration sound*
10	Rigors*	23	Auscultation abnormality breath sound*
11	Purulent sputum*	24	Auscultation focal wheezing*
12	Bloody sputum*	25	Abnormality in 20–24, if 2+, in single locus*
13	Cold/influenza signs*		

*Indicator variate (1 if feature at issue is present, 0 if absent).

URI, upper respiratory infection.

Table 2 Panel-specific median probabilities for the presence of pneumonia

Case	EU	US	Can	EU-US	All 3	Case	EU	US	Can	EU-US	All 3
1	60.0	65.0	60.0	61.7	61.4	21	38.0	42.5	42.5	39.5	40.0
2	25.0	17.5	40.0	22.5	25.7	22	32.5	25.0	37.5	30.0	31.4
3	20.0	30.0	17.5	23.3	23.3	23	10.0	5.0	7.5	8.3	8.2
4	10.0	15.0	12.0	11.7	11.7	24	30.0	12.5	15.0	24.2	22.5
5	35.0	40.0	27.5	36.7	35.0	25	50.0	27.5	20.0	42.5	38.4
6	30.0	25.0	15.0	28.3	25.9	26	35.5	17.5	15.0	29.5	26.9
7	18.0	12.5	7.5	16.2	14.6	27	20.0	22.5	17.5	20.8	20.2
8	59.5	50.0	50.0	56.3	55.2	28	28.5	20.0	27.5	25.7	26.0
9	17.5	7.5	17.5	14.2	14.8	29	50.0	70.0	55.0	56.7	56.4
10	20.0	15.0	22.5	18.3	19.1	30	65.0	55.0	55.0	61.7	60.5
11	27.5	17.5	27.5	24.2	24.8	31	65.0	70.0	75.0	66.7	68.2
12	60.0	75.0	60.0	65.0	64.1	32	26.5	17.5	20.0	23.5	22.9
13	30.0	17.5	15.0	25.8	23.9	33	40.0	30.0	30.0	36.7	35.5
14	23.5	20.0	12.5	22.3	20.5	34	25.0	10.0	17.5	20.0	19.5
15	50.0	15.0	47.5	38.3	40.0	35	28.0	20.0	20.0	25.3	24.4
16	24.0	30.0	25.0	26.0	25.8	36	27.5	12.5	15.0	22.5	21.1
17	50.0	50.0	65.0	50.0	52.7						
18	60.0	60.0	60.0	60.0	60.0		34.5	29.5	31.1	32.8	32.5
19	21.5	7.5	17.5	16.8	17.0					Means of medians	
20	29.5	35.0	50.0	31.3	34.7						

Expressed as percentages, the panel-specific probabilities for each of the 36 hypothetical cases, specified in Appendix 1, are shown. The fourth column gives the weighted means of the European and US medians (12:6 weighting), and the last column gives the counterpart of this for all three panels (12:6:4 weighting).

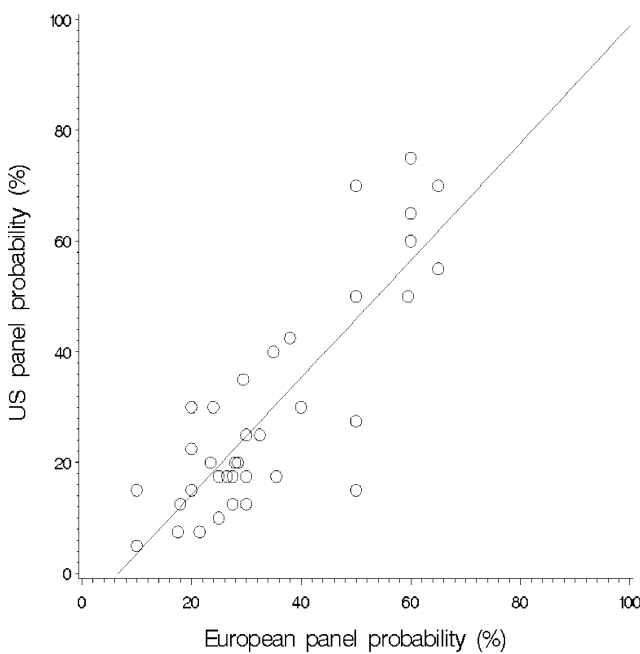


Figure 1 The median probabilities of the US panel plotted against their European counterparts. Also shown is the regression line (linear) representing the mean of the US medians as a function of the European median.

