The Role of Physicians' First Impressions in the Diagnosis of Possible Cancers without Alarm Symptoms

Olga Kostopoulou, PhD, Miroslav Sirota, PhD, Thomas Round, MBBS, Shyamalee Samaranayaka, MD, Brendan C. Delaney, MD

Background. *First impressions are thought to exert a disproportionate influence on subsequent judgments; however, their role in medical diagnosis has not been systematically studied. We aimed to elicit and measure the association between first impressions and subsequent diagnoses in common presentations with subtle indications of cancer.* ***Methods.*** *Ninety UK family physicians conducted interactive simulated consultations online, while on the phone with a researcher. They saw 6 patient cases, 3 of which could be cancers. Each cancer case included 2 consultations, whereby each patient consulted again with nonimproving and some new symptoms. After reading an introduction (patient description and presenting problem), physicians could request more information, which the researcher displayed online. In 2 of the possible cancers, physicians thought aloud. Two raters coded independently the physicians' first utterances (after reading the introduction but before requesting more information) as either acknowledging the possibility of cancer or not. We measured the association of these first impressions with the final diagnoses and management decisions.* ***Results.*** *The raters coded 297 verbalizations with high inter-rater agreement (Kappa = 0.89). When the possibility of cancer was initially verbalized, the odds of subsequently diagnosing it were on average 5 times higher (odds ratio 4.90 [95% CI 2.72 to 8.84], P < 0.001), while the odds of appropriate referral doubled (OR 1.98 [1.10 to 3.57], P = 0.002). The number of cancer-related questions physicians asked mediated the relationship between first impressions and subsequent diagnosis, explaining 29% of the total effect.* ***Conclusion.*** *We measured a strong association between family physicians' first diagnostic impressions and subsequent diagnoses and decisions. We suggest that interventions to influence and support the diagnostic process should target its early stage of hypothesis generation.* ***Key words:*** *family medicine; provider decision making; cognitive psychology; heuristics and biases.* ***(Med Decis Making 2017;37:9–16)***

It is a characteristic of human judgment to formulate hypotheses quickly.[1] Research in social and cognitive psychology has found a disproportionate influence of early hypotheses on final judgments.[2–6] This has been attributed to the working memory being less loaded at the start of a judgment task; therefore, initial information receives more attention and is better encoded.[3] Early impressions can be maintained and carried through to the final judgment via biased information search[7] and/or biased information processing.[8] Physicians, too, have been found to generate hypotheses in the first few seconds of the clinical encounter and with little information.[9] Physicians' self-reports suggest an association between

Received 30 November 2015 from the Department of Surgery and Cancer, Imperial College London (OK, BCD); Department of Psychology, University of Essex, UK (MS); Department of Primary Care and Public Health Sciences, King's College London, UK (TR); and Department of Family Medicine, University of Sri Jayewardenepura, Sri Lanka (SS). This study was funded by a Cancer Research UK project award (National Awareness and Early Diagnosis Initiative) to Olga Kostopoulou, grant C33754/A12222. Ethical approval was granted by the Proportionate Review Subcommittee of the West London (REC 2), reference 11/LO/0079. Revision accepted for publication 17 March 2016.

DOI: 10.1177/0272989X16644563

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

Address correspondence to Olga Kostopoulou, Imperial College London, 5th Floor Medical School Building, St Mary's Campus, Norfolk Place, London W2 1PG, UK; telephone: (+44) 20 7594 9120; e-mail: o.kostopoulou@imperial.ac.uk.

these early hypotheses and subsequent diagnosis.[10,11] Nevertheless, this association has not been systematically studied and measured in practicing physicians' diagnostic judgments.

We aimed to elicit and measure this association in situations of known diagnostic difficulty, namely, that of family physicians diagnosing common presentations with subtle indications of cancer. Family physicians are tasked with the diagnosis of potentially serious but as yet undifferentiated problems. Diagnosing cancer is inherently difficult: fewer than half of all cancer patients present with so-called ''alarm symptoms'' or ''red flags,''[12–14] which are, supposedly characteristic features of cancer. Most such symptoms have generally low positive predictive values.[15] Diagnostic delays in cancer can lead to significant patient harm, as the disease can become less treatable with time. Retrospective studies of diagnostic errors, using record screening and analysis of case reports,[16] as well as analyses of medicolegal cases,[17] have highlighted cancers as being a common diagnostic problem that may lead to error. Using short vignettes, 1 study found that family physicians generated common diagnoses first, before they generated more serious and less common possibilities.[18] We expect that these initial judgments will exert disproportionate influence on final diagnoses, especially in the absence of strongly diagnostic information.

