Family Practitioners' Diagnostic Decision-Making Processes Regarding Patients with Respiratory Tract Infections: An Observational Study

Thomas Fischer, MD, Susanne Fischer, MD, Wolfgang Himmel, PhD, Michael M. Kochen, MD, MPH, PhD, FRCGP, Eva Hummers-Pradier, MD, PhD

Background. *The influence of patient characteristics on family practitioners' (FPs') diagnostic decision making has mainly been investigated using indirect methods such as vignettes or questionnaires. Direct observation—borrowed from social and cultural anthropology—may be an alternative method for describing FPs' real-life behavior and may help in gaining insight into how FPs diagnose respiratory tract infections, which are frequent in primary care.* ***Objective.*** *To clarify FPs' diagnostic processes when treating patients suffering from symptoms of respiratory tract infection.* ***Methods.*** *This direct observation study was performed in 30 family practices using a checklist for patient complaints, history taking, physical examination, and diagnoses. The influence of patients' symptoms and complaints on the FPs' physical examination and diagnosis was calculated by logistic regression analyses. Dummy variables based on combinations of symptoms and complaints were constructed and tested against saturated (full) and backward regression models.* ***Results.*** *In total, 273 patients (median age 37 years, 51% women) were included. The median number of symptoms described was 4 per patient, and most information was provided at the patients' own initiative. Multiple logistic regression analysis showed a strong association between patients' complaints and the physical examination. Frequent diagnoses were upper respiratory tract infection (URTI)/common cold (43%), bronchitis (26%), sinusitis (12%), and tonsillitis (11%). There were no significant statistical differences between "simple heuristic" models and saturated regression models in the diagnoses of bronchitis, sinusitis, and tonsillitis, indicating that simple heuristics are probably used by the FPs, whereas "URTI/common cold" was better explained by the full model.* ***Conclusion.*** *FPs tended to make their diagnosis based on a few patient symptoms and a limited physical examination. Simple heuristic models were almost as powerful in explaining most diagnoses as saturated models. Direct observation allowed for the study of decision making under real conditions, yielding both quantitative data and "qualitative" information about the FPs' performance. It is important for investigators to be aware of the specific disadvantages of the method (e.g., a possible observer effect).* ***Key words:*** *direct observation study; respiratory tract infection; diagnostic strategy; clinical decision making; simple heuristics.* ***(Med Decis Making 2008;28:810–818)***

Respiratory tract infections (RTIs) are the most frequent cause for patients to contact family practitioners (FPs) in Germany and generate considerable economic costs (e.g., antibiotic prescriptions, work absences, etc.).[1] Nevertheless, knowledge about FPs' cognitive processes (especially with regard to decision making) with patients suffering from symptoms of RTIs is limited and is mainly focused on the differential diagnoses between upper and lower RTIs and between bronchitis and pneumonia.[2] Overprescribing of antibiotics has been the central issue of numerous studies concerning FPs' therapy of RTIs. It was recognized that treatment decisions were influenced by the patients' and FPs' characteristics; however, the main factor is the diagnosis.[3] For example, antibiotic prescription rates tended to be higher for sinusitis and tonsillitis and lower for acute upper

Received 3 September 2004 from the Department of General Practice/Family Medicine, Georg-August University Göttingen, Göttingen, Germany (TF, SF, WH, MMK) and Hannover Medical School, Hannover, Germany. Revision accepted for publication 14 December 2007.

Address correspondence to Thomas Fischer, MD, Department of General Practice/Family Medicine, Georg-August University Göttingen, Humboldt-Allee 38, 37073 Göttingen, Germany; e-mail: tfische@gwdg.de.

DOI: 10.1177/0272989X08315254

respiratory tract infections (URTIs/common cold).[4–6] It is therefore necessary to understand which factors influence FPs' decisions about diagnoses, which in turn result in predictable therapy choices.

FPs' diagnostic strategies have already been investigated using chart reviews, questionnaires, or written case simulations (vignettes).[7–11] All these approaches gathered data indirectly, resulting in some specific disadvantages. Questionnaires (e.g., asking FPs about their procedures in hypothetical situations) and vignettes have been criticized as reflecting knowledge (and attitudes) rather than performance, whereas chart reviews tend to underestimate real performance (particularly because of documentation deficits).[12,13] In some studies, FPs were requested to fill out a documentation sheet about patients' characteristics and symptoms by themselves.[9,14,15] As in chart review studies, this design is prone to documentation deficits. Furthermore, by using vignettes, decision making can only be examined based on a preselection of analyzable items (predictors) regarded as relevant by the investigator. This implies the risk of ignoring relevant predictors that play a decisive role in real patient contacts. Alternatively, it is possible that predictors have been chosen that were not relevant under real conditions.