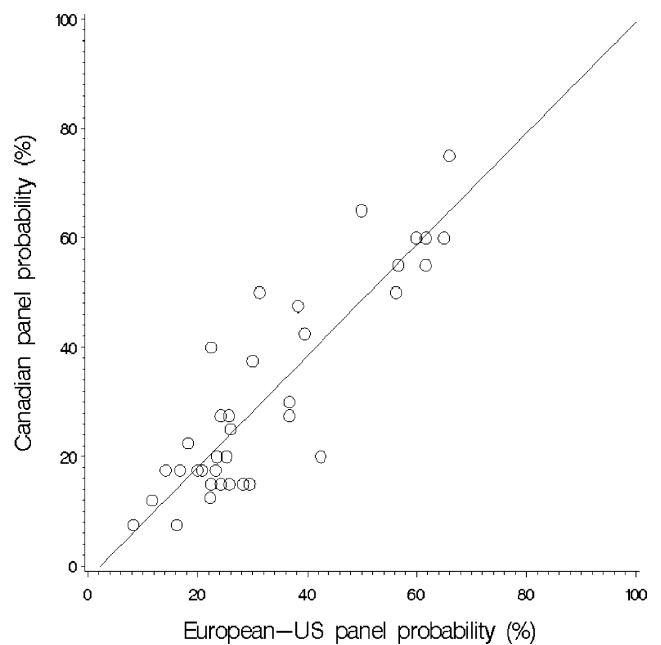


Figure 2 The median probabilities of the Canadian panel plotted against their European-US counterparts. Also shown is the regression line (linear) representing the mean of the US medians as a function of the European median.

in Table 2, and so are the counterparts of these for all three of the panels combined (12:6:4 weighting). Of special note in Table 2 is the intrapanel variability in the median probabilities for the duplicate cases – numbers 1 and 30, 2 and 26, 3 and 35, and 4 and 36 (cf. Appendix 2) – together with the convergence of these in the

medians for the three panels combined, except for the pair constituted by cases 4 and 36.

The interrelations of the medians of the probabilities set by the responding members of the three panels are addressed further by Figs 1 and 2, both based on the data in Table 2. The scattergram in

Fig. 1 relates the US medians to the European ones, while that in Fig. 2 relates the Canadian medians to those of the European and US medians combined. These patterns show no ‘regression towards the mean’: for the regression line in Fig. 1 the slope is 1.06 (SE: 0.12), and the counterpart of this in Fig. 2 is 1.02 (SE: 0.09).

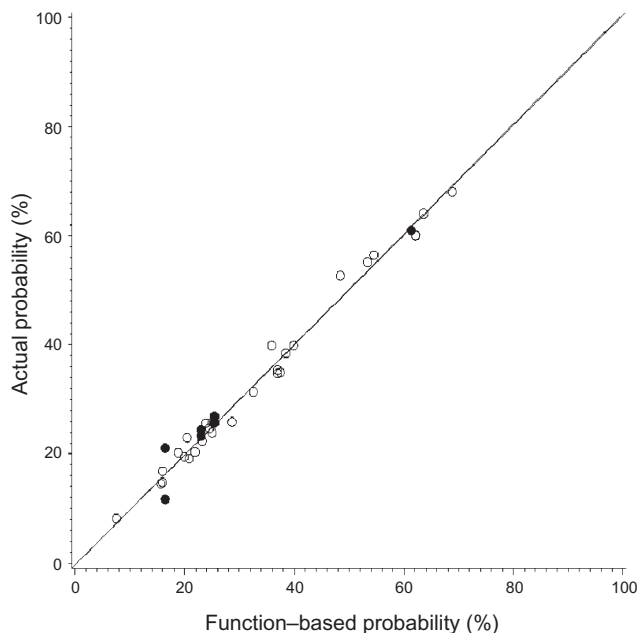


Figure 3 Goodness-of-fit of the reduced regression function. The actual probabilities are the medians from the three panels combined (shown in Table 1). The solid circles correspond to the duplicated cases (numbered 1 and 30, 2 and 28, 3 and 35, and 4 and 36). The line corresponds to identity of the two types of probability.

