Digital Healthcare Empowering Europeans
R. Cornet et al. (Eds.)

doi:10.3233/978-1-61499-512-8-90

Use of eye-tracking technology in clinical reasoning: a systematic review

Katherine BLONDON[a,b,1,2], Rolf WIPFLI[a,2] and Christian LOVIS[a,b]
[a] *University Hospitals of Geneva*
[b] *University of Geneva, Medical faculty*

Abstract. Achieving a better understanding of the clinical reasoning process is an important approach to improve patient management and patient safety. Although clinical psychologists have used talk-aloud or stimulated recall approaches, these methods have biases. Recently, researchers have been exploring eye-tracking technology to gain “live” insight into clinicians’ reasoning processes in certain fields of medicine (radiology, dermatology, etc.). We present a systematic review of eye-tracking literature used for clinical reasoning. We performed a literature search using the terms “eye” or “gaze tracking”, “clinical” or “diagnostic reasoning”, and “physician” in Pubmed, Embase, Psychinfo, Web of Science and ACM databases. Two investigators screened the abstracts, then full-text articles to select 10 pertinent studies. The studies evaluated medical decision making in four different medical domains using mostly experimental, observational approaches. A total of 208 participants were enrolled for the selected experiments. Paths for further studies are discussed that may extend the use of eye trackers in order to improve understanding of medical decision making.

Keywords. Gaze tracking, clinical reasoning, usability, medical expertise, systematic review

Introduction

Clinical reasoning processes have traditionally been studied by cognitive psychologists using think aloud or stimulated recall techniques. Both of these methods have intrinsic biases, however: thinking aloud takes additional time, and may add features, modifying the reasoning process. Stimulated recall takes place after the resolution, which may affect the recall or report of the reasoning process. Eye-tracking provides a potential for new insight into the reasoning process, hopefully without these intrinsic biases.

Eye-tracking measures gaze behaviour during task execution allowing researchers to gather data about the cues used during reasoning. Gaze behaviours are not all consciously remembered, and thus may allow researchers to detect unconscious cues that may affect the reasoning process. Eye-trackers measure gaze behaviour during task execution, visualize what areas on a screen are inspected, and thus provide clues on what information was included in the decision making process. They are used in educational or cognitive psychology to understand expert performance for example in sports, aviation, or car driving[1]. They provide quantitative measures including gaze

[1] katherine.blondon@hcuge.ch
[2] first co-authors

coordinates, dwelling time on regions of interests (ROI), and number of inspections per ROI that can be interpreted accordingly[2]. Most of the eye-trackers used today are non-intrusive infrared reflectance cameras that measure corneal reflections of emitted infrared light. The eye tracking method may be accompanied by the think aloud method where participants explain their reasoning either during task execution or after when reviewing their scan path video[3]. In the medical domain, eye-tracking has been used in clinical trials with patients and with health professionals. For this paper, we present a systematic review on the use of eye-tracking among health care professionals to understand their clinical reasoning and work processes.

1. Methods

For this systematic review, our goal was to report all published studies that used eye-tracking to explore clinical reasoning. We performed a literature search of PubMed, PsycInfo, ACM, and Web of Science using the keywords or MeSH terms "eye" or "gaze tracking", "clinical" or "diagnostic reasoning", and "physician." The search included published papers in English, from Jan 2000 to Nov 2014. We used review papers to identify other papers or authors in the references. We limited the search to articles published in year 2000 or later, due to significant changes in eye-tracking technology.

Two reviewers screened the selected studies independently, first based on titles, then abstracts, and full papers. We included studies that enrolled on physicians as participants and that used eye-tracking to understand the reasoning process. We included all reasoning processes, from the diagnostic reasoning to errors, management and decision-making processes as outcomes. We excluded articles that used eye-tracking to understand visual analytics without implications for the reasoning process, such as a comparison between two eye-tracking approaches. Reviewers had to reach a consensus for inclusion or exclusion of articles, and discordances were discussed. We extracted the following data from the included studies: medical field, study population, intervention, comparators, outcome, and study design according to the PICOS approach. For comparators, we used the principal methods used for gaze data analysis.

2. Results

Our initial database search yielded 132 abstracts with another 37 abstracts from other resources, mainly from reference lists in review articles (n=165 without duplicates). We excluded 155 articles that did not use eye-tracking to explore clinical reasoning, which resulted in the inclusion of 10 articles in our systematic review.

We summarize the results of our qualitative research in Table 1. Nine out of ten articles were published in 2012 or later. Overall, the included studies had a total of 208 participants. Two studies had a very small sample size[11,13] one of which had inconclusive results in their comparison of 3 levels of expertise.

These publications used eye tracking in four medical domains in radiology, dermatology, neurology or use of electronic medical records. Most of the studies sought to understand abnormal pattern identification while reading of static medical images like mammography[10,11,13] or skin lesions[7,12]. Two studies used dynamic presentations of videos: a looped 3-D CT colonography[8] and video patient cases in paediatric neurology[4]. Finally, we identified two articles that studied the use electronic medical records (EMR) for handoffs[5], and for decision-making[6].