Therefore, our objective was to explore FPs' diagnostic strategies in patients suffering from symptoms of RTIs through the direct observation of real consultations. This concept has been borrowed from social and cultural anthropology, where it has been developed since the late 1950s: starting from an unstructured (qualitative) approach claiming to allow open, unbiased data collection, a more systematic, structured method using checklists (quantitative approach) has been developed over time, particularly in nursing research.[16–18] Because direct observation is logistically difficult to accomplish, the method is infrequently used in primary care research but provides interesting insights into real practice.[12]In our study with the data produced by the direct observation method, we analyze the influence of presenting complaints and symptoms on physical examination techniques, as well as the dependency of diagnoses on complaints, symptoms, and physical examination techniques and findings. We use multiple logistic regression to model these dependencies. In constructing these models, we use different models to represent the different theories of what it is that physicians do: make use of all available information or use a more integrative diagnostic approach[19] with reduced predictors or simple heuristics.[20,21]

METHODS

A structured observational study was performed in family practitioners' practices in a medium-sized city in Lower-Saxony and the rural areas of North Rhine-Westphalia (Germany). Of 62 FPs asked, 30 participated in this study. The FPs were informed that consultations with patients showing symptoms of RTIs would be documented using the direct observation method. To avoid biasing FPs' behavior, no specific hypotheses were shared with the participants. FPs were visited for 1 day by SF (medical student at time of observation). Because German patients regularly contact their FPs with symptoms of RTIs without an appointment, 1-day visits proved to be sufficient for acquiring a representative number of patients. All patients (older than 14 years of age) with symptoms consistent with RTIs were included. All patients were informed that a medical student would participate in the consultation.

A checklist was developed to record information on the FP-patient interaction in patients with suspected RTI (for details, see Fischer[22]). Items were based on history taking and physical examination protocols, and the checklist was evaluated and adapted in a pilot study. Data collection focused on patient complaints, physical examination findings, further diagnostic procedures, and diagnoses. Physical examination findings were coded according to the level of precision obtained: if the FP indicated that "something was wrong" without giving a precise description, "abnormal finding" was marked (lower level of precision). If more information was provided during or immediately after the consultation, this description (i.e., "coated tonsils") was then added (higher level of precision). Data were mostly acquired through silently observing and listening to the consultation. If necessary, FPs were asked to verbalize results of the physical examination immediately after the consultation. The student did neither intervene with nor actively participate in history taking or the examination; thus, only the resulting findings could be recorded but not the precise process or order of information gathering.

Respiratory tract infections were classified according to the International Classification for Primary Care (ICPC):[23] upper respiratory tract infection (URTI)/common cold (R74), sinusitis (R75), tonsillitis (R76),

laryngitis (R77), and bronchitis (R78). Because exacerbations of chronic lung diseases are not defined in the ICPC, we constructed a dummy variable (including R91, R95, and R96). Multiple diagnoses were accepted. All diagnoses were recorded as stated by the FPs directly after the consultations.

Multiple logistic regression models were used to check to what extent the decision to perform physical examinations (backward elimination with $P < 0.05$ for exclusion) depended on patients' complaints. Degree of effect of a sign or symptom upon an examination or of a sign, symptom, or finding upon a diagnosis was reported as an odds ratio (OR) with a 95% confidence interval (CI). To analyze the process of decision making (in the 4 most common diagnoses: URTI/common cold, bronchitis, sinusitis, tonsillitis), we used different logistic regression models. First, we calculated saturated (full) models with the most frequent complaints (see Table 1), abnormal findings in the physical examinations (see Table 2), age, smoking status, and duration of symptoms. Second, we performed backward analyses ($P < 0.05$ for exclusion) to represent a more integrative diagnostic approach. Third, we wanted to analyze if there were specific patterns or simple heuristics explaining the diagnoses. Therefore, 1-predictor analyses were performed to identify relevant independent predictors (inclusion criterion: Wald chi-square, $P < 0.05$). On the basis of these significant predictors, we calculated conjunction patterns (AND-combination of all significant predictors [e.g., bronchitis diagnosed if there were rales AND wheezing AND fatigue AND sputum AND no smoking]), disjunction patterns (OR-combination of all significant predictors [e.g., bronchitis diagnosed if there were rales OR wheezing OR fatigue OR sputum OR no smoking]), and several "k of n" patterns (becoming "1" if k of n predictors were 1; e.g., bronchitis diagnosis if at least 2 [or 3] of the 5 signs mentioned above were present). Thereafter, these pattern dummy variables were entered into logistic regression models. To compare the overall performance of these models, we calculated the area under the models' receiver operating characteristic curve (AUC).[24] To avoid bias in the classification accuracy for any one model, we selected the cutoff score that maximized classification accuracy for each model to identify optimal sensitivity and specificity. In addition, we calculated R^2statistics, although the explanatory power of this parameter in logistic regression models is still under debate.[25] To see if the different number of included patients per FP influenced the results, we repeated the analysis with a reduced data set based on the first 5 patients seen on the day in each practice ($n = 150$ patients).