A bit of ‘exploratory data analysis’ led to the logistic probability (P) function, $P = 1/[1 + \exp(-S)]$, involving

$$S = -79.19 - 0.0054X_1 + 0.146X_2 + 1.890X_5 - 1.34X_6 + 2.63X_7 + 0.52X_{10} + 0.0230X_{14} + 0.144X_{16} + 1.23X_{17} - 1.40X_{19} + 0.84X_{20} + 0.97X_{21} + 0.41X_{22} + 0.52X_{23} + 0.00017(X_1 - 45)^2 - 0.00912(X_2 - 10)^2 - 0.83(X_5 - 38.5)^2 - 0.33(X_{16} - 38.5)^2$$

Each of the deleted variates, when added to those in this scoring function one at a time, changed the score value by ± 0.20 at most.

The ‘goodness-of-fit’ of that probability function, merely additive and based on a substantially reduced set of inputs, is addressed in Fig. 3, indicating how well this function characterizes the typical (median) expert diagnoses in the 36 hypothetical cases.

The extent to which findings from chest radiography might, in the extreme, change the clinical diagnosis of pneumonia (its probability) is addressed in Table 3, for a start. In it, the 36 hypothetical cases are ordered according to the 22 experts’ median probabilities for pneumonia (given in Table 2); and the table shows, for each of those cases, this probability together with the median of the minimum and of the maximum of the possible post-radiography probabilities specified by the 22 experts. The way in which these latter probabilities are joint functions of the (reduced set of) diagnostic indicators is addressed in Appendix 4.

With all of the results addressed above pertaining to the panel members’ medians for the probabilities at issue, the interexpert variability of the probabilities remains to be addressed. Among the 22 experts, the range of the expert-specific clinical probabilities for pneumonia for a given case had, across the cases, a minimum of 30% and a maximum of 90%, with a median of 70%. For the post-radiography minima the corresponding statistics were 25, 90 and 70, and for the maxima, 60, 100 and 90. The extremes in the case-specific ranges were not due to a few experts setting exceptional

Table 3 Possible impact of the findings from chest radiography on the probability of pneumonia

Case	Clinical probabilities	Post-chest x-ray		Case	Clinical probabilities	Post-chest x-ray	
		Minima	Maxima			Minima	Maxima
23	8.2	0	40	6	25.5	9	80
4	11.7	5	60	28	26.0	8	80
7	14.6	5	60	26	26.9	10	80
9	14.8	4	65	22	31.4	18	88
19	17.0	5	70	20	34.7	10	78
10	19.1	5	58	5	35.0	15	83
34	19.5	4	80	33	35.5	10	83
27	20.2	8	78	25	38.4	18	80
14	20.5	5	78	15	40.0	5	85
36	21.1	5	55	21	40.0	10	90
3	22.3	10	70	17	52.7	20	95
24	22.5	10	90	8	55.2	15	90
32	22.9	10	65	29	56.4	30	93
13	23.9	5	85	18	60.0	28	93
35	24.4	5	80	30	60.5	28	95
11	24.8	5	55	1	61.4	20	95
2	25.7	5	80	12	64.1	40	95
16	25.8	4	78	31	68.2	30	95

Corresponding to the clinical probabilities in Table 2, shown are the medians for the minima and maxima for the post-radiography probabilities specified by the 22 experts.

probabilities across the set of cases. Extreme clinical probabilities were set by 16 of the 22 experts, in one case by eight of them.

Discussion

In this Information Age, it is realistic to dream of knowledge-based, objective probability-setting in diagnosis [1,2]: the knowledge base is codified in electronic form, and the doctor approaches it by specifying the case presentation, say recent cough with fever. Back come specifications of the relevant elements in the clinical-diagnostic profile. The doctor ascertains and enters the datum on each. Upon the completion of this, the knowledge base specifies the differential-diagnostic set together with the probability/diagnosis of each of the illnesses in this set. It also specifies what laboratory test(s) might be invoked first; and if the result(s) is (are) entered, it gives the updated diagnoses and specifies the next possible test(s), etc. At the end of this process, an orderly narrative of the case, suitable for entry into the clinical record, automatically emerges.