A large majority of studies were experimental and observational and conducted in a laboratory setting. We found one usability test that aimed to test a prototype[6], and one in-situ study where the eye-tracker was installed at the participant's work place[9].

The majority of studies were designed to sample two (or three) populations with different levels of medical expertise. Their aim was to learn about expert clinical reasoning by comparing the behaviours of experts to that of novices[4,6-8,10-13]. Three studies even enrolled lay people as a comparison group[4,7,12].

The raw data provided by eye-trackers were fixations and saccade length. The analyses, however, differed between studies. Some studies used eye gaze visualizations like heat maps to qualitatively learn about screen reading through patterns, such as skipped passages of the EMR report[6]. Most of the studies defined regions of interest and calculated fixation times or fixation numbers to test their hypotheses[4,5,8,10-12]. One paper chose to measure how often regions of interest were reinspected after a first inspection[12]. In order to better understand the underlying clinical reasoning process, some studies also included a concurrent think aloud approach[4,5,7] where physicians explained their actions while performing them.

Overall, the studies show that expert physicians have a better diagnostic accuracy, need less time to first spot the abnormal pattern and subsequently look more often at the patterns. However, experts are not necessarily more efficient in the diagnostic task[4,10]. Rather, they generate more theories and possible diagnosis in the given time[4].

3. Discussion

Our literature search identified 10 studies using eye-tracking for clinical reasoning, Most articles were published in 2012 and later, suggesting that the use of eye-tracking technology to understand clinical reasoning is a relatively novel, current research topic. This research has been conducted in four different medical domains, two of which require strong visual analytical skills: radiology (X-rays and other imaging) and dermatology. It was therefore not unexpected to find these domains in eye-tracking research. Other studies about expertise during X-ray interpretation have been published, but were not included if they did not include (1) physicians and (2) an exploration of clinical reasoning, i.e.[14]. We expected to find studies in other domains, such as pathology (high dependence on visual analytics) or in internal medicine, which has been well studied in cognitive psychology.

The use of eye-tracking technology to explore the progress note reading is a blend between usability research and cognitive psychology. The design or data visualization implications can potentially be integrated into future EMR designs, thus hopefully improving both usability and clinical reasoning. This is illustrated in the study on antibiotic prescription on an EMR, which used usability testing of the prototype to gain insight into different approaches by level of expertise. With the growing implementation of EMRs, future studies on EMR usability should consider including an exploration of the reasoning process to guide future designs.

Table 1. Articles with study characteristics that were included in the qualitative analysis

Reference	Domain	Study type	Participants	Task	Analysis	Results
Balslev et al. 2012[4]	Pediatric neurology	Experimental observational study	N=43 (15 medical students, 16 residents, 12 experts)	Diagnose with 4 patient video cases	Diagnostic accuracy, eye gazes, think aloud statements	Experts: more time on relevant ROI, not faster in executing tasks, more theory building and evaluation, less data exploration
Brown et al. 2014[5]	Hospital medicine	Experimental observational study	N=10 (hospitalists)	Hand-off 3 patients after reading the progress notes in an electronic medical record	Eye gazes, transcription of verbalized handoff	Most glances in sections impression and plan (67%), laboratory results (8%) and medication profile (7%)
Forsman et al. 2013[6]	Intensive care	Multimethod: ethnographic research, participatory design, prototyping, usability test	For usability test: N=12 (6 ICU specialists, 2 usability experts, 4 residents)	For usability test: 15 navigation tasks + 8 clinical tasks using an interface for antibiotics prescription	For usability test: eye gazes on different sections of the interface	For usability test: resident physicians more often inspected graphical representations compared to specialists).
Li et al. 2012[7]	Dermatology	Experimental observational, master-apprentice model	N=28 (11 attending dermatologists + 4 residents compared to 13 undergraduate lay people)	Experts: diagnose while thinking aloud; novices: act as explaining to a physician over phone	Hierarchical dynamic model of gaze data; analysis of transcriptions	Successful application of hierarchical dynamic model which is coupled with transcribed verbal reports
Mallet et al. 2014[8]	Colonography	Experimental, observational study	N=65 (27 experienced, 38 inexperienced radiologists)	Identify polyps in 23 fly-through (no stopping for further inspection) 3-D CT colonographies	Eye gazes during looped video play	Experts performance: higher identification rate, shorter time to first pursuit, but gaze patterns similar to inexperienced participants
Nielson et al. 2013[9]	Emergency department	Observational in-situ study	N=14 (physicians)	Perform clinical duties on usual work station equipped with a remote eye tracker	Eye gazes	Per shift: 21.7 times laboratory displayed; per display of laboratory results: 13.9s fixation of values
Nodine et al. 2002[10]	Mammography	Experimental, observational study	N=9 (6 radiology trainees, 3 mammographers)	Mark location of lesion and confidence level when reading 40 cases (20 positives, 20 negatives)	Diagnostic accuracy, fixation on true lesions	Same decision time, trainees had more false positives when looking at normal cases for more than 25 s; experts had more inspections on ROI
Tourassi et al. 2013[11]	Mammography	Experimental, observational study	N = 6 (3 experts + 3 4th year residents)	Detect lesions in 20 mammographies	Diagnostic accuracy (predicting behavior with machine learning)	85.7% diagnostic accuracy for malign masses and 96.7% for benign masses
Vaidyanathan et al. 2014[12]	Dermatology	Experimental observational, master-apprentice model	N=29 (16 dermatologists, 13 undergraduates with no dermatology training)	Examine and describe 42 dermatological images	Recurrence quantification analysis of gaze pattern	Experts: less recurrent fixations, less repetitions of short inspection sequences, but revisited same regions after longer time intervals
Voisin et al. 2013[13]	Mammography	Experimental, observational study	N=6 (2 experts, 2 intermediates, 2 novices)	Assess anomaly probability in 40 mammograms (20 positives, 20 negatives) and difficulty of task	Gaze pattern, confidence rating	No consistent results for gaze patterns; longer reading time and higher number of fixations result in less confidence in answers

