Gníomhaireacht Bainistíochta an Chisteáin Náisiúnta
National Treasury Management Agency

An Ghníomhaireacht um Éilimh ar an Stát
State Claims Agency

What 20 Years of Research on Diagnosis Have Taught Me (that others should know)

Dr Olga Kostopoulou
Applied Cognitive Psychologist,
Imperial College London

#SCALearning25

IMPERIAL

What 20 years of research on diagnosis have taught me (that others should know)

Olga Kostopoulou, PhD
Reader in Medical Decision Making

State Claims Agency - Clinical Risk Conference, Dublin 1/10/25

Diagnostic error

Overlooked in patient safety?

- *“Diagnosis apparently gets overlooked in most efforts to ensure quality and safety”* (Leape et al., *JAMA* 2002)
- IoM report 1998 “To Err Is Human”: the term *medication error* was mentioned 70 times, while *diagnostic error* appeared only twice.
- IoM report 2015 “Improving Diagnosis in Health Care”: *“Most people will experience at least one diagnostic error in their lifetime”*.

Diagnostic error

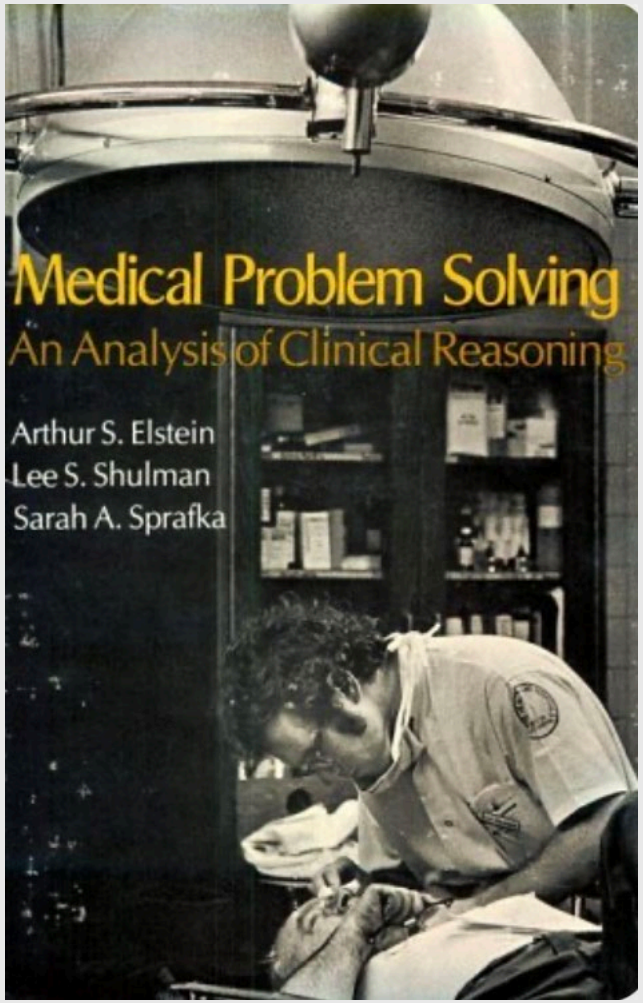
A Cognition-System interplay

- Cognitive causes are common (Gandhi et al., 2006; Lyratzopoulos, et al., 2015)
- Cognition is relevant to essentially every aspect of diagnosis and medical practice generally (Croskerry & Graber, 2025)

- E.g.,
  - Failing to consider a broad enough differential
  - Focusing on one diagnosis too early
  - Not asking enough questions
  - Dismissing things that don't 'fit'
- *Not specific to medicine*

Ample evidence that

- People formulate hypotheses quickly and with little information (Baron, 2000)
- Automatically, even if they are instructed not to do so. *A psychological necessity*.
- Clinicians generate diagnostic hypotheses within seconds (Elstein et al. 1978)
- Diagnostic hypotheses guide information search and interpretation



Diagnosis as a cognitive process

- Where do diagnostic hypotheses come from?

Memory.

- How do they get activated?

From observations (e.g., symptoms, signs, risk factors, FHx, GP's referral letter, etc.)

- Which hypotheses will be activated first?

Those that are more strongly linked with the observations (common diagnoses, serious diagnoses)

- People will generate only a small subset of hypotheses - no more than 4.

First impressions matter

And clinicians know it

- Physicians acknowledge their importance in interviews (e.g., Balla et al. 2012; Bösner et al. 2014)
- the first impression ‘…*is* ***crucial*** *regarding the decision on what I will do with this patient and on what further diagnostic procedures I will initiate.” (*Bösner et al. 2014)

Initial hypotheses are compelling

- Rested working memory
- More attention, better encoding
- Disproportionate weight (‘primacy effect’)

- Crowd out competitors

- The lens through which incoming information is evaluated

First impressions matter

How much?

Kostopoulou et al. (2017) *Medical Decision Making*, **37**(1), 9-16.

ORIGINAL ARTICLE

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

- Are GPs' initial diagnostic impressions associated with subsequent diagnosis and referral decisions?
- Study aimed to elicit and measure this association

Design

- 90 UK GPs
- 6 detailed patient scenarios (3 possible cancers)
  - Repeated consultations
- Interactive via phone communication with researcher

Henry Strauss

Patient information

- AGE: 61 years old
- ETHNICITY: Caucasian