SPIE Medical Imaging 50th anniversary: history of the Image Perception, Observer Performance, and Technology Assessment Conference

Elizabeth A. Krupinski [a],[*] **and Harold L. Kundel** [b]

aEmory University, Department of Radiology and Imaging Sciences, Atlanta, Georgia, United States

bUniversity of Pennsylvania, Department of Radiology, Atrium, Philadelphia, Pennsylvania, United States

Abstract. As SPIE Medical Imaging celebrates its 50th anniversary, we reflect on the history of the Image Perception, Observer Performance, and Technology Assessment Conference and its importance within the SPIE Medical Imaging Symposium and the greater medical imaging community. © 2022 *Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JMI.9.S1 .012202]

1. Introduction

Just remember that the things that you put in your head are there forever, he said.
You might want to think about that.
You forget some things don't you?
Yes. You forget what you want to remember and you remember what you want to forget.
  Cormac McCarthy – *The Road.*

A few radiologists and imaging scientists were always interested in human factors as a source of error and variation in the interpretation of images.[1,2] However, the emphasis of the larger imaging community was on physical and engineering approaches to improving diagnostic accuracy. Make the images better and performance will improve, as opposed to making the humans better and improve performance. Between the physical and psychological approaches stood a common need for a reliable method for evaluating performance.

  As early as 1971 at the very first SPIE Medical Imaging-related meeting (Quantitative Imagery in the Biomedical Sciences, Volume 26 of the Proceedings), image perception was present, with Arne Troelstra giving a paper titled "The Eye as a Detector"[3] and Ziskin, Shea, Kundel, and Revesz presenting on "Accuracy of Radiologists' Decision Making."[4] Then, at the 1993 SPIE Medical Imaging Image Processing Conference, chaired by Murray Lowe, PhD (George Washington University), Kundel presented the keynote paper "Perception and Representation of Medical Images."[5] He stressed the human component of the output of image interpretation. There were also a number of other papers about various components of error, variation, and diagnostic accuracy in the Physics of Medical Imaging and Image Processing conferences, and Samuel Dwyer, PhD (UCLA), chairman of the Medical Imaging Conference, asked Kundel to pull them together in 1994 as a dedicated Perception Conference. Tony Franken, MD, from the University of Iowa, and Hal Kundel had been sponsoring an informal meeting of "Perceptionists" at the annual Radiological Society of North America (RSNA) meeting, so the mailing list of the "Perceptionist" meeting was added to the regular request for papers sent out by SPIE to invite presentations at the newly formed Image Perception Conference.

---

*Address all correspondence to Elizabeth A. Krupinski, ekrupin@emory.edu

## 1.2 *Medical Image Perception through the Years*

The first meeting of the "Image Perception" conference in 1994 in Newport Beach, California (CA) was international in flavor with participants from the United States, United Kingdom, Netherlands, France, Canada, Russia, and Mexico. The objective of the conference was to discuss core principles and techniques unique to medical imaging. The abstracts of the presentations (SPIE Proceedings Volume 2166) fit roughly into the following categories:

(1) Modeling visual tasks: from noise to discrimination (6 abstracts)
(2) Specific image properties that affect observer performance (5 abstracts)
(3) Visual search and eye tracking (3 abstracts)
(4) Human factors that influence performance: workload, expertise (2 abstracts)
(5) Evaluation of observer performance: 2AFC, receiver operating characteristic, Bayesian, etc. (8 abstracts).

It was clear from the initial and subsequent conferences that metrics of observer performance were of interest and, beyond the evaluation of various aspects of human perception, had practical value in the evaluation of emerging imaging technology. As a consequence, for the 1999 conference in San Diego, CA (SPIE Proceedings Volume 3663), the name of the conference was changed to "Image Perception and Performance" and in 2002 (SPIE Proceedings Volume 4686) the name was again changed to "Image Perception, Observer Performance, and Technology Assessment." With each name change, there was an increase in the number of abstracts submitted, papers presented, and the number of papers in the proceedings (Fig. 1). The ebb and flow of submissions and accepted presentations over the years is likely a multi-factorial phenomenon and likely includes, but is not limited to, such factors as changes in imaging technology, influx of attendees from different fields (e.g., psychology, mathematics, and other clinical specialties than radiology), fluctuations in funding, and other conferences incorporating similar topics into programs. It should also be noted that, to some degree, the CAD and Digital Pathology conferences are spin-offs of the Image Perception and the other Medical Imaging conferences, potentially attracting some authors.