Using a comparative approach between two populations with different levels of expertise is common, although there may be significant variability among experts. Research in cognitive psychology has shown that experts tend have little overlap in their reasoning process, because of "shortcuts" acquired over time[15]. Therefore, differentiating the expert-novice characteristics due to level of expertise may be affected by inter-individual variations as well.

Finally, the included studies focus mainly on the diagnostic reasoning process, except for the two studies on EMR data. Although diagnostic reasoning is an important component of patient management, further studies should also investigate other reasoning processes, such as handoffs and decision-making. These components may seem less visual per se, but can be studied through EMR use.

There is evidence of interest in using eye tracking as a new approach to clinical reasoning, which appears to be a novel, current area of research. A deeper understanding of clinical reasoning can help improve medical education. Further studies are needed to address potential case specificity issues, and to confirm the results of the smaller sized studies. Future studies are also needed to compare results from eye-tracking studies to traditional approaches (think-aloud, stimulated recall).

References

1. Gegenfurtner A, Lehtinen E, Säljö R. Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains, Educ Psychol Review 2011;23:523-52
2. Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H, Van de Weijer J. Eye tracking: A comprehensive guide to methods and measures. Oxford UK: Oxford University Press; 2011.
3. Olsen A, Smolentzov L, Strandvall R. Comparing different eye tracking cues when using the retrospective think aloud method in usability testing. Proceedings of the 24th BCS Interaction Specialist Group Conference 2010:45-53.
4. Balslev T, Jarodzka H, Holmqvist, de Gravee W, Muijtjens AMM, et al. Visual expertise in paediatric neurology. Eur J Paediatr Neurol 2012;16:161-6.
5. Brown PJ, Marquard JL, Amster B, Romoser M, Friderici J, Goff S, et al. What do physicians read (and ignore) in electronic progress notes? Appl Clin Inform 2014;5:430-44.
6. J Forsman, N Anani, A Eghdam, Falkenhav M, Koch S. Integrated information visualization to support decision making for use of antibiotics in intensive care: design and usability evaluation. Inform Health Soc Care 2013;38:330-53.
7. Li R, Pelz J, Shi P, Alm CO, Haake A. Learning Eye Movement Patterns for Characterization of Perceptual Expertise. Proceedings of the Symposium on Eye Tracking Research and Applications. NewYork NY: ACM; 2012;393-6.
8. Mallett S, Phillips P, Fanshawe TR, Helbren E, Boone D, Gale A, et al. Tracking Eye Gaze during Interpretation of Endoluminal Three-dimensional CT Colonography: Visual Perception of Experienced and Inexperienced Readers. Radiology 2014;273(3):783-92.
9. Nielson JA, Mamidala RN, Khan J. In-situ eye-tracking of emergency physician result review. Stud Health Technol Inform 2013;192:1156.
10. Nodine CF, Mello-Thoms C, Kundel HL, Weinstein SP. Time course of perception and decision making during mammographic interpretation. American journal of roentgenology 2002;179:917-23.
11. Tourassi G, Voisin S, Paquit V, Krupinski E. Investigating the link between radiologists' gaze, diagnostic decision, and image content. J Am Med Inform Assoc 2013;20:1067-75.
12. Vaidyanathan P, Pelz J, Alm C, Pengcheng S, Haake A. Recurrence Quantification Analysis Reveals Eye-movement Behavior Differences Between Experts and Novices. Proceedings of the Symposium on Eye Tracking Research and Applications. New York NY: ACM;2014:303-6.
13. Voisin S, Pinto F, Xu S, Morin-Ducote G, Hudson K, Tourassi GD. Investigating the association of eye gaze pattern and diagnostic error in mammography. Proceedings of SPIE 2013;8673:867302-8.
14. Anderson B, Shyu CR. Studying Visual Behaviors from Multiple Eye Tracking Features Across Levels of Information Representation. AMIA Annual Symposium Proceedings 2011:72.
15. Nendaz MR, Gut AM, Perrier A, Reuille O, Louis-Simonet M, Junod AF, et al. Degree of concurrency among experts in data collection and diagnostic hypothesis generation during clinical encounters. Med Educ 2004; 38, 25-31.