METHODS

Materials

We conducted evidence reviews in relation to symptoms and signs of colorectal cancer, lung cancer, and myeloma. On the basis of these, we constructed 3 patient cases in which cancer was a possible diagnosis and included sufficient detail so that they could be employed in an interactive simulated consultation (example in the supplemental appendix). We also used 3 cases from a previous study: a man with typical symptoms of gout, a small child with typical asthma symptoms, and another small child with fever.[19] These were employed as decoys to prevent participants from forming an impression that all consultations were about possible cancers but also as practice cases to get participants used to the methodology. Only data from the cancer consultations were analyzed and are presented.

The patients in the cancer cases were older than 60 y and presented with 1 main, persistent symptom: constipation for 1 mo (colorectal cancer case), cough for 6 wk (lung cancer case), and back pain for more than 2 mo (myeloma case). The symptom could be explained by a more common, preexisting diagnosis: irritable bowel syndrome (colorectal cancer case), exacerbation of chronic obstructive pulmonary disease (lung cancer case), and mechanical back pain (myeloma case). There were no ''alarm symptoms,'' such as rectal bleeding, hemoptysis, or severe weight loss. All cancer patients (as well as 1 decoy patient) consulted twice. The second consultation was described to take place either 2 wk (colorectal and lung cases) or 6 wk after the first consultation (myeloma case). At the second consultation, the main symptom (cough, constipation, back pain) had not improved despite any treatment prescribed at the first consultation. The patients reported new symptoms, such as increased fatigue and breathlessness, and the results of some investigations (if ordered) could suggest an abnormality (e.g., slight anemia and inflammation). The lack of improvement in the patients' main symptom, the additional symptoms, and the abnormal test results constitute information that is incompatible with the more common competing diagnoses and warrant referral to specialist or referral for specialist investigations (e.g., colonoscopy, computed tomography [CT] scan).

Methodologies

We used 2 process-tracing methodologies: active information search[20] and think aloud.[21,22] Active information search involves participants requesting information in a step-by-step fashion, as they see fit, rather than being presented with the information all at once or in a sequence determined by the researcher. The methodology is well suited for the study of medical diagnosis, which is interactive and involves a stepwise search for information. The think aloud methodology allows researchers some access to covert thinking processes, such as hypotheses, assumptions, and inferences.

Procedure

Data collection took place remotely over the Internet, using a Web tool designed specifically for the study. Participants were on the phone with a researcher (M.S.), who operated the site and guided them through the task during a single session. All participants followed the same sequence and consulted with all the patients (Figure 1). The presentation order of the noncancer patients was fixed, while the presentation order of the cancer patients was

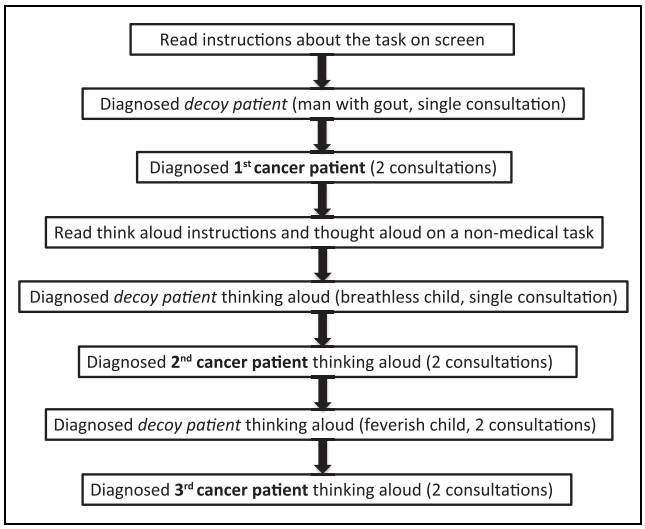

Figure 1 The sequence followed by all participants in the study. The presentation order of the cancer patients was randomized, while that of the decoy patients was fixed.