In 2021, the growing interest in artificial intelligence (AI) was clearly reflected in topics covered at the meeting (which was virtual due to the COVID-19 pandemic):

(1) AI, machine, and deep learning (11 abstracts)
(2) Modeling visual signal detection and task performance (11 abstracts)
(3) Observer performance assessment (4 abstracts)
(4) Visual search and perceptual processes (3 abstracts)
(5) Technology evaluation (1 abstract).
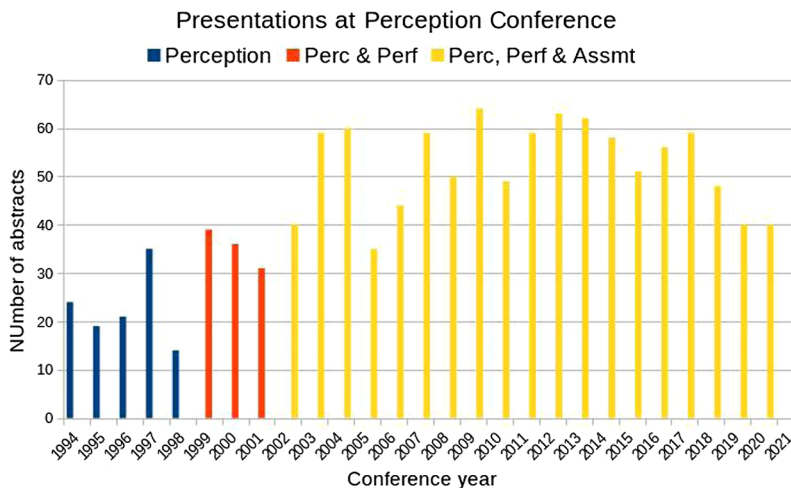
**Fig. 1** Number of presentations at the Perception Conference by year.
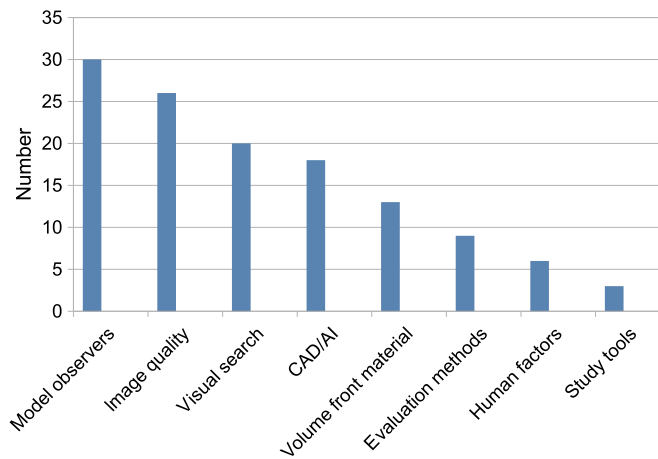
Fig. 2 Topic areas within which the top five downloaded papers per year fall. CAD, computer aided detection; AI, artificial intelligence. "Volume front material" refers to the volume description and table of contents.

An analysis of the top five downloaded conference proceedings papers for each year (Fig. 2) of the Perception Conference since its inception shows that the top five topics are model observers (30), image quality assessment (26), visual search and perception (20), computer-based detection methods (18) (e.g., CAD/AI), and observer performance assessment (15). The most downloaded paper was "Hologram stability evaluation for Microsoft HoloLens" by Vassalo et al. in 2017[6] at 1002 (as of May 2021), a technology evaluation paper, perhaps reflecting a growing interest in augmented and virtual reality environments and tools for viewing medical images clinically and/or during training. Aside from front matter, the top downloads in each topic are summarized in Table 1.

Table 1 Most downloaded papers per topic area.

