Erschienen in: Medical Decision Making ; 37 (2017), 1. - S. 27-34
https://dx.doi.org/10.1177/0272989X16653401

ORIGINAL ARTICLE

The Phenomenology of the Diagnostic Process: A Primary Care–Based Survey

Norbert Donner-Banzhoff, MD, MSc, Judith Seidel, MD, Anna Maria Sikeler, MD, Stefan Bösner, MD, MPH, Maria Vogelmeier, MD, Anja Westram, PhD, Markus Feufel, PhD, Wolfgang Gaissmaier, PhD, Odette Wegwarth, PhD, Gerd Gigerenzer, PhD

Background. *While dichotomous tasks and related cognitive strategies have been extensively researched in cognitive psychology, little is known about how primary care practitioners (general practitioners [GPs]) approach ill-defined or polychotomous tasks and how valid or useful their strategies are.* ***Objective***. *To investigate cognitive strategies used by GPs for making a diagnosis.* ***Methods***. *In a cross-sectional study, we videotaped 282 consultations, irrespective of presenting complaint or final diagnosis. Reflective interviews were performed with GPs after each consultation. Recordings of consultations and GP interviews were transcribed verbatim and analyzed using a coding system that was based on published literature and systematically checked for reliability.* ***Results***. *In total, 134 consultations included 163 diagnostic episodes. Inductive foraging (i.e., the initial, patient-guided search) could be identified in 91% of consultations. It contributed an average 31% of cues obtained by the GP in 1 consultation. Triggered routines and descriptive questions occurred in 38% and 84% of consultations, respectively. GPs resorted to hypothesis testing, the hallmark of the hypothetico-deductive method, in only 39% of consultations.* ***Limitations***. *Video recordings and interviews presumably interfered with GPs' behavior and accounts. GPs might have pursued more hypotheses and collected more information than usual.* ***Conclusions***. *The testing of specific disease hypotheses seems to play a lesser role than previously thought. Our data from real consultations suggest that GPs organize their search for information in a skillfully adapted way. Inductive foraging, triggered routines, descriptive questions, and hypotheses testing are essential building blocks to make a diagnosis in the generalist setting.* ***Key words:*** *general practice; decision making; hypothesis testing; qualitative research; diagnosis, family medicine; cues.* ***(Med Decis Making 2017;37:27–34)***

Making a diagnosis is perhaps the most intellectually challenging task of the physician. Often within minutes or even seconds, experienced physicians manage to narrow the range of possible diseases. While invasive and costly diagnostic procedures are widely debated by researchers and health planners, it is the patient's history that provides the most important material for this task.[1]

To a physician making a diagnostic assessment, plenty of information is available, ranging from the patient's utterances, knowledge of the patient's history, disease prevalences, visual impression of the patient, and findings from the physical examination. Given the limited capacity of the human brain and that much information is noise,[2] physicians can and should process only part of it. But how do they collect, select, and, inevitably, ignore information? When do they stop and when do they continue to collect more?

In medicine, the most influential model of the diagnostic process has perhaps been the hypothetico-deductive method proposed by Elstein and others.[3] According to this view, early in the encounter with the patient, physicians form diagnostic hypotheses in their minds. These guide further data collection, which aims at confirmation or disconfirmation of possible explanations for the patient's problem. This may result in an iterative process of rejection and reformulation hypotheses until an adequate conclusion has been found.[4]

It remains unclear, however, which processes precede the formulation of the first hypothesis. Once they are entertaining one or more hypotheses, physicians have narrowed the range of possible

DOI: 10.1177/0272989X16653401

explanations from several hundreds, or even thousands, to perhaps 3 or 4 at most. This reduction of uncertainty must be regarded as a considerable achievement, especially in generalist settings, such as primary care or hospital emergency departments.

We have proposed ''inductive foraging'' to have a central role in the first stage of diagnostic data collection. Inductive foraging refers to an initial open search for information guided by the patient. This stage usually starts with an open prompt by the physician, giving the patient the opportunity to elaborate on his or her problem. It can be terminated by the patient or by the physician interrupting the patient.[5] Inductive foraging is more than the patient stating a presenting complaint. With the diagnostic task represented as the search through an almost infinite problem space, inductive foraging helps define and limit the problem space. It is only the patient who can provide this indispensable initial guidance. For the first time, we here report empirical data about the occurrence of this and related strategies (for the definition and operationalization of strategies, see Table 1 and Figure 1).