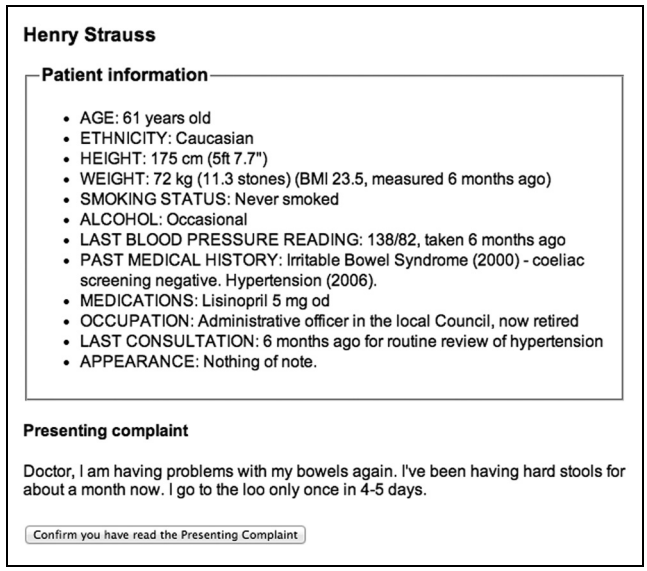

Figure 2 The initial information that all participants read: patient description and presenting problem (example from the colorectal cancer case).

randomized per participant. Participants were asked to think aloud while diagnosing the last 2 cancer patients that they encountered, while the first cancer patient was used as silent control.

At the start of each consultation, all participants read the same initial information about the patient: a short description and the presenting problem (Figure 2). They could then request more information in relation to history, physical examination, and investigations (that did not require referral to a specialist). We had prepared a set of answers to potential questions for each patient. After each question, the researcher chose the appropriate answer and displayed it on the participant's screen. If participants asked questions for which there was no predetermined answer, the researcher typed in the question, so that it was recorded, and selected appropriately from a set of generic responses (e.g., ''no,'' ''normal''). When participants wished to finish the consultation, they were asked to type in their working diagnosis (''What is your main working diagnosis? Enter only one'') and their differential (''If you have any other differential diagnoses, enter them below''). They were then asked to select their management from a list of options (more than 1 could be selected): prescribe, refer to specialist/for specialist tests, arrange follow-up, and/or ask patient to come back if symptoms persist. The system automatically recorded all information gathered, time, diagnoses, and management decisions. After participants gave their diagnosis and management at the first consultation, the patient presented again for a second consultation, unless he or she had already been referred to a specialist.

When think aloud was required, the researcher asked participants to read aloud the patient description and presenting problem shown on the screen. Given our focus on the initial phase of the diagnostic process, he always prompted them to keep talking after they had finished reading, unless they did so spontaneously. The researcher also prompted them to keep talking at various points during the process but not after each question, to avoid interfering with the diagnostic task. The think aloud protocols were audio recorded and transcribed verbatim.

Sample Size and Recruitment

The cases differed widely in their content. We thus assumed independence of responses within participants, as observed in our previous studies in which family physicians diagnosed a range of different patient cases.[19,23] Using the software G*Power 3.1, we estimated that in a 2-tailed logistic regression with a binary predictor, a conservative expected effect size (odds ratio of 2), 50% probability of the null hypothesis, 5% probability of type I error, and 80% power, 270 responses (90 physicians diagnosing 3 patients) would be sufficient to detect a relationship between initial hypotheses and final diagnosis.

We invited family physicians from London and southwest England to participate in ''a study of

clinical reasoning'' and did not mention cancer. We recruited participants either by e-mailing family physicians, who had taken part in other studies by our group, or via local clinical research networks. Recruitment continued until the required sample size was achieved. Participants received recompense for an estimated 3-h involvement at standard clinical rates. Data were collected between October 2013 and November 2014.

Analyses

The main outcome measure was whether the physician recorded a cancer diagnosis at the end of each consultation, either as the working diagnosis or in the differential. We coded this as either 1 (cancer diagnosed) or 0 (cancer not diagnosed). Management was coded as either 1 (appropriate referral) or 0 (no/inappropriate referral). We coded as 1 all referrals to the appropriate specialist: colorectal surgeon, gastroenterologist, and gastrointestinal team in the colorectal cancer case; respiratory or chest physician in the lung cancer case; and hematologist, rheumatologist, oncologist, and orthopedics in the myeloma case. We also coded as 1 all referrals for appropriate investigations: colonoscopy and sigmoidoscopy in the colorectal cancer case, CT scan and bronchoscopy in the lung cancer case, and magnetic resonance imaging and bone scan in the myeloma case. Referrals to the appropriate specialist were either for suspected cancer or for further investigations. Referrals to a different specialist for non–cancer-related reasons (e.g., to a cardiologist for echocardiogram, to a smoking cessation clinic, or for pulmonary physiotherapy) as well as no referrals were scored as 0.