SAS software (Version 8.2) was used to analyze the data.[26] Because SAS is not able to calculate confidence intervals for AUC, we used a specific SAS macro.[27]

Table 1 Most Common Reported Complaints (Multiple Answers Possible)

Symptoms	*n* (% of 273 Patients)
Cough	215 (78.8)
Sneezing/nasal congestion	135 (49.5)
Sore throat	95 (34.8)
Headache	73 (26.7)
Fever	69 (25.3)
Fatigue	58 (21.2)
Hoarseness	43 (15.8)
Myalgia	36 (13.2)
Earache	35 (12.8)
Facial pain	24 (8.8)
Other	118 (56.0)

RESULTS

Doctors and Patients

The FPs' median age was 48 years, their median experience in family practice was 12 years, and 17% of the FPs were women. A total of 273 patients (51% women) were included (representing 21.4% of all patients visiting their FP during the period of observation). The median number of included patients with symptoms of RTI was 9 per FP (mean 9.1; range, 5–17). Patients' median age was 37 years (mean 42.4; range, 14–88). The median duration of patients' complaints before consultation was 5 days. The most commonly reported complaints or symptoms are shown in Table 1. The median number of symptoms described was 4 per patient.

History Taking

Patients provided most information on their own. FPs' questions focused on just a few topics. The temperature was ascertained in only 68% of the 69 patients reporting fever. The method of measurement (oral, rectal, tympanic) was not ascertained. In total, 181 (84%) of the 215 patients reporting cough were asked if their cough was productive, but only 49 (23%) were queried about a circadian rhythm. Of

Table 2 Frequency of Performed Physical Examinations and Abnormal Results

Physical Examinations	*n* (% of 273 Patients)
Lung auscultation	223 (81.7)
Rhonchi	34 (12.5)
Rales	73 (26.7)
Inspection of mouth and throat	206 (75.5)
Red/inflammatory throat	180 (65.9)
Tonsils inflamed and/or coated	186 (68.1)
Paranasal sinus palpation	83 (30.4)
Tenderness	44 (16.1)
Otoscopy	73 (26.7)
Inflammatory auditory canal	6 (2.2)
Inflammatory tympanic membrane	16 (5.9)
Cervical lymph node palpation	68 (24.9)
Lymphadenopathy	44 (16.1)
Lung percussion	37 (13.6)
Abnormal findings	1 (0.4)
Vocal fremitus	10 (3.7)
Abnormal findings	0

those reporting a productive cough ($n = 102$), 73% were asked about the color of the sputum.

Physical Examination

Almost all patients were physically examined (271 of 273). The most frequent examination techniques were lung auscultation (often performed without undressing the patient) and inspection of the mouth and throat (Table 2). Most techniques lasted only a few seconds. Because physical examinations in family practice tend to be symptom oriented, we analyzed the relationship between symptoms and physical examination. A strong association between common symptoms and physical examination procedures was found (e.g., earache–otoscopy) (Table 3). In each of the multiple logistic regressions, the physical examination dependent variable was not affected by factors such as age, smoking habits, or the duration of symptoms.

Further Diagnostic Procedures

Additional diagnostic procedures were performed with some patients ($n = 32$, 11.7%), most frequently complete blood counts (18 patients). In 9 patients, pulmonary function was tested; 7 patients were referred to specialists (otorhinolaryngology, pulmonology, radiology), and 4 x-rays were ordered (2 of the chest, 2 of the paranasal sinuses). No cultures were initiated.