Although the initial diagnostic assessment achieved by physicians is important, it is difficult to investigate, resulting in a plethora of discussion and educational statements, but only a few empirical studies of the process itself. We therefore undertook a survey of real-life primary care consultations. Our objective was to explore cognitive strategies used by general practitioners (GPs) for their diagnostic assessment. Based on our data, we propose a phenomenology that has implications for the generalist setting and any other setting where a large number of diagnostic possibilities must be considered.

Received 23 September 2015 from the Department of Primary Care, Philipps University of Marburg, Marburg, Germany (ND-B, JS, AMS, SB, MV); Harding Center for Risk Literacy, Max-Planck-Institute for Human Development, Berlin, Germany (AW, MF, OW, GG); and Department of Psychology, University of Konstanz, Konstanz, Germany (WG). This research was supported by the German Research Foundation (DFG) with grants DO 513/2-1 and GI 170/8-1. All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that ND-B, JS, AMS, SB, MV, AW, MF, WG, OW, and GG have no relationships with companies that might have an interest in the submitted work in the previous 3 years; their spouses, partners, or children have no financial relationships that may be relevant to the submitted work; and ND-B, JS, AMS, SB, MV, AW, MF, WG, OW, and GG have no nonfinancial interests that may be relevant to the submitted work. The funder of the study had not influence on study design, collection, analysis, and interpretation of data; on the writing of the report; or the decision to submit the article for publication. All authors had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author (ND-B) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. Data sharing: no additional data available. Revision accepted for publication 13 May 2016.

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://mdm.sagepub.com/supplemental.

Address correspondence to Stefan Bösner, MD, Department of Primary Care, Philipps University of Marburg, Karl-von-Frisch-Str. 4, D-35043 Marburg, Germany; telephone: +49 6421 286 5122; fax: +49 6421 286 5121; e-mail: boesner@staff.uni-marburg.de.

METHODS

Twelve full-time GPs in the Marburg-Biedenkopf district of Hessen, Germany, were asked to take part in the study and all agreed to participate. We required participants to have been in practice for at least 5 years and to have under- or postgraduate teaching experience.

We covered 3 half-day sessions with each participating GP. Patients were included irrespective of their symptoms or possible diagnoses. Only consultations exclusively planned for nondiagnostic reasons, such as chronic disease monitoring or follow-up for a previously identified problem, were excluded beforehand.

Participating GPs informed each patient about the study and asked for written consent to participate and have his or her consultation video-recorded. GPs were instructed not to address the patient's presenting complaint(s) at this stage. After consent was obtained, consultations proceeded as usual. Sessions were scheduled so that after each consultation, GPs had sufficient time for a semistructured interview to explain their diagnostic reasoning. These interviews were also video-recorded. GPs were asked to use the initial utterance by the patient as the starting point of their reflection. GPs elaborated on their first impression and previous knowledge of each patient, as well as diagnostic hypotheses considered.

Recordings of consultations and GP interviews were transcribed verbatim. Transcripts were coded with MAXQDA software for qualitative data analysis.[6] We defined a consultation as containing a diagnostic episode when the patient brought up a new complaint, which resulted in some kind of data collection by the GP, such as taking a history or examining the patient.