We coded participants' questions (both those that the system recorded automatically, i.e., questions with a predetermined answer, and those that the researcher had typed in) as either cancer related or not. Cancer-related questions were those that could provide some evidence for cancer, irrespective of the patient's answer, based on the agreement of the 3 clinical authors (T.R., S.S., B.C.D.). For example, asking about blood in sputum, tiredness, appetite, and weight loss were all coded as cancer-related, although they differ in the strength of evidence that they can provide.

In the think aloud protocols, we singled out the participants' initial utterances, after they read the initial patient description and presenting problem (Figure 2) and before asking further questions. Two raters (O.K. and B.C.D.) coded these utterances independently as either 1 (cancer mentioned) or 0 (cancer not mentioned). After agreement of the clinical authors, we also coded as 1 instances in which participants did not mention cancer but mentioned ''malignancy,'' ''tumor,'' ''carcinoma,'' ''neoplasm,'' ''something sinister,'' and ''red flags'' (red flags were thought to refer to cancer only in the colorectal and lung cases; in myeloma, red flags could refer also to other conditions, such as central disc prolapse). A third coding category was used for instances in which verbalization was not sufficient to enable coding of 1 or 0. We used logistic regressions to explore the relationship between initial utterances (''first impressions'') and subsequent diagnoses and decisions. To test whether this relationship was explained by information search, we constructed a simple mediation model with the number of cancer-related questions as the mediator.

To ensure that the assumption of independence of responses within participants held, we repeated all the analyses as 2-level logistic regression models with random intercept and patient/consultation as a repeated measure. We also checked for any influences of thinking aloud by comparing performance on the consultations in which participants thought aloud (i.e., in the last 2 cancer cases encountered) with that on the consultations without thinking aloud (i.e., in the first cancer case encountered, used as silent control). We expected increased time but no differences in diagnoses and decisions. STATA 13.1 was used in all the analyses.

RESULTS

We recruited 90 family physicians: 50 were men (55.6%) and had an average experience of 12 y in family medicine (s = 8.8, median = 10, range = 0–36 y). Across patients and consultations, cancer was diagnosed on 51% of occasions. On 22.5% of occasions, cancer was the working diagnosis. Appropriate referrals were made on 42% of occasions (Table 1). As expected, thinking aloud significantly increased the time taken ($\bar{x}$ = 8.33 min v. 7.02 min for think aloud v. silent consultations, respectively; beta 1.31 [95% CI 0.61 to 2.00], $P < 0.001$) but had no influence on the outcome measures (diagnosis and referral). The simple and the 2-level logistic regression models produced almost identical results, validating our assumption of independence of responses. Below, we report results from the simple regression models.

There was a strong association between cancer diagnosis and appropriate referral: odds ratio (OR) 9.01 [5.78 to 14.04], $P < 0.001$. A significant increase

Table 1 Cancer Diagnosis Frequency, No. (%)

Patient		First Consultation			Second Consultation		
		Cancer Diagnosis	Appropriate Referral	Total	Cancer Diagnosis	Appropriate Referral	Total
Colorectal	Working/differential	54 (60)	32 (36)	90	42 (72)	39 (67)	58
	Working	18 (20)			21 (36)		
Lung	Working/differential	28 (31)	9 (10)	90	23 (30)	33 (43)	76
	Working	0			5 (7)		
Myeloma	Working/differential	51 (57)	49 (54)	90	25 (69)	25 (69)	36
	Working	37 (41)			18 (50)		
Total	Working/differential	133 (49)	90 (33)	270	90 (53)	97 (57)	170
	Working	55 (20)			44 (26)		

Note: Cancer diagnoses (working or in differential, and working only) and appropriate referrals by simulated patient and consultation. Working diagnoses of cancer were always followed by appropriate referral.

in appropriate referrals was observed at the second patient consultation: OR 2.66 [1.79 to 3.95], $P < 0.001$. In contrast, there was no significant increase in cancer diagnoses at the second consultation: OR 1.16 [0.79 to 1.70]. This suggests that a second consultation had an independent effect on referral, unrelated to the diagnosis. This was confirmed in a regression model with both diagnosis and consultation as predictors of referrals: diagnosis OR 10.29 [6.41 to 16.54] and consultation OR 3.39 [2.11 to 5.45], both at $P < 0.001$.