Table 3 Relations between Complaints and Elements of the Physical Examination

Physical Examination and Associated Complaints	Odds Ratio	95% Confidence Interval
Inspection of mouth and throat		
Sore throat	11.69	4.39–31.12
Fever	3.60	1.40–9.30
Headache	3.34	1.35–8.27
Sneezing/nasal congestion	2.36	1.22–4.58
Lung auscultation		
Cough	49.90	20.88–119.26
Fever	3.92	1.44–10.67
Sore throat	0.28[a]	0.14–0.54
Facial pain	0.25	0.10–0.64
Otoscopy		
Earache	182.07	23.30–999.99
Sneezing/nasal congestion	1.98	1.15–3.42
Fever	1.86	1.01–3.43
Paranasal sinus palpation		
Facial pain	60.54	7.88–465.42
Earache	2.58	1.13–5.87
Sneezing/nasal congestion	2.48	1.36–4.53
Cervical lymph node palpation		
Pain at neck	10.42	2.50–43.39
Earache	3.08	1.41–6.74
Fever	2.38	1.24–4.57
Fatigue	0.43	0.19–0.96
Lung percussion		
Cough	5.48	1.28–23.48

a. An odds ratio lower than 1.0 means that the physicians are less likely to perform the examination, given the complaint. For example, sore throat is an obstacle to lung auscultation. For included variables, see Methods section.

Diagnoses

For 230 (84.2%) patients, the physician assigned just 1 diagnosis. The most common diagnoses were URTI/common cold, bronchitis, tonsillitis, and sinusitis (Table 4), and 87.9% patients were labeled with at least 1 of these diagnoses.

Diagnostic Process

Logistic regression analyses (with backward elimination) clearly confirmed that diagnoses were

Table 4 Frequency of Diagnoses

Diagnoses	Number of Patients (% of 273 Patients)
URTI/common cold	117 (42.9)
Bronchitis	70 (25.6)
Sinusitis	33 (12.1)
Tonsillitis	30 (11.0)
Acute exacerbation of chronic lung diseases	24 (8.8)
Otitis media	17 (6.3)
Laryngitis	12 (4.4)
Other	16 (5.9)

Multiple diagnoses were possible. In total, 319 diagnoses were given. URTI, upper respiratory tract infection.

associated with certain symptoms and complaints as well as certain results of physical examination (Table 5). The patients' age had no influence on the diagnoses. Smoking reduced the probability of the diagnosis of bronchitis.

In addition to the backward models, we calculated logistic regression models based on combinations of those predictors previously identified as significant in 1-dimensional analyses (see Methods section and Table 5). It was not possible to construct reliable regression models based on conjunction patterns (AND-combinations) of all significant predictors because a complete set of signs and symptoms was recorded in less than 5% of the patients. Disjunction patterns (OR-combinations) of all significant predictors showed weak model performance because of the low calculated specificity in all 4 examined diagnoses. The best-fitting reduced models consisted of only a few predictors (k of n patterns; see Methods); for example, in the diagnosis of bronchitis, the most powerful model was cough in combination with any abnormal result in lung auscultation (rhonchi or rales). We could not demonstrate a significant difference in the prediction between saturated (full) and best-fitting (k of n patterns) models in the diagnoses of bronchitis, sinusitis, and tonsillitis ($P < 0.05$) (see Table 6). For the diagnosis of URTI/common cold, the saturated (full) model predicted results with significantly more accuracy.

The analysis based on the reduced data set ($n = 150$, according to the first 5 patients seen per FP) showed no statistically significant differences when compared with the results of the full data set (data not shown).

Table 5 Diagnostic Predictors of the 4 Most Common Diagnoses

Patient Complaints and Physical Examination Results Associated with the Following	Odds Ratio	95% Confidence Interval
URTI/common cold		
Sneezing/nasal congestion	3.35	1.95–5.77
Cough	2.94	1.33–6.50
Earache	2.92	1.29–6.59
Hoarseness	2.63	1.27–5.45
Rales	0.03[a]	0.01–0.10
Paranasal sinus tenderness	0.11	0.04–0.28
Wheezing	0.14	0.05–0.47
Sputum	0.33	0.18–0.61
Facial pain	0.34	0.12–0.94
Bronchitis		
Rales	12.01	5.25–27.50
Wheezing	5.69	2.12–15.27
Fatigue	3.53	1.53–8.17
Sputum	2.74	1.22–6.18
Current smoking	0.28	0.10–0.76
Tonsillitis		
Sore throat	44.53	5.63–352.03
Abnormal findings in mouth and throat inspection	13.32	1.74–102.07
Abnormal findings in cervical lymph node palpation	4.9	1.8–13.36
Sneezing/nasal congestion	0.06	0.02–0.26
Cough	0.13	0.04–0.39
Sinusitis		
Paranasal sinus tenderness	163.98	34.16–787.18
Facial pain	20.03	6.39–62.82
Headache	4.12	1.56–10.85
Cough	0.35	0.12–0.96

For included variables, see Methods section. URTI, upper respiratory tract infection.
a. An odds ratio lower than 1.0 means, for example, that rales were an obstacle to the diagnosis of URTI.