Table 1 Cognitive Strategies: Definitions and Operationalization

	Definition	Operationalization	Illustrative Example
Inductive foraging	The patient is guiding the physician to relevant problem areas. Search of a wide problem space is thus possible (see Figure 1a). This is the only phase controlled by the patient.	Open question or invitation by physician regarding the reason for visit; subsequent time period with patient elaborating spontaneously. Usually occurs at the beginning of the consultation but occasionally marking the beginning of a second diagnostic episode within the same consultation. Terminated either by the patient pausing or suggesting a diagnosis himself or herself or by the physician interrupting, usually with a closed question.	GP: *What brings you here today?* Patient: *What I mentioned before; I stop breathing at night.* GP: *Yes.* Patient: *While I'm sleeping. . . . My wife says, when it happens, it happens several times, and it takes 30 seconds until after she has given me a kick, until I start breathing again.* GP: *Are you snoring?* [continued below "hypothesis testing"] (0310-P)
Descriptive questioning	Physician obtains descriptions of symptoms or problems. Phase controlled by physician.	Closed questions aiming at descriptions of a symptom or problem already mentioned by the patient.	GP: *And where, exactly, do you feel the pain?* (0212-P)
Triggered routines	Physician obtains information on problem area becoming relevant because of symptom or problem mentioned by patient (see Figure 1b). Phase controlled by physician.	Closed questions, exploring a problem area, such as an organ system. Triggered by symptoms mentioned by the patient. Formal clinical prediction rules also fulfill this definition.	[Patient complaining of diarrhea] GP: *Blood in your stool?* Patient: *No.* GP: *Do you feel sick? Vomiting?* Patient: *Not really, but I feel a bit funny—not all the time though. . . . And there was a loud noise in my tummy last night, but that's gone now.* GP: *OK, any temperature?* Patient: *No, I didn't take it.* (0301-P)
Hypothesis testing	Physician evaluates defined hypotheses (diseases) potentially explaining the patient's problem. Following the space metaphor, this corresponds to deep digging for findings or abnormalities specifically associated with hypothesis (Figure 1c). Phase controlled by physician.	Closed questions related to a specific disease, aiming at confirmation or disconfirmation. Hypothesis either explicitly mentioned in interview or questions asked relate to a specific disease as assessed by a medically trained coder.	GP: *Are you snoring? Has your wife mentioned that you snore?* Patient: [always answers in the affirmative] GP: *How do you feel when you get up in the morning? Tired? Exhausted?* Patient: . . . GP: *You are saying, when you sit comfortably, you easily fall asleep. . . . There is a possibility that you have sleep apnea.* (0310-P)

GP, general practitioner.

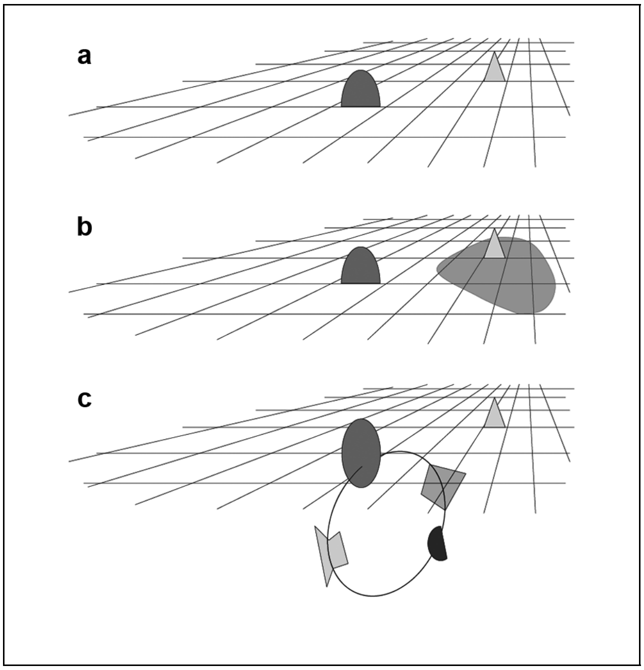

Figure 1 Illustration of cognitive strategies: (a) inductive foraging, (b) triggered routine, and (c) hypothesis testing. The plane stands for the problem space of the patient. Symptoms, which the patient would present if given enough time, are shown above the surface (geometrical shapes). Findings accessible only by directed questioning or search are hidden below the surface (c). An initial search guided by the patient will lead physicians to problem areas where directed search ("digging") will be highly efficient. Efficiency is lost if physicians start directed search (hypothesis testing) too early.

Drawing from previously published work,[3,5,7] we developed a coding tree composed of categories describing GPs' diagnostic reasoning and data-collecting behavior (available upon request). Extensive discussion in our group and several iterative loops of coding and modification of the coding tree resulted in operational definitions of categories. After completing this phase, we tested the reliability of our coding procedure on 3 consultations and 3 interviews, all randomly selected. Two blinded independent raters double-coded the material with high reliability, resulting in a weighted mean of 84% agreement.[8] Remaining discrepancies in coding were resolved by discussion with senior members of the research team.