We coded 297 instances of initial verbalizations: 180 at first consultation (90 physicians thinking aloud on 2 cancer cases) and 117 at second consultation—there were fewer second consultations because, on 63 think aloud occasions, physicians had referred the patient at the first consultation. Interrater agreement was very high: colorectal cancer case Kappa 0.88, lung cancer case Kappa 0.90, and myeloma case Kappa 0.87, with an overall Kappa of 0.89. Discrepancies were resolved by discussion. There were 85 instances of insufficient verbalization, in which participants talked about what questions they wanted to ask the patient and what investigations they were going to order, but they did not explain ''why'' in diagnostic terms. These 85 instances were dropped from further analysis, which left 212 instances of first impressions: 108 instances where cancer was initially mentioned (51%) and 104 instances where it was not (49%). When cancer was initially mentioned, it was subsequently diagnosed in 62% of the consultations (67/108); when it was not initially mentioned, it was diagnosed in only 25% of the consultations (26/104). First impressions were strongly associated with subsequent diagnosis: when cancer was initially mentioned, the odds of a cancer diagnosis were on average 5 times higher than when it was not initially mentioned (OR 4.90 [2.72 to 8.84], $P < 0.001$). The odds of appropriate referral were doubled when cancer was initially mentioned (OR 1.98 [1.10 to 3.57], $P = 0.002$).

In 60% of the instances of insufficient verbalization (51/85), cancer was diagnosed in the end. We performed sensitivity analyses using all 297 instances of initial verbalization, including the 85 instances of insufficient verbalization that had been dropped from the analyses above. First, we coded all 85 instances as 1 (cancer mentioned), which resulted in 65% (193/297) of verbalization instances where cancer was mentioned. This did not alter the relationship between first impressions and diagnosis (OR 4.74 [2.78 to 8.02], $P < 0.001$). When we coded the 85 instances as 0 (cancer not mentioned), this resulted in 36% (108/297) of verbalization instances where cancer was mentioned. The strength of the relationship between first impressions and diagnosis was reduced but remained significant: OR 2.38 [1.43 to 3.86], $P < 0.001$.

We detected some associations with physician experience. Specifically, physicians with more years in family medicine were less likely to mention cancer at the start (OR 0.96 [0.93 to 0.99], $P = 0.008$) and give it later as their working diagnosis (OR 0.97 [0.95 to 0.99], $P = 0.018$). The association with the inclusive measure of diagnosis (cancer as working or in differential) was borderline (OR 0.98 [0.96 to 1.00], $P = 0.051$), whereas no association between experience and appropriate referral was found (OR 0.99 [097 to 1.01]). As the study was not designed and powered to detect experience-related differences, these associations should be interpreted with caution.

The number of cancer-related questions mediated the relationship between first impressions and

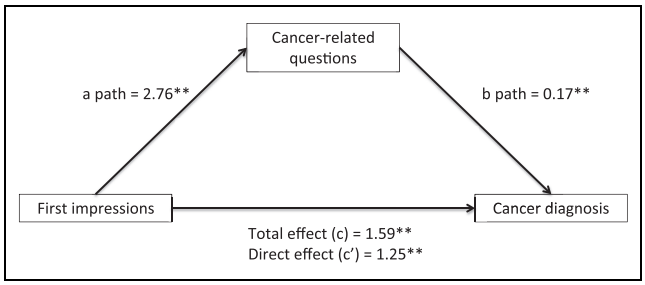

Figure 3 Mediation model and standardized regression coefficients for the relationship between first impressions and final cancer diagnosis as mediated by the number of cancer-related questions.

diagnosis (Figure 3). The standardized regression coefficients both between first impressions and cancer-related questions (2.76 [1.80 to 3.72]) and between cancer-related questions and final diagnosis (0.17 [0.08 to 0.26]) were significant. The standardized indirect effect was (2.76)*(0.17) = 0.47 [0.22 to 0.81] (confidence intervals estimated using bootstrapping with 10,000 samples) and explained 29% of the total effect (0.47/1.59). No associations were found between either first impressions or diagnosis with noncancer questions.