DISCUSSION

Our study showed a strong association between patients' complaints and FPs' choice of physical examination procedures. Diagnoses were associated with both typical patient complaints and findings

Table 6 Explanatory Power of Different Logistic Regression Models

Model	Bronchitis			Sinusitis			Tonsillitis			URTI		
	Full Model	Backward	Simple Heuristic[a]	Full Model	Backward	Simple Heuristic[b]	Full Model	Backward	Simple Heuristic[c]	Full Model	Backward	Simple Heuristic[d]
Variables	14	5	3	14	4	2	14	5	2	14	9	4
AUC (95% confidence interval)	0.948 (0.920–0.976)	0.935 (0.90–0.969)	0.868 (0.818–0.928)	0.982 (0.967–0.996)	0.943 (0.899–0.986)	0.964 (0.93–0.997)	0.968 (0.937–0.998)	0.943 (0.893–0.965)	0.875 (0.809–0.942)	0.893 (0.852–0.933)	0.884 (0.842–0.926)	0.752 (0.705–0.799)
Sensitivity	0.886	0.814	0.814	0.97	0.939	0.97	0.931	0.862	0.862	0.855	0.821	0.53
Specificity	0.871	0.916	0.921	0.95	0.946	0.929	0.938	0.889	0.889	0.826	0.852	0.974
Adjusted R^2	0.701	0.675	0.592	0.779	0.704	0.73	0.704	0.597	0.476	0.558	0.536	0.421

URTI, upper respiratory tract infection. AUC, Receiver operating characteristic curve.
a. Combination of cough and rhonchi or rales in lung auscultation.
b. Combination of pathologic findings in paranasal sinus palpation and facial pain.
c. Combination of sore throat and no sneezing/nasal congestion.
d. Combination of sneezing/nasal congestion and no wheezing and no rales in lung auscultation and no pathologic findings in paranasal sinus palpation.

on physical examinations. FPs tended to make their diagnosis based on a few patient symptoms and a limited physical examination: there were no significant statistical differences between "simple heuristic" models based on few cues and saturated regression models in the prediction of bronchitis, sinusitis, or tonsillitis. The full model, however, better explained URTI/common cold.

Observation Method

In this report, we demonstrate that an observation-based study is a useful approach for gaining insights on primary care patient contacts and the process of FPs decision making. As with indirect methods (such as chart reviews, vignettes, etc.), observation-based studies can reveal quantitative data (e.g., frequencies of symptoms and diagnoses). However, observation-based studies can also provide more detailed information on FP-patient interaction. Further research is needed to confirm these results, especially to compare observation-based results with indirectly gathered information. This investigation was planned as a nonparticipating observation, and the observer usually could maintain this role. It was possibly helpful that the observer was a medical student, as the FPs probably did not feel their performance was being assessed. In 1994, Pretzlik suggested the impossibility of planning the exact extent of the observer's involvement.[17] In some situations, FPs tried to involve the observer—for example, by demonstrating diagnostic findings. However, our observer spent most of the time silently recording information with a very limited proportion of time spent in discussions with the FP. According to Gold,[16] this amounts to nonparticipation.

One possible disadvantage of such a study methodology is the "observer effect" (or Hawthorne effect).[28] It seemed, however, that even major distractions are quickly forgotten by both doctor and patient.[29] Nevertheless, we cannot exclude the possibility that an FP's behavior (e.g., the high rate of performed physical examinations) was partly due to an "observer effect."

History Taking

Most patients with the symptom of cough were asked about the productivity and the character of sputum. Because productive cough (sputum) is associated with the diagnosis of bronchitis (OR = 2.74), this is possibly an effort to distinguish patients with this "complication" from those with

the diagnosis of URTI. Overall, FPs rarely asked for a detailed description of symptoms.