To quantify the yield of cognitive strategies, we identified all verbal and nonverbal diagnostic cues presented by patients. A cue is any piece of information made available to the GP, either spontaneously or during questioning. Nonverbal cues were usually GPs' visual impressions and findings from the physical examination. For this kind of analysis, we selected a random sample of 5 consultations from each practice, including GP interviews (n = 58). In one practice, where only 3 consultations were included, we analyzed all consultations in this manner.

Each item of diagnostic information gathered by the GP during the first diagnostic episode of a consultation was coded irrespective of whether it was obtained spontaneously or in response to a closed question. Information modifying a previous cue in a relevant way, such as the quality of pain or specific triggers causing a symptom, was counted as a separate cue. We could thus quantify to what degree different strategies contributed to the information acquired from each patient. For each strategy, we report frequency of occurrence within the consultation, including 95% confidence intervals.

Two authors (ND-B and GG) developed the general design of the study. Details of the study protocol, such as recruitment and data collection and analysis, were discussed by the entire study team. Two authors (MV and AW) collected data in participating practices and lead interviews with GPs, and 2 qualified physicians (JS, AMS) coded and analyzed the text material. They were directly supervised by ND-B and SB. Selected text passages were discussed by the entire team, which included researchers with medical (ND-B, JS, AMS, SB, MV) and cognitive psychology (WG, MAF, OW, AW, GG) backgrounds. ND-B is the guarantor for this work. The study obtained ethical approval by the Ethics Committee of the Faculty of Medicine, University of Marburg (39/2010).

RESULTS

Participating GPs and Patients

Of 12 participating GPs, 5 were female. On average, they were 53 years old and had been in primary care practice for 21 years (for details, see Table 2).

After the exclusion of consultations with technically unsatisfactory recordings and/or without diagnostic content, 134 consultations with 163 diagnostic episodes were available for analysis (see flowchart in Figure 2; for patients' characteristics, see Table 3 and Suppl. Table SA). Practices contributed between 3 and 16 consultations with diagnostic content. Their average duration was 9 minutes 59 seconds (range, 2 minutes 45 seconds to 28

Table 2 Participating General Practitioners (n = 12)

Characteristic	Value
Sex: proportion female, *n*	5/12
Age, median (range), y	53 (49–62)
Practice, *n*	
Single handed	2
Group (2 partners and more)	10
Years in primary care practice, median (range)	21 (9–30)
Location of practice, *n*	
Urban—small town	9
Rural	3

Table 3 Participating Patients (n = 134)

Characteristic	Value
Age, mean (standard deviation), y	54.6 (4.8)
Sex: female, *n* (%)	85 (63)
Family, *n* (%)	
Single/divorced/widow	47 (35)
Married/living with partner	81 (60)
Adolescent living with parents	6 (5)
Level of education,[a] *n* (%)	
Low	31 (23)
Intermediate	73 (55)
High	30 (22)

a. Levels defined according to German educational system: low, basic (*Hauptschule*) or no secondary education; intermediate, equivalent to O-levels (*Realschule, mittlere Reife*, etc.); high, graduation from grammar school or equivalent (*Abitur*).

minutes 15 seconds). Reflective interviews lasted between 2 and 18 minutes (median, 6 minutes 35 seconds).

Inductive Foraging

Inductive foraging (i.e., the initial search guided by the patient) could be identified in 122 (91%) consultations (for definition and operationalization of strategies, see Table 1 and Figure 1). In 5 of the remaining 12 cases, the beginning of the consultation was not recorded for technical reasons. Inductive foraging is thus likely to occur even more frequently. The median duration of this phase was 34 seconds (range, 6–176 seconds). Inductive foraging took a proportion of 14.6% (median) of diagnostic episodes (range, 1.7%–93.1%). See online supplemental material for the effect of practice and presenting symptoms (Suppl. Table SD and SE).