DISCUSSION

Using process-tracing methodologies, we elicited and measured a strong association between family physicians' first diagnostic impressions, as evident from their concurrent verbalizations, and their subsequent diagnosis and referral decisions in common presentations with subtle indications of cancer. Participants who, after reading a brief description about the patient and the presenting problem, and before requesting more information, did not explicitly acknowledge cancer as a diagnostic possibility were considerably less likely to diagnose it later and to refer the patient appropriately. A second presentation of the nonimproving patient increased the odds of appropriate referral but not of diagnosis. It is possible that considerations, such as patient satisfaction and regret avoidance, affect only referral decisions without influencing diagnosis.

When cancer was acknowledged explicitly as a possibility at the start of a consultation, it led to more cancer-related questions asked. This suggests that an initial concern about a possible cancer drove physicians to ask more questions about it, which enabled them to build a picture of cancer as a viable hypothesis and manage the patients accordingly. It is also likely that first impressions influenced the interpretation of the information subsequently gathered. The patients did not present with alarm symptoms for cancer but with subtle ones. If cancer was not considered at the start, symptoms such as fatigue or borderline anemia could well be dismissed or normalized.[24,25] The weak, negative association between physician experience and first impressions deserves further study, as it can have implications for medical education.

We took great care to minimize the likelihood that participants would perceive this as a cancer-related study, which would influence the behavior of interest: we included as many noncancer cases as cancer cases; 1 noncancer case included 2 consultations, like the cancer cases; and we asked participants to think aloud during that noncancer case, too. Our clinical cases were rich in detail and contained both diagnostically relevant and irrelevant information, meticulously developed to satisfy participants' information requests. This type of simulated interactive consultation on computer, where answers to physicians' questions are provided in real time, has been used in previous studies by the first author[19,26,27] and is the closest to a clinical consultation, short of using standardized patients. Nevertheless, there may still be a concern that medical scenarios presented in written form do not sufficiently represent real-life clinical encounters. Written scenarios used to study medical decision making intend to elicit and measure aspects of the decision-making processes that physicians use in real life. The elicited behaviors should not be taken as a reflection of real-life behavior ''but rather as strong predictors or proxies for such behavior.''[28] There is now substantial evidence that clinicians behave similarly both in written scenarios and approximate real-life situations.[28]

We employed the think aloud methodology to gain access to participants' initial hypotheses without having to ask them directly, as this would likely change their usual way of dealing with the cases. By using one cancer case as silent control (i.e., diagnosed without thinking aloud), we also ascertained that the think aloud methodology did not interfere with the outcome measures (diagnosis and decision) in a measurable way. Nevertheless, thinking aloud cannot reveal the entire contents of a participant's working memory; it is possible that some participants considered cancer as a possibility at the start but did not verbalize this. The sensitivity analyses that we performed by including instances of insufficient verbalization in the analyses go some way to tackle this limitation of concurrent verbal data, as they

demonstrate the strength of the relationship between initial verbalizations and final diagnoses. Whether physicians did not elicit the cancer hypothesis at the start or elicited it but considered it unlikely, the fact remains that they explored it less extensively than those who explicitly acknowledged it as a possibility.

The weak mediation effect of information search (number of cancer-related questions) suggests that a sizeable portion of the influence of initial impressions on final diagnosis is likely also to be mediated by the biased interpretation of information subsequently encountered. Because of incomplete verbalizations (participants were not prompted for their thoughts after each question), the concurrent verbal protocols from this study cannot be used for a systematic and unbiased exploration of information interpretation following first impressions. There is, however, substantial evidence for predecisional information distortion in the literature,[29,30] which suggests that, as a judgment, hypothesis or preference emerges, information gets distorted to support it (either bolstered or denigrated or both),[31–33] and that this happens with not only ambiguous but also diagnostic information.[34,35] In a series of experiments on diagnostic reasoning, in which students were taught the probabilistic relationships between fictitious chemicals and resulting health symptoms and were subsequently asked to identify the chemical that had caused the presenting symptoms, Rebitschek and colleagues[36] found a strong primacy effect: once an initial, leading hypothesis was established, it determined the final diagnosis, even in cases in which subsequent information was inconsistent. The authors attributed their findings to information distortion: participants changed the subjective value of the sequentially presented information to maintain coherence with their initial hypothesis, a phenomenon also supported by a number of other studies.[37–40]