At the core of this system would be diagnostic probability functions, either scientific or quasi-scientific [1,2]. A scientific counterpart of the function we derived for clinical diagnosis of pneumonia would be based on prevalence research in the domain of recent cough and fever in a person 20–70 years of age – on experts' collective interpretation of the results of such research [3]. The function we derived is only quasi-scientific: it is of the scientific function's form, but its empirical content derived, informally, from experts' personal experiences with diagnostic practice in that domain, not from prevalence research.

Hopstaken *et al.* [4], using the scientific approach, 'evaluated the diagnostic value of symptoms, signs, ESR [erythrocyte sedimentation rate], and CRP [C-reactive protein] for pneumonia in adult patients presenting to a GP [general practitioner] with LRTI [lower respiratory tract infection]'. They pointed out that 'classical signs and symptoms of pneumonia, derived from hospital studies, are of limited value in everyday general practice, because of the lower incidence and smaller extent of disease found there'. And consonant with this a-priori idea, their 'conclusion' from their study was that 'Most symptoms and signs traditionally associated with pneumonia are not predictive of pneumonia in general practice'.

That a-priori idea, commonly held, we regard as a misunderstanding: patients from a given diagnostic domain unquestionably are prone to have different distributions according to the (subdomain-defining) diagnostic indicators between hospital and general practices. But conditionally on a given profile, encompassing and adequately addressing all of the diagnosis-relevant reasons for the case's coming to the fore, the prevalence of pneumonia (or whatever) is the same between those two settings. A given profile thus should be translated into the same knowledge-based probability of pneumonia in whatever setting, ideally representing the correct one – corresponding to the prevalence of pneumonia in instances of this profile in general. (We use the term 'knowledge' in reference to experts' shared belief, different from philosophers 'standard, widely accepted, Platonic definition of knowledge' [5].)

Whatever is 'concluded' about the informativeness of 'symptoms and signs traditionally associated with pneumonia' thus needs to refer equally to hospital and general-practice settings; and that 'conclusion' of their uninformativeness we take to be misinterpretation of easily misleading evidence. The study domain/base was constituted by instances in which a GP had diagnosed – in

whatever manner – LRTI in a patient at least 18 years of age who presented with recent cough together with at least one symptom/sign from each of two sets of possible ones. The evidence thus had very little to do with that 'conclusion' or, otherwise put, with whether some particular item is contributory to a clinical-diagnostic probability function for pneumonia in what we took to be a reasonable domain for this.

A scientific probability function obviously would be preferable to the corresponding quasi-scientific one, and for pneumonia diagnosis it is feasible to study insofar as truth about the presence/absence of this disease is taken to be ascertainable (by imaging); but a study of it, even when feasible, is much more demanding than is the production of its quasi-scientific counterpart. The latter, in turn, is feasible only insofar as experts do exist; and the existence of expertise requires that the illness be reasonably common in some type(s) of practice, and that truth about its presence/absence tends to emerge on the basis of later experiences with instances of the diagnostic domain.

Among our experts, there was a remarkable degree of variability in the individual expert diagnoses in any given case, even if not very surprisingly [6–9]. It underscored the importance of having, in the development of the knowledge base of quasi-scientific diagnosis, expert panels of suitably large sizes, commonly panels with several dozens of members. And it implies that, in the context of a given diagnostic profile, the proper question for a non-expert to consider is not, what *an* expert's diagnosis – diagnostic probability – regarding a particular illness would be; the proper question is about a *typical* expert's diagnosis. An expert, even, should be concerned with the latter question and, hence, with the availability of the type of answer provided by work (quasi-scientific) of the sort reported here. A supreme expert diagnostician is, arguably at least, one whose case-specific diagnostic probabilities generally are typical of those that would be set by top experts.