Title	Authors	Year	Volume	Topic
Hologram stability evaluation for Microsoft HoloLens	R. Vassallo, A. Rankin, E. C. S. Chen, and T. M. Peters	2017	10136	Technology evaluation
Nature of expertise in searching mammograms for breast masses	C. F. Nodine, H. L. Kundel, S. C. Lauver, and L. C. Toto	1996	2712	Visual search
A deep learning model observer for use in alterative forced choice virtual clinical trials	M. Alnowami, G. Mills, M. Awis, P. Elangovanr, M. Patel, M. Halling-Brown, K. C. Young, D. R. Dance, and K. Wells	2018	10577	CAD/AI
Task-based evaluation of deep image super-resolution in medical imaging	V. A. Kelkar, X. Zhang, J. Granstedt, H. Li, and M. A. Anastasio	2021	11599	Image quality
Stabilized estimates of Hotelling-observer detection performance in patient-structured noise	H. H. Barrett, C. K. Abbey, B. D. Gallas, and M. P. Eckstein	1998	3340	Model observers
Validation of no-reference image quality index for the assessment of digital mammographic images	H. C. R. De Oliveira, B. Barufaldi, L. R. Borges, S. Dabarda, P. R. Nakic, A. D. A. Maidment, H. Schiabel, and M. A. C. Vieira	2016	9787	Evaluation methods
Does reader visual fatigue impact interpretation accuracy?	E. A. Krupinski and K. S. Berbaum	2010	7627	Human factors
Producing lesions for hybrid mammograms: extracted tumors and simulated microcalcifications	A. E. Burgess and S. Chakraborty	1999	3663	Study tools
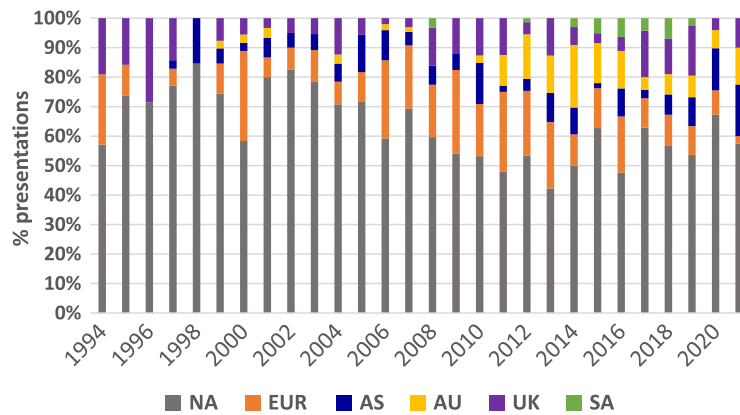
Fig. 3 First/corresponding author country/region. NA, North America (USA, Canada, Mexico); EUR, Europe, including Russia and Middle East; AS, Asia; AU, Australia; UK, United Kingdom and Northern Ireland; and SA, South America.

Like other conferences, the Perception Conference hosted a variety of workshops, including a couple that brought in eye-tracking technologies to demonstrate live how visual search patterns can be recorded and analyzed. In many respects, these on-site demonstrations led to the implementation of a dedicated Image Perception Laboratory at the annual RSNA meeting where researchers can run time-limited perception studies using volunteer RSNA attendees as observers.

1.3 Impact of the Perception Conference on Attendees and the Field

The steady growth in the Perception Conference over the years has fostered the careers of numerous investigators both in academia and industry. The organizing committee grew from a single person to a multidisciplinary committee (e.g., radiologists, psychologists, engineers, mathematicians, and physicists) headed by two co-chairs. Many committee members have gone on to serve on a variety of SPIE committees including the Medical Imaging Symposium Committee, Awards Committee, and Publications Committee, to name a few. The conference itself has always been welcoming to student/trainee presentations (both oral and poster), and for long-time attendees it has been wonderful to watch them evolve into full-time faculty members with their own students doing the presenting. For many, SPIE was their very first major conference presentation and exposure to a community of international researchers (Fig. 3) dedicated to medical image perception.

In 2019, SPIE created the Harrison H. Barrett Award in Medical Imaging, recognizing the impact Harry has made on the field. Harry and his students have been a part of the Perception Conference from the start, contributing a number of papers especially in the model observer area. The recipients of this award have all been key to the success of the Perception Conference, through their innovative research, leadership, and contributions to the meeting and the field. The winners to date have been Harrison Barrett, Arthur Burgess, Charles Metz, and Robert Wagner in 2019, Harold Kundel in 2020, and Kevin Berbaum in 2021. Research by these pioneers and their trainees fills the Perception Conference proceedings and it is well worth everyone's time to periodically go back and review some of these seminal papers.