Our study adds to the literature on first impressions, specifically in the area of diagnostic reasoning, using physicians as study participants diagnosing detailed clinical cases in an interactive manner that reflects real-life consultations. The study also establishes early diagnostic impressions as one reason for diagnostic delay in cases of possible cancers presenting with subtle symptoms and no red flags. Our findings suggest that attempts to reduce diagnostic delays should target the earliest stages of the diagnostic process. Hogarth[41] advises us to be critical of first impressions and to ask ourselves why our first idea might be wrong. Larrick[42] suggests considering the opposite, as a way of avoiding confirmation bias and reducing overconfidence. Rebitschek and colleagues[36] found a reduction of the primacy effect when participants assessed each symptom in relation to each competing, potential cause. Educators could consider how such strategies can be formally and systematically introduced to the medical curricula. Nevertheless, people who make decisions under time pressure, are multitasking, or are faced with too much and poorly structured information are less likely to question their first impressions. External decision aids may be more effective in such pressured and busy working environments. In 2 recent randomized controlled trials in the United Kingdom and Greece, using the same methodology for presenting the materials as in this study, physicians who simply read on their screen a list of differential diagnoses at the start of their interactive consultation with a computerized patient, and before gathering any further information (i.e., at the exact stage where we elicited first impressions in this study), were more accurate than controls across a range of diagnostic difficulty.[26,27] These 2 trials, conducted in 2 different countries with different medical training and health care systems, suggested that simple, external aids aimed at the initial stage of hypotheses generation can successfully influence first impressions and reduce diagnostic error.

REFERENCES

1. Baron J. Thinking and Deciding. 3rd ed. Cambridge: Cambridge University Press; 2000.

2. Asch SE. Forming impressions of personality. J Abnorm Soc Psych. 1946;41(3):258–90.

3. Crano WD. Primacy versus recency in retention of information and opinion change. J Soc Psychol. 1977;101(1):87–96.

4. Barrick MR, Swider BW, Stewart GL. Initial evaluations in the interview: relationships with subsequent interviewer evaluations and employment offers. J Appl Psychol. 2010;95(6):1163–1172.

5. Forgas JP. Can negative affect eliminate the power of first impressions? Affective influences on primacy and recency effects in impression formation. J Exp Soc Psychol. 2011;47(2):425–9.

6. Jackson A, Greene RL. Impression formation of tests: retrospective judgments of performance are higher when easier questions come first. Mem Cogn. 2014;42(8):1325–32.

7. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol. 1998;2(2):175–220.

8. Brownstein AL. Biased predecision processing. Psychol Bull. 2003;129(4):545–568.

9. Elstein AS, Shulman LS, Sprafka SA. Medical Problem Solving: An Analysis of Clinical Reasoning. Cambridge, MA: Harvard University Press; 1978.

10. Balla J, Heneghan C, Goyder C, Thompson M. Identifying early warning signs for diagnostic errors in primary care: a qualitative study. BMJ Open. 2012;2(5).

11. Bösner S, Hartel S, Diederich J, Baum E. Diagnosing headache in primary care: a qualitative study of Gps' approaches. Br J Gen Pract. 2014;64(626):E532–E537.

12. Nielsen TN, Hansen RP, Vedsted P. Symptom presentation in cancer patients in general practice. Ugeskr Laeger. 2010;172(41): 2827–31.

13. Ingebrigtsen SG, Scheel BI, Hart B, Thorsen T, Holtedahl K. Frequency of 'warning signs of cancer' in Norwegian general practice, with prospective recording of subsequent cancer. Fam Pract. 2013;30(2):153–60.

14. Scheel BI, Holtedahl K. Symptoms, signs, and tests: the general practitioner's comprehensive approach towards a cancer diagnosis. Scand J Prim Health Care. 2015;33(3):170–7.

15. Jones R, Latinovic R, Charlton J, Gulliford MC. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using general practice research database. BMJ. 2007;334(7602): 1040–4.

16. Singh H, Giardina TD, Meyer AND, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. JAMA Intern Med. 2013;173(6):418–25.

17. Kern KA. Medicolegal analysis of the delayed diagnosis of cancer in 338 cases in the United States. Arch Surg. 1994;129(4): 397–403.

18. Weber EU, Bockenholt U, Hilton DJ, Wallace B. Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. J Exp Psychol Learn Mem Cogn. 1993; 19(5):1151–64.

19. Kostopoulou O, Oudhoff J, Nath R, et al. Predictors of diagnostic accuracy and safe management in difficult diagnostic problems in family medicine. Med Decis Making. 2008;28(5):668–80.