As we were concerned not to unduly burden our panel members (unpaid), and as we thus used only three dozen cases, opportunity for sufficiently multiparameter characterization of the experts' typical diagnostic probability as a function of the diagnostic indicators involved was not guaranteed. But the resulting probability function turned out to be well descriptive of the diagnoses of the overall panel (Fig. 3), and this with a substantially reduced set of inputs. It gives, for cases not presented to the panel, probabilities as low as 0% and as high as 100% within the quantitative indicators' ranges in the cases addressed by the panel. But as the median probabilities in the cases addressed by the panel ranged from 8% to 68% (Table 3), the reliability of those more extreme probabilities is dubitable. The validity issue here is one of extrapolation of knowledge beyond what the experts addressed, not of the 'overparametrization' that in the context of descriptions of experience calls for 'shrinkage' of a multiparameter function fitted to sparse data [10]. Given an expert panel's typical probability in a particular type of case, whatever the level of this probability, there is no regression-towards-the-mean principle asserting that the corresponding probability by another, similar panel would tend to be closer to what is typical of the domain (cf. Figs 1 & 2).

We dare propose that, until something better comes along, clinical diagnosis of pneumonia is best based on the expert function presented here, perhaps applying the software at <http://www.evimed.ch/pneumonia>

Regarding the profile-specific ranges of possible post-radiography probabilities for pneumonia, we retain some uncertainty as to whether all of the panel members truly grasped what was at issue. To wit, for case number 31 the clinical probability of pneumonia was 68% (Tables 2 & 3), and the updated probability in the face of opacity-free radiographs was typically given a value as high as 30%, 10 days into the course of the sickness (Appendix 1).

Acknowledgements

We owe a debt of gratitude to the Presidents of the European, American, and Canadian societies of pulmonary/thoracic medicine, and especially to the collaborating expert-panel members nominated by them. All of these colleagues are specified in Appendix 4. Dr. Igor Karp lent his expertise on modern software for graphics and the fitting of regression functions, while Ms. Diane Rose Legault gave dedicated and competent clerical support.

References

- Miettinen, O. S. (2005) On the needs for innovation in physicians' schooling [in Finnish]. *Nuori Lääkäkäri*, 42 (16–19), 36–38.
- Miettinen, O. S. (2006) Evidence-based medicine, case-based medicine; scientific medicine, quasi-scientific medicine. *Journal of Evaluation in Clinical Practice*, 12, 260–264.
- Niiniluoto, I. (1984) The nature of science. In *Is Science Progressive?* (ed. I. Niiniluoto), pp. 1–9. Dordrecht, NL: D. Reidel Publishing Company.
- Hopstaken, R. M., Muris, J. W. M., Knottnerus, J. A., *et al.* (2003) Contributions of symptoms, signs, erythrocyte sedimentation rate, and C-reactive protein to a diagnosis of pneumonia in acute lower respiratory tract infection. *British Journal of General Practice*, 53, 358–364.
- Boghossian, P. (2006) *Fear of Knowledge. Against Relativism and Constructivism*. Oxford: Oxford University Press.
- Dolan, J. G., Bordley, D. R. & Mushlin, A. I. (1986) An evaluation of clinicians' subjective prior probability estimates. *Medical Decision Making*, 6, 216–223.
- Attia, J. R., Balakrishnan, R. N., Sibbritt, D. W., *et al.* (2004) Generating pre-test probabilities: a neglected area in clinical decision making. *Medical Journal of Australia*, 180, 449.
- Phelps, M. A. & Levitt, M. A. (2004) Pretest probability estimates: a pitfall to the clinical utility of evidence-based medicine? *Academic Emergency Medicine*, 11, 692–694.
- Puhan, M. A., Steurer, J., Bachmann, L. M. & ter Riet, G. (2004) Variability in diagnostic probability estimates [Letter]. *Annals of Internal Medicine*, 141, 578–579.
- Harrel, F. E. Jr (2001) *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer-Verlag.

Appendix 1 Clinical-diagnostic profiles in 36 hypothetical cases

The variates are specified in Table 1.