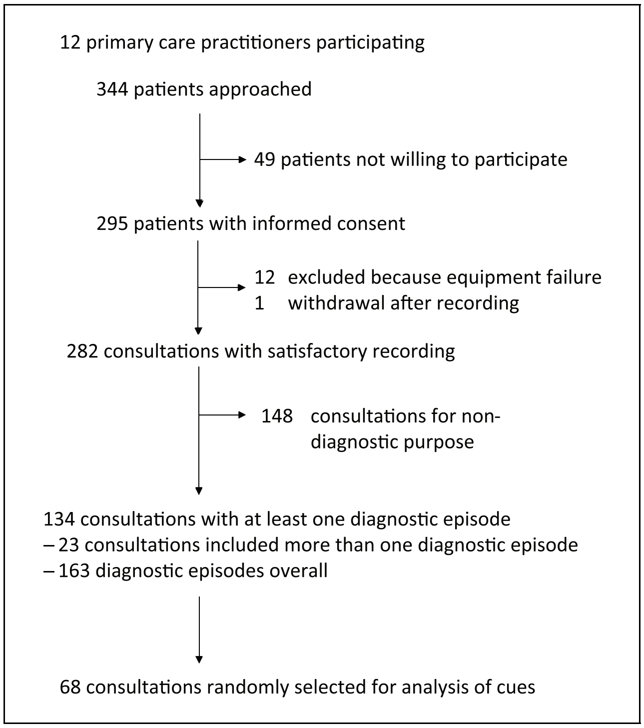

Figure 2 Flowchart of participating physicians and patients.

In diagnostic episodes beginning without inductive foraging, GPs started with testing a hypothesis derived from previous encounters or from information obtained by receptionists.

In 70 (57%) cases, inductive foraging ended by the GP interrupting the patient, usually by asking a closed question. In 6 consultations, there was a return to inductive foraging at a later stage.

GPs employed several tactics to support patients in their foraging for complaints and symptoms; paraphrasing, verbal prompts, and nonverbal encouragements were the most frequently used.

Triggered Routines and Descriptive Questions

We identified an intermediate stage in the diagnostic process in which GPs explored a limited problem area. GPs consistently used these "triggered routines" for identical symptoms (see Table 1 and Figure 1 for definitions of strategies). This strategy was observed in 62 (38%) of 163 diagnostic episodes.

GPs asked descriptive questions in 137 (84%) diagnostic episodes. A physical examination was conducted in 120 (89%) consultations.

Table 4 Cognitive Strategies: Occurrence and Contribution (Diagnostic Cues)

Cognitive Strategy	Diagnostic Episodes with Strategy Occurring (n = 163), n (%; 95% CI)	Average Contribution, % of Cues Presented in Each Diagnostic Episode (n = 68)
Inductive foraging	128 (79; 73–85)	31
Descriptive questions	137 (84; 87–90)	25
Triggered routines	62 (39; 31–45)	12
Hypothesis testing	63 (39; 31–46)	12
Physical examination[a]	138 (85; 79–90)	20

a. Although the physical examination is not a cognitive strategy per se, a relevant number of cues were obtained during that phase.

Hypothesis Testing

Collecting information confirming or disconfirming specific diseases is the hallmark of the hypothetico-deductive method. However, GPs used this strategy in only 63 of 163 diagnostic episodes (39%).

Relative Contribution of Cognitive Strategies

Occurrence and duration of the aforementioned strategies differed by GP and consultation. In the subsample of consultations analyzed quantitatively by diagnostic cues, inductive foraging contributed an average of 31% of diagnostic cues presented in each consultation/episode. Only 12% of cues were obtained by hypothesis testing (see Table 4).

DISCUSSION

In our study of 134 consultations, we found that GPs use inductive foraging, triggered routines, and hypothesis testing for the diagnostic evaluation of their patients. In this primary care sample, the contribution of focused hypotheses testing was limited, whereas the more open strategies, such as inductive foraging, descriptive questions, or triggered routines, contributed the majority of diagnostic cues obtained by GPs.

How to Explain the Limited Reliance on Hypothesis Testing: Adaptive Cognitive Strategies

It may be surprising that GPs organized their information search using specific hypotheses in only 39% of consultations and obtained an average of 12% of cues this way. However, this behavior can be understood as an adaptation to the primary care work environment.