20. Huber O, Wider R, Huber OW. Active information search and complete information presentation in naturalistic risky decision tasks. Acta Psychol. 1997;95:15–29.

21. Ericsson KA, Simon HA. Protocol Analysis: Verbal Reports as Data. Revised edition. Cambridge, MA: MIT Press; 1993.

22. Ericsson KA, Moxley JH. Thinking aloud protocols: concurrent verbalizations of thinking during performance on tasks involving decision making. In: Schulte-Mecklenbeck M, Kühberger A, Ranyard R, eds. A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide. Hove, UK: Psychology Press; 2011.

23. Woolley A, Kostopoulou O, Delaney BC. Can medical diagnosis benefit from ''unconscious thought''? Med Dec Making. 2016; 36:541–9.

24. Kostopoulou O, Devereaux-Walsh C, Delaney BC. Missing celiac disease in family medicine: the importance of hypothesis generation. Med Decis Making. 2009;29(3):282–90.

25. Kostopoulou O, Mousoulis C, Delaney BC. Information search and information distortion in the diagnosis of an ambiguous presentation. Judgm Decis Making. 2009;4(5):408–18.

26. Kostopoulou O, Rosen A, Round T, Wright E, Douiri A, Delaney B. Early diagnostic suggestions improve accuracy of Gps: a randomised controlled trial using computer-simulated patients. Br J Gen Pract. 2015;65(630):E49–54.

27. Kostopoulou O, Lionis C, Angelaki A, Ayis S, Durbaba S, Delaney BC. Early Diagnostic suggestions improve accuracy of family physicians: a randomized controlled trial in Greece. Fam Pract. 2015;32(3):323–8.

28. Evans SC, Roberts MC, Keeley JW, et al. Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies. Int J Clin Health Psychol. 2015;15(2):160–70.

29. Russo JE, Medvec VH, Meloy MG. The distortion of information during decisions. Organ Behav Hum Dec. 1996;66(1):102–10.

30. Dekay Ml. Predecisional information distortion and the self-fulfilling prophecy of early preferences in choice. Curr Dir Psychol Sci. 2015;24:405–11.

31. Blanchard SJ, Carlson KA, Meloy MG. Biased predecisional processing of leading and nonleading alternatives. Psychol Sci. 2014;25(3):812–6.

32. Dekay ML, Miller SA, Schley DR, Erford BM. Proleader and antitrailer information distortion and their effects on choice and postchoice memory. Organ Behav Hum Dec. 2014;125(2):134–50.

33. Nurek M, Kostopoulou O, Hagmayer Y. Predecisional information distortion in physicians' diagnostic judgments: strengthening a leading hypothesis or weakening its competitor? Judgm Decis Making. 2014;9(6):572–85.

34. Russo JE, Carlson KA, Meloy MG. Choosing an inferior alternative. Psychol Sci. 2006;17(10):899–904.

35. Kostopoulou O, Russo JE, Keenan G, Delaney BC, Douiri A. Information distortion in physicians' diagnostic judgments. Med Dec Making. 2012;32(6):831–9.

36. Rebitschek FG, Bocklisch F, Scholz A, Krems JF, Jahn G. Biased processing of ambiguous symptoms favors the initially leading hypothesis in sequential diagnostic reasoning. Exp Psychol 2015;62:287–305.

37. Russo JE, Carlson KA, Meloy MG, Yong K. The goal of consistency as a cause of information distortion. J Exp Psychol Gen. 2008; 137(3):456–70.

38. Glöckner A, Betsch T, Schindler N. Coherence shifts in probabilistic inference tasks. J Behav Decis Making. 2010;23(5):439–62.

39. Simon D, Holyoak KJ. Structural dynamics of cognition: from consistency theories to constraint satisfaction. Pers Soc Psychol Rev. 2002;6(4):283–94.

40. Hagmayer Y, Kostopoulou O. A probabilistic constraint satisfaction model of information distortion in diagnostic reasoning. In: Knauff M, Pauen M, Sebanz N, Wachsmuth I, eds. Proceedings of the 35th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society; 2013. p 531–6.

41. Hogarth RM. Educating Intuition. Chicago: University of Chicago Press; 2001.

42. Larrick RP. Debiasing. In: Koehler DJ, Harvey N, eds. Blackwell Handbook of Judgment and Decision Making. Oxford, UK: Blackwell; 2004. p 316–37.