Case	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19	x20	x21	x22	x23	x24	x25
1	20	3	0	0	39	1	1	1	0	1	1	0	1	0	0	38.5	0	24	1	1	0	1	0	0	0
2	20	3	0	0	40	1	1	1	0	0	1	0	1	0	0	38.5	0	24	1	0	0	0	0	0	0
3	70	3	1	1	39	0	0	1	0	0	1	0	0	0	0	37.5	0	12	0	0	0	1	0	0	1
4	70	3	1	1	39	0	0	1	0	0	1	0	0	0	0	37.5	0	12	0	0	0	0	0	0	0
5	70	30	0	1	39	0	0	0	1	0	0	1	1	40	0	37.5	0	24	1	1	0	0	0	1	1
6	70	30	1	1	39	1	0	0	0	0	1	0	1	20	0	39.5	1	12	0	0	0	0	0	0	0
7	20	30	1	0	38	1	1	0	0	0	0	1	0	40	1	37.5	0	12	1	0	0	0	0	0	0
8	45	10	1	0	38	0	0	1	0	0	0	1	1	40	0	38.5	1	12	1	1	0	1	0	0	1
9	20	30	0	0	38	0	0	0	0	0	0	0	0	0	0	37.5	0	12	0	1	0	0	0	0	0
10	70	30	1	1	38	1	1	0	0	0	0	0	1	0	0	38.5	0	24	0	0	0	0	0	0	0
11	45	10	1	0	38	0	0	1	0	0	0	1	1	40	1	38.5	1	12	1	0	0	0	0	0	0
12	70	30	1	1	38	1	1	0	0	1	0	0	0	0	0	38.5	0	24	0	1	0	0	1	0	1
13	20	3	1	0	38	0	0	0	0	0	1	1	0	40	1	37.5	1	24	0	0	0	0	0	0	0
14	70	3	1	1	39	1	0	1	0	0	0	0	0	20	0	37.5	1	24	0	0	0	0	0	0	0
15	45	10	0	0	38	1	1	1	0	0	1	1	1	0	1	38.5	0	24	0	0	0	0	0	0	0
16	20	30	1	0	38	1	1	0	0	0	0	1	0	40	0	37.5	0	12	1	0	0	0	1	0	0
17	70	30	1	1	39	1	0	0	0	0	1	0	1	20	0	39.5	1	12	0	1	0	0	0	1	0
18	45	10	0	1	40	0	0	1	1	0	0	1	0	0	0	38.5	1	24	1	0	0	1	0	1	1
19	20	30	0	0	39	1	0	0	0	0	1	0	0	20	0	37.5	1	24	1	0	0	0	0	0	0
20	45	10	0	1	38	1	1	1	0	0	1	1	0	40	0	38.5	0	12	1	0	0	0	1	1	0
21	20	3	0	0	39	1	0	1	0	0	0	0	1	20	0	39.5	1	12	0	0	1	0	0	0	0
22	20	30	0	0	39	1	0	0	0	1	1	0	0	20	0	37.5	1	24	1	0	0	1	0	0	0
23	20	30	0	0	38	0	0	0	0	0	0	0	0	0	0	37.5	0	12	0	0	0	0	0	0	0
24	45	10	0	1	38	1	1	1	0	0	1	1	0	40	1	37.5	0	12	1	0	0	0	0	0	0
25	45	30	0	1	40	0	0	1	0	0	0	1	0	0	1	38.5	0	24	0	0	0	0	0	0	0
26	20	3	0	0	40	1	1	1	0	0	1	0	1	0	0	38.5	0	24	1	0	0	0	0	0	0
27	70	3	0	1	38	1	1	0	0	0	1	1	1	40	1	37.5	1	12	1	0	0	0	0	0	0
28	20	3	1	0	38	0	0	0	1	0	1	1	0	40	0	37.5	1	24	0	0	0	0	0	1	0
29	70	3	1	1	39	1	0	1	0	1	0	0	0	20	0	37.5	1	24	0	0	0	1	1	0	1
30	20	3	0	0	39	1	1	1	0	1	1	0	1	0	0	38.5	0	24	1	1	0	1	0	0	0
31	45	10	1	0	38	1	1	1	1	0	1	1	1	0	0	38.5	0	24	0	1	0	0	1	0	0
32	70	30	0	1	39	0	0	0	0	0	0	1	1	40	1	37.5	0	24	1	0	0	0	0	0	0
33	70	3	0	1	38	1	1	0	0	0	1	1	1	40	0	37.5	1	12	1	0	0	1	1	0	0
34	20	3	0	0	39	1	0	1	0	0	0	0	1	20	0	39.5	1	12	0	0	0	0	0	0	0
35	70	3	1	1	39	0	0	1	0	0	1	0	0	0	0	37.5	0	12	0	0	0	1	0	1	0
36	70	3	1	1	39	0	0	1	0	0	1	0	0	0	0	37.5	0	12	0	0	0	0	0	0	0

Inadvertently, specification of whether the fever was present each day (variate no. 4) was left out of the narrative for case no. 18. The value 0.5 was imputed to this missing datum. (No member of the panels remarked on this omission.)