1.4 Opportunities for the Image Perception Conference

It's difficult to make predictions, especially about the future.
Attributed to Niels Bohr, Mark Twain, Yogi Berra, and others (we prefer Yogi Berra)

Neural networks and deep learning have appeared in abstracts since computer aided diagnosis started becoming competitive with plain human interpretation, mainly of mammograms.

Recently, AI, which for the most part is machine learning, has started to appear as a subject of interest. A recent American College of Radiology bulletin features an article about AI,[7] in which the enthusiasm for AI is tempered by noting that there are biases in the data that are used to train AI algorithms and that accuracy may vary with different populations.

In addition to the traditional topics, the 2022 SPIE "Call for Papers" included a request for papers on the "Impact of AI technologies on provider clinical decisions & workflow." It also includes a notice as follows: "New for 2022: Joint Session with the Computer-Aided Detection Conference: 'Translation of CAD-AI methods to clinical practice; are we there yet?' We invite papers on comparisons in performance between CAD-AI and humans, retrospective studies comparing CAD-AI output to original clinical decision, reader studies, and studies of CAD-AI in clinical practice."

Defining the accuracy of AI has one of the main difficulties of ROC analysis—namely establishing "the truth." Validating AI will have the same difficulty. Here is a good opportunity to deal with methods for defining and measuring accuracy (whatever that is) and for building models that characterize AI based on human perception. Assessing the impact on observer performance, development of expertise, reader fatigue, and the host of other topics the Perception Conference has covered over the years comes to bear with AI implementation and evaluation as well.

"Image Perception, Observer Performance, and Technology Assessment" may incorporate yet another focus area over time, but it will always remain relevant to medical imaging. Technology development does not stand still. AI is only a small part of what medical imaging will become, with technologies such as photon-counting computed tomography, digital pathology, telemedicine, augmented and virtual reality, true 3D displays, and what has yet to be thought of emerging daily. At least for now (and likely forever), the human operator/user is required in the medical decision-making and patient care process, thus research focused on how these users interact with and are impacted by the technologies they use and the information provided to them will remain as vital and relevant as it has for at least the past 50 years.

Disclosures

No conflicts of interests, financial, or otherwise, are declared by the authors.

References

1. L. H. Garland, "On the scientific evaluation of diagnostic procedures," *Radiology* **52**(3), 309–328 (1949).
2. W. J. Tuddenham, "Roentgen image perception—a personal survey of the problem," *Radiol. Clin. N. Am.* **7**(3), 499–501 (2010).
3. A. Troelstra, "The eye as a detector," *Proc. SPIE* **0026**, 11–14 (1971).
4. M. C. Ziskin et al., "Accuracy of radiologists' decision making," *Proc. SPIE* **0026**, 41–48 (1971).
5. H. L. Kundel, "Perception and representation of medical images," *Proc. SPIE* **1898**, 1–11 (1993).
6. R. Vasallo et al., "Hologram stability evaluation for Microsoft HoloLens," *Proc. SPIE* **10136**, 1013614 (2017).
7. C. Hudnall, "Thinking about AI?" *Am. Coll. Radiol. Bull.* **6**, 9–12 (2021).

Elizabeth A. Krupinski, PhD, is Professor and Vice Chair for Research at Emory University in the Departments of Radiology & Imaging Sciences, Psychology, and Medical Informatics. She received her BA degree from Cornell, MA from Montclair State, and PhD from Temple, all in Experimental Psychology. Her interests are in medical image perception, observer performance, medical decision making, and human factors. She is past chair of the SPIE Medical Imaging Conference.

Harold L. Kundel, MD, a retired diagnostic radiologist, is currently Professor Emeritus of Radiology at the Perelman School of Medicine of the University of Pennsylvania. His research interests were centered on understanding image reader performance, using human performance for the evaluation of imaging technology and the development of decision aids for radiologists. He was among the first investigators in radiology to study visual search using gaze tracking and to use ROC analysis for technology evaluation.