The hypothetico-deductive approach requires awareness of at least 1 defined hypothesis and subsequently selected tests (mostly items from the history) confirming or disconfirming the hypothesis. If several hypotheses are entertained, cues have to be evaluated with regard to each hypothesis. This procedure is cognitively extremely demanding and, with several hypotheses evaluated simultaneously, perhaps unachievable. In contrast, individuals in complex and uncertain environments have repeatedly been shown to employ fast and frugal strategies adapted to their setting.[2,9] Inductive foraging and triggered routines seem to be appropriate strategies to search the problem space without premature restriction. With inductive foraging, GPs are especially receptive to patients leading them to areas of concern. Focused hypothesis testing as a cognitively more demanding strategy is used only if relevant information is still lacking. In most consultations, GPs apparently regard information obtained by open strategies as sufficient. Moreover, an early switch to hypothesis testing restricts the range of possible explanations to those explicitly considered by the GP. Once the GPs decide to evaluate specific hypotheses, they inevitably control the communication with the patient by asking closed questions. At that stage, the patient will then only mention cues regarded as relevant and asked for by the GP.

Although physicians may already think of specific diseases at the inductive foraging and triggered routine stages, hypotheses are not required for information searches at this point. We suggest that collecting data according to specific hypotheses should be suspended at these stages. Otherwise, GPs will interrupt the patient prematurely, take control of communication, and inordinately restrict the range of possible explanations.[5]

Comparison with the Literature

In their seminal study, Elstein and others[3,10] showed that hypotheses form early in the minds of physicians taking a history from a patient and guide subsequent data gathering. The contribution of hypothesis testing in this study may have been somewhat inflated by the laboratory setting.

Ridderikhoff[11,12] analyzed primary care practitioners' diagnostic processes. He already observed that hypothesis testing was rare and felt that searching for data served the purpose of evoking further hypotheses rather than testing them.

Heneghan and others[7] suggested that diagnostic reasoning starts with an initiation stage, followed by a refinement stage. They included the inductive foraging stage under the heading of "presenting complaint." However, our recordings of consultations show that inductive foraging at this stage produces much richer data than the presenting complaint itself. The strategies that Heneghan and others[7] proposed for the refinement stage (e.g., restricted rule-outs or stepwise refinement) could not be reliably identified in our sample of GPs who were not previously primed to use these categories.

Routines are often mentioned negatively as a fallback tool for the inexperienced[13] or as a strategy for mature physicians to gain time.[14,15] Instead, we postulate a positive role for triggered routines and descriptive questions: they help explore areas of interest that emerge during the consultation when data collection guided by specific hypotheses would unnecessarily reduce the problem space.

Strengths and Weaknesses

Our research differs from most published work in this field in that we investigated real patient-physician encounters. We included consultations concerning any symptom or disease occurring in primary care. The results thereby reflect processes being used across the entire spectrum of patients' problems. The sample is purposely biased toward experienced practitioners and those actively involved in teaching, thus enabling us to obtain in-depth accounts of experienced GPs' reasoning. However, it is unlikely that cognitive processes among our convenience sample are relevantly different from the population of primary care physicians.

Despite our efforts to preserve the natural character of the setting studied, video recordings and interviews presumably interfered with GPs' behavior and accounts. From informal feedback and our own impressions, we learned that sometimes GPs were more active, pursuing more hypotheses and collecting more information than usual. As a result, the frequency of hypothesis testing is perhaps overestimated in our study.

The differentiation between cognitive strategies was not always clear. Determining whether a question asked by the GP should be understood as triggered routine or hypothesis evaluation was sometimes difficult. We iteratively developed clear definitions for the concepts studied, resulting in a concise format (see Table 1). This, together with extensively training medically qualified observers, led to a high reliability of our data analysis. We thus avoided more invasive probing of whether a specific hypothesis was entertained at a certain point in time or not. This would have interfered with GPs' reasoning to an unacceptable degree. We report cues obtained by the physical examination separately, although this is not a specific cognitive strategy. Physical examination in primary care is always focused, usually consists of a fixed sequence of acts, and thus resembles triggered routines. On the other hand, GPs may have been influenced by specific hypotheses at this stage of the diagnostic episode. Similar considerations apply to investigations or referral (see Suppl. Appendix C). When GPs asked questions while they examined the patient, these verbal expressions were coded accordingly.

Although the quantitative investigation of cues obtained by GPs is a major strength of our study, we could only evaluate their number, not their relevance. One can surmise that cues obtained by triggered routines or hypothesis testing were more important for the diagnosis than those provided during inductive foraging.