Appendix 2 Narratives of the cases

The narratives of the hypothetical cases (Appendix 1) presented to the expert panels are exemplified by these two:

Case 1

A 20-year-old complains about cough with fever. Cough's duration, 3 days; history of chronic cough (last 6 months+), negative. Fever not daily (if untreated); maximum surface temperature (measured core t. minus 1°C), 38.5°C. New/worsened (n/w) dyspnea, yes, at effort only; n/w wheezing, yes. Pleuritic-type pain (on inspiration), no; rigors, yes. Purulent (yellow/green) sputum, yes; bloody (red/rusty spots or streaks) sputum, no. Right before episode: indications of common cold or influenza, yes; smoking (cigarettes), none. History of pneumonia (in last year), negative.

Has fever now: surface temperature, with no antipyresis, 38.5°C. Signs of upper respiratory infection, no. Respiratory rate, 24 minute⁻¹; expiration prolonged. Percussion: pneumonia-type (unibasilar) dullness, yes. Auscultation: friction rub (pleural), no; focally diminished inspiratory sound, yes; focally abnormal (bronchial) breath sound, no; focal wheezing, no. Percussion/auscultation abnormalities not in single locus.

Case 3

A 70-year-old complains about worsened cough with fever. Worsened cough's duration, 3 days; history of chronic cough (last 6 months+), positive. Fever daily (if untreated); maximum surface temperature (measured core t. minus 1°C), 38.5°C. New/worsened (n/w) dyspnea, no; n/w wheezing, yes. Pleuritic-type pain (on inspiration), no; rigors, no. Purulent (yellow/green) sputum, yes; bloody (red/rusty spots or streaks) sputum, no. Right before epi-

sode: indications of common cold or influenza, no; smoking (cigarettes), none. History of pneumonia (in last year), negative.

Has fever now: surface temperature, with no antipyresis, 37.5°C. Signs of upper respiratory infection, no. Respiratory rate, 12 minute⁻¹; expiration not prolonged. Percussion: pneumonia-type (unibasilar) dullness, no. Auscultation: friction rub (pleural), no; focally diminished inspiratory sound, yes; focally abnormal (bronchial) breath sound, no; focal wheezing, yes. Percussion/auscultation abnormalities not in single locus.

Appendix 3 The expert panels

The European panel

J. Steurer wrote to Dr. Ronald Dahl, President of the European Society for Respiratory Diseases, on our concern to have a European panel of top experts on clinical diagnosis of pneumonia. Dr. Dahl proposed that the members might best be nominated by Dr. Tom Schaberg. Dr. Schaberg agreed, specifying 19 colleagues. Two of these, unable to participate, nominated five others, for a total of 22 nominees.

J. Steurer wrote to each of these 22, attaching documents explaining what the project was about and attaching also the file of the 36 narratives of hypothetical case profiles together with the response form. The response form involved three columns. In the first column the panel member was to give his best understanding of the percentage of cases of the specified type (case no. 1, etc.) such that the patient has pneumonia. In the other two columns were to be specified the corresponding minimum and maximum of the possible diagnostic probabilities when findings from chest radiography are available as additions to the diagnostic profile.

Of the 22 nominees for the panel, 12 responded with their diagnoses for each of the 36 hypothetical cases. These colleagues were: Dr. Wim Boersma, Dr. Peter Greminger, Dr. Martin Krause, Dr. Max Kuhn, Dr. Hartmut Lode, Dr. Regina Lüthy, Dr. Erich Russi, Dr. Tom Schaberg, Dr. Martin Studnicka, Dr. Antonio Torres, Dr. John Wiggins and Dr. Mark Woodhead.