Physical Examination

Regression analysis confirmed that the physical examination procedures performed by the FPs depend on the patients' initial complaints. Although scientific evidence for the utility of most of these physical examination procedures is weak, FPs may have expected to improve patient satisfaction with consultations when physical examinations were performed.[30]Furthermore, this practice may represent an effort to identify reliable "red flags."[31]In this context, most studies describing physical examination in RTIs focused on discriminating uncomplicated (mostly viral) RTIs from bacterial sinusitis, otitis media, or tonsillitis.[32–36] However, the diagnostic value of physical findings is uncertain (e.g., paranasal sinus tenderness increases the likelihood of acute sinusitis) but has been shown to have low sensitivity and specificity for the diagnosis.[34] Therefore, a combination of abnormal findings and symptoms (scores) has often been recommended to increase diagnostic accuracy.[32] However, in this study, FPs apparently did not aim at collecting all information required by "sore throat" scores.[37,38] In only 1 of 30 patients subsequently given the diagnosis of tonsillitis was all the sore throat score information collected by the FPs (examination of cervical lymph nodes, throat inspection, history of fever, and cough).[38] This may result in overdiagnosing "tonsillitis" if FPs use this term as a synonym for bacterial tonsillitis or strep throat, which has been shown to result in high rates of antibiotic prescriptions in this study[39] and elsewhere.[9,40,41]

In addition to the quantitative results of just counting the examination techniques performed, watching the performance provided some additional information about the "quality" of these techniques to the observer. This "qualitative" information is an advantage of the direct observation method. For example, lung auscultation was frequently done without completely undressing the patient and lasted only a few breaths. Consequently, the observer had the impression that this procedure represented a symbolic activity. This subjective observer impression was consistent with the regression analyses showing that, for example, bronchitis was diagnosed if lung auscultation revealed abnormality, irrespective of the classification of the diagnostic findings (rhonchi or rales). Furthermore, in patients with sore throat combined with no sneezing/nasal congestion, the diagnosis of tonsillitis was made, irrespective of the result of the throat inspection.

Diagnoses

The proportion of diagnoses is comparable to previous studies involving German FPs (obtained from questionnaires).[9]A higher frequency for tonsillitis (24% v. 11% in our study) was shown in a recent Swedish FP study, which could have been due to extensive testing according to their guidelines and to the inclusion of children.[7] The latter might be responsible for the high rate of otitis media in other studies of Nordic countries, too.[42] In contrast to diagnoses made by American family physicians, age was not a contributory factor.[36] Smoking habits had an influence only on the diagnosis of bronchitis (OR = 0.28). In smokers, coughing was apparently considered "normal."

Decision Making Regarding Diagnoses

FPs' diagnostic decision-making processes have been described by several theoretical concepts—for example, by simple heuristics (or fast and frugal models), prototype models, script models (or "mindlines"), pattern recognition, or even rules of thumb.[19,20,21,43]All of these studies showed FPs' efforts to simplify and accelerate the decision-making process. This corresponds to our observations: short phases of history taking to ascertain the patient's leading symptoms (resulting in a median of only 4 symptoms per patient), rare efforts to ask for detailed descriptions, followed by symptom-oriented (and sometimes even symbolic) physical examinations. However, because we did not use "think-aloud" protocols or (structured) interviews with the FPs, to avoid interfering with the normal consultation process, our observations alone could not identify the theoretical concepts with which FPs' decisions most closely corresponded.[44]

Statistical analyses gave further evidence for a "simplified" process. Because there were only a few patients in the sample for whom all of the relevant predictors were identified, using history taking and physical examination, we expected that logistic regression models containing all relevant predictors (saturated model) would be significantly overfitted.[45]By the same token, models based on fewer predictors could be expected to include more cases and hence be more robust and yet explain a smaller proportion of the cases.[24,25,27] We could not demonstrate a relevant difference in model prediction between

saturated (based on up to 14 predictors) or "best identified" reduced models (with 2–4 predictors) in the diagnoses of bronchitis, sinusitis, and tonsillitis (see Table 6); only the diagnosis of URTI/common cold was better predicted by a saturated model.[24,25,27] Thus, the cues included in the reduced models corresponded well with "simple heuristics." For example, the "rule" to explain the diagnosis of bronchitis (which might be simplified as follows: if a patient describes the symptom of coughing, perform a lung auscultation, and if there is anything conspicuous, call it bronchitis) can be interpreted as a fast and frugal classification tree using binary features.[21] The combination of sore throat and no sneezing for the diagnosis of tonsillitis may be another simple heuristic, representing probably an effort to distinguish between a localized infection (here possibly a streptococcal tonsillitis) and a more systemic disease (probably viral caused by URTI/common cold if sneezing is present). The absence of a simple heuristic for URTI/common cold might demonstrate that the diagnosis of URTI/common cold was made by exclusion, probably if no other simple heuristic applies.