Physicians use a broad spectrum of mental processes to assess what is wrong with their patients. Our study had to concentrate on a limited number of aspects, leaving open how physicians construct patterns of findings obtained[16] or how emotions, gut feelings, or other nonanalytical strategies factor into the process.[17,18] However, we provide a phenomenology of how GPs organize their information search during consultation. We regard our combination of observed GPs' behavior and subsequent reflective interviews as a valid way to triangulate findings on a difficult research topic. Moreover, the large number of consultations allowed for the quantitative estimation of strategies used by GPs.

CONCLUSION

Based on our analysis of 134 consultations by experienced practitioners, we propose the sequence of inductive foraging, descriptive questioning, triggered routines, and deductive testing as strategies adapted to primary care and other generalist settings. This phenomenology clarifies the stages preceding and often replacing the testing of specific diagnostic hypotheses.

Our findings also have implications for teaching. Feedback and examination formats tend to be biased toward hypothesis testing and a directive style of inquiry.[19–21] The concepts derived from our study may motivate clinical teachers to encourage the use of a broader array of strategies.

By allowing the patient to control the initial data collection process and taking control only at later stages, GPs adapt to a setting with multiple diagnostic possibilities. ''Inductive foraging'' not only eases the patient's sense of distress but is essential to define the diagnostic problem space.[5] The worlds of patient-centered medicine and diagnostic reasoning can thereby be reconciled.

REFERENCES

1. Hampton JR, Harrison MJ, Mitchell JR, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. BMJ. 1975;2:486–9.

2. Gigerenzer G, Todd PM; ABCRG. Simple Heuristics That Make Us Smart. New York: Oxford University Press; 1999.

3. Elstein A, Schulman L, Sprafka S. Medical Problem-Solving: An Analysis of Clinical Reasoning. Cambridge, MA: Harvard University Press; 1978.

4. Norman G, Barraclough K, Dolovich L, Price D. Diagnosis in general practice iterative diagnosis. BMJ. 2009;339:b3490.

5. Donner-Banzhoff N, Hertwig R. Inductive foraging: improving the diagnostic yield of primary care consultations. Eur J Gen Pract. 2014;20(1):69–73.

6. MAXQDA. Berlin, Germany: VERBI Software; 2010.

7. Heneghan C. Diagnostic strategies used in primary care. BMJ. 2009;338:b946.

8. Holsti OR. Content Analysis for the Social Sciences and Content Analysis for the Social Sciences and Humanities. Reading, MA: Addison-Wesley; 1969.

9. Gigerenzer G, Gaissmaier W. Heuristic decision making. Annu Rev Psychol. 2011;62:451–82.

10. Elstein AS, Kagan N, Shulman LS, Jason H, Loupe MJ. Methods and theory in the study of medical inquiry. J Med Educ. 1972;47:85–92.

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

12. Ridderikhoff J. Problem-solving in general-practice. Theor Med. 1993;14:343–63.

13. Leaper DJ, Gill PW, Staniland JR, Horrocks JC, de Dombal FT. Clinical diagnostic process: an analysis. BMJ. 1973;3:569–74.

14. Barrows HS, Bennett K. The diagnostic (problem solving) skill of the neurologist: experimental studies and their implications for neurological training. Arch Neurol. 1972;26:273–7.

15. Gale J, Marsden P. The role of the routine clinical history. Med Educ. 1984;18:96–100.

16. Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise—theory and implications. Acad Med. 1990;65:611–21.

17. Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: the role of experience. Med Educ. 2007;41: 1140–5.

18. Stolper E, van de Wiel M, van Royen P, van Bokhoven M, van der Weijden T, Dinant GJ. Gut feelings as a third track in general practitioners' diagnostic reasoning. J Gen Intern Med. 2011;26: 197–203.

19. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing: let's preach what we practice. N Engl J Med. 1983; 309:921–3.

20. van Puymbroeck H, Remmen R, Denekens J, Scherpbier A, Bisoffi Z, Van Den Ende J. Teaching problem solving and decision making in undergraduate medical education: an instructional strategy. Med Teacher. 2003;25:547–50.

21. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. N Engl J Med. 2006;355:2217–25.