The North American panels

In terms completely analogous to those in Europe, K. Flegel approached Dr. Sharon I. S. Rounds, President of the American Thoracic Society, and Dr. P. Gerard Cox, President of the Canadian Thoracic Society, about nominees for the US and Canadian panels, respectively. They specified, respectively, 23 and 12 top experts.

Of the 23 nominees from the USA, six responded with what was requested. These colleagues were: Dr. John W. Kreit, Dr. Bimalin Lahiri, Dr. James Patterson, Dr. Alan L. Plummer, Dr. Randall R. Reeves and Dr. George L. Stewart.

Of the 12 nominees from Canada, five responded with completed data forms, but one had to be discarded on the grounds of the nature of the responses: for case no. 1 the probability was given as 50%; the next highest value was 10%, for case no. 12; and among the others, the given highest probability was 5%, for five cases, the remaining probabilities ranging from 1% to 3%. The other, useable responses came from: Dr. Don Cockcroft, Dr. Neil Colman, Dr. Peter MacLeod and Dr. Nigel Patterson.

Appendix 4 Potential of radiography

The functions for the medians of the perceived post-radiography minima and maxima of the probabilities for pneumonia were based on the data in Table 3 in conjunction with the same set of independent variates as in the clinical-diagnostic function given in the Results. For case no. 23 the value 2% was used for the minimum probability instead of 0%. The respective fitted values for the parameters are given in Table A1, and the functions' goodness-of-fit is depicted in Fig. A1.

Table A1 For the functions specifying the perceived post-radiography minima and maxima of the probability of pneumonia, the fitted coefficients for the variates involved (coefficient of $X_0 \equiv 1$ being the intercept)

Variate	Coefficient for		Variate	Coefficient for	
	Minima	Maxima		Minima	Maxima
X_0	-52.01	-40.37	X_{19}	-1.14	-1.87
X_1	0.0033	-0.0077	X_{20}	0.85	1.25
X_2	0.129	0.198	X_{21}	1.03	0.84
X_5	1.687	1.855	X_{22}	0.17	0.43
X_6	-1.31	-1.07	X_{23}	0.33	0.06
X_7	2.47	3.05	$(X_1 - 45)^2$	-3.3 (10) ⁻⁴	4.1 (10) ⁻⁴
X_{14}	0.0210	0.0343	$(X_2 - 10)^2$	-7.6 (10) ⁻³	-13.6 (10) ⁻³
X_{16}	-0.421	-0.811	$(X_5 - 38.5)^2$	-0.42	-0.27
X_{17}	1.13	1.19	$(X_{16} - 38.5)^2$	0.19	0.70

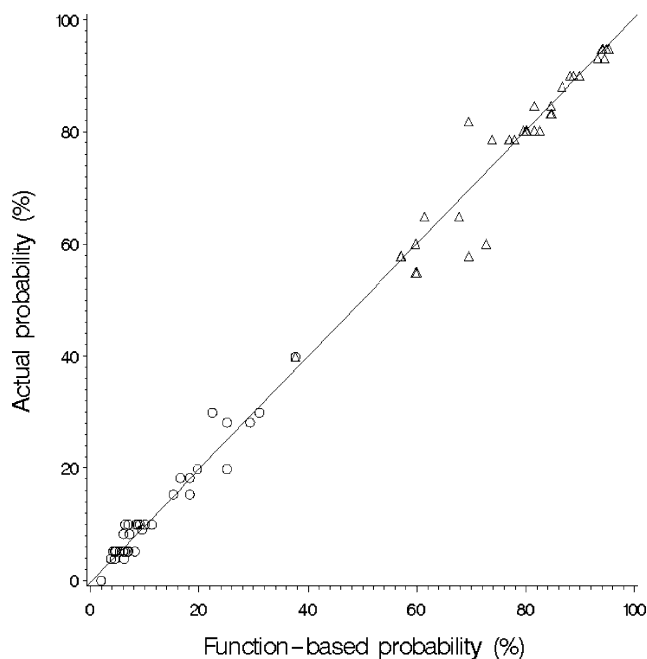


Figure A1 Goodness-of-fit of the regression functions for post-radiography minima (circles) and maxima (triangles). The actual probabilities are the medians from the three panels combined (shown in Table 3). The line corresponds to identity of the two types of probability.