Limitations of the Study

Despite a relatively high participation rate (half of all invited FPs in this noncommercial study), we cannot exclude selection effects of the participants. An observer (Hawthorne) effect due to the presence of a student in the consulting room cannot be excluded. Furthermore, the statistical analysis based on all 273 patients (representing independent entries in the logistic model) is a simplification, which neglects the variation in patient number per FP. However, performing logistic regression analyses based on clustered data (grouped by individual FPs) did not result in significant models because of the low case number per FP (data not shown). Results of an analysis of a data set reduced to 5 patients per FP yielded results very similar to those of the full sample.

Conclusions

Methodologically, this study has shown that direct observation is a feasible method to obtain insights into FPs' diagnostic processes, yielding both quantitative data and additional qualitative information about the FP's way of working. Concerning content, FPs' diagnostic decision-making processes in patients with respiratory tract infections were predominantly based on simple cues in line with a concept of simple heuristics. As prescription of antibiotics in RTIs strongly depends on the diagnostic decision, these simple heuristics should be considered in efforts to modify FPs' prescribing behavior.

ACKNOWLEDGMENTS

We thank all participating FPs who opened their practices to S. Fischer. We also thank Stephanie Heinemann for her valuable assistance concerning the language and format in this article.

REFERENCES

1. Schnur S, Exner H. Acute respiratory tract infections in general practice [in German]. Z Allgemeinmed. 1997;73:477–9.

2. Coenen S, van Royen P, Vermeire E, Hermann I, Denekens J. Antibiotics for cough in general practice: a qualitative decisions analysis. Fam Pract. 2000;17:380–5.

3. Sorum PC, Stewart TR, Mullet E, et al. Does choosing a treatment depend on making a diagnosis? US and French physicians' decision making about acute otitis media. Med Decis Making. 2002;22:394–402.

4. Gonzales R, Steiner J, Sande M. Antibiotic prescribing for adults with colds, upper respiratory tract infections and bronchitis by ambulatory care physicians. JAMA. 1997;278:901–4.

5. Mainous AG, Hueston WJ, Clark JR. Antibiotics and upper respiratory infections: do some folks think there is a cure for the common cold? J Fam Pract. 1996;42:357–61.

6. Kuyvenhoven M, de Melker R, van der Velden K. Prescription of antibiotics and prescribers' characteristics: a study into prescription of antibiotics in upper respiratory tract infections in general practice. Fam Pract. 1993;10:366–70.

7. André M, Odenholt I, Schwan A; the Swedish Study Group on Antibiotic Use. Upper respiratory tract infections in general practice: diagnosis, antibiotic prescribing, duration of symptoms and use of diagnostic tests. Scand J Infect Dis. 2002;34:880–6.

8. de Melker RA, Kuyvenhoven MM. Management of upper respiratory tract infections in Dutch family practice. J Fam Med. 1994; 38:353–7.

9. Hummers-Pradier E, Pelz J, Himmel W, Kochen MM; participants of the 1997 Freiburg/Goettingen general practitioners' research course. Treatment of respiratory tract infections: a study in 18 general practices in Germany. Eur J Gen Pract. 1999;1:15–20.

10. Kassirer JP, Gorry GA. Clinical problem solving: a behavioral analysis. Ann Intern Med. 1978;89:245–55.

11. Gonzalez-Vallejo C, Sorum PC, Stewart TR, Chessare JB, Mumpower JL. Physicians' diagnostic judgments and treatment decisions for acute otitis media in children. Med Decis Making. 1998;18:149–62.

12. Rethans JJ, Westin S, Hays R. Methods for quality assessment in general practice. Fam Pract. 1996;13:468–76.

13. Dresselhaus TR, Peabody JW, Lee M, Wang MM, Luck J. Measuring compliance with preventive care guidelines: standardized

patients, clinical vignettes, and the medical record. J Gen Intern Med. 2000;15:782–8.

14. Hamm RM, Hicks RJ, Bemben DA. Antibiotics and respiratory infections: do antibiotic prescriptions improve outcomes? J Oklahoma State Med Assoc. 1996;89:267–74.

15. Hamm RM, Hicks RJ, Bemben DA. Antibiotics and respiratory infections: are patients more satisfied when expectations are met? J Fam Pract. 1996;43:56–62.

16. Gold RL. Roles in sociological field observations. Soc Forces. 1958;36:217–23.

17. Pretzlik U. Observational methods and strategies. Nurse Res. 1994;2:13–21.

18. Mulhall A. In the field: notes on observation in qualitative research. J Adv Nurs. 2003;41:306–13.

19. Ridderikhoff J. Medical problem-solving: an exploration of strategies. Med Educ. 1991;25:196–207.

20. André M, Borgquist L, Foldevi M, Molstad S. Asking for 'rules of thumb': a way to discover tacit knowledge in general practice. Fam Pract. 2002;19:617–22.

21. Dhami MK, Harries C. Fast and frugal versus regression models in human judgement. Thinking Reasoning. 2001;7:5–27.

22. Fischer S. Diagnosis and treatment of respiratory tract infections in general practice: survey of 273 patients at 30 general practitioners. Doctoral dissertation. Available from: http://webdoc.sub.gwdg.de/diss/2003/fischer_susanne/fischer_susanne.pdf

23. Lamberts H, Wood M, Hofmans-Okkes, eds. The International Classification of Primary Care in the European Community. Oxford, UK: Oxford University Press; 1993.

24. Pepe MS. Receiver operating characteristic methodology. J Am Stat Assoc. 2000;95:308–11.

25. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting. J Educ Res. 2002;96:3–14.

26. SAS Institute, Inc. SAS/STAT. User's Guide Version 8. Cary, NC: SAS Institute; 1999.

27. Muche R, Ring C, Ziegler C. Development and Validation of Prognostic Models Basing on Logistic Regression [in German]. Aachen, Germany: Shaker Press; 2005.

28. Mangione-Smith R, Elliott MN, McDonald L, McGlynn EA. An observational study of antibiotic prescribing behaviour and the Hawthorne effect. Health Serv Res. 2002;37:1603–23.

29. Buchan IC, Richardson IM. Time Study of Consultations in General Practice. Scottish Health Service Studies No. 27. Edinburgh: Scottish Home and Health Department; 1973.

30. Robbins JA, Bertakis KD, Helms LJ, Azari R, Callahan EJ, Creten DA. The influence of physicians practice behaviour on patient satisfaction. Fam Med. 1993;25:17–20.

31. Simel DL, Rennie D. The clinical examination: an agenda to make it more rational. JAMA. 1997;277:572–4.

32. Blomgren K, Hytönen M, Peillinen J, Relander M, Pitkäranta A. Diagnostic accuracy of acute maxillary sinusitis in adults in primary care. Scand J Prim Health Care. 2002;20:40–4.

33. Johansson L, Mansson NO. Rapid test, throat culture and clinical assessment in the diagnosis of tonsillitis. Fam Pract. 2003;20:108–11.

34. Lindbaeck M, Hjortdahl P. The clinical diagnosis of acute purulent sinusitis in general practice—a review. Br J Gen Pract. 2002;52:491–5.

35. Klaudt MR, Steinbach WJ, Sectish TC. Clinical considerations in the diagnosis of otitis media. Curr Allergy Asthma Rep. 2003;3:313–20.

36. Hueston WJ, Eberlein C, Johnson D, Mainous AG. Criteria used by clinicians to differentiate sinusitis from viral upper respiratory tract infection. J Fam Pract. 1998;46:487–92.

37. McIsaac WJ, Goel V. Effect of an explicit decision-support tool on decisions to prescribe antibiotics for sore throat. Med Decis Making. 1998;18:220–8.

38. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Med Decis Making. 1981;3:239–46.

39. Fischer T, Fischer S, Kochen MM, Hummers-Pradier E. Influence of patient symptoms and physical findings on general practitioners' treatment of respiratory tract infections: a direct observation study. BMC Fam Pract. 2005;6:6.

40. van Duijn HJ, Kuyvenhoven MM, Butler CC, Coenen S, Verheij TJ. Outpatient antibiotic use in three European countries: exploration of possible determinants. Eur J Gen Pract. 2005;11:139–40.

41. Hutchinson JM, Jelinski S, Hefferton D, Desaulniers G, Parfrey PS. Role of diagnostic labeling in antibiotic prescription. Can Fam Physician. 2001;47:1217–24.

42. Rautakorpi UM, Lumio J, Huovinen P, Klaukka T. Indication-based use of antimicrobials in Finnish primary health care: description of a method for data collection and results of its application. Scand J Prim Health Care. 1999;17:93–9.

43. Hamm RM. Medical decision scripts: combining cognitive scripts and judgement strategies to account fully for medical decision making. In: Hardman D, Macchi L, eds. Thinking: Psychological Perspectives on Reasoning, Judgement and Decision Making. London: John Wiley; 2003. p 315–45.

44. Hamm RM. Theory about heuristic strategies based on verbal protocol analysis: the emperor needs a shave. Med Decis Making. 2004;24:681–6.

45. Tabachnik BG, Fidell LS. Using Multivariate Statistics. 5th ed. Needham Heights, MA: Allyn & Bacon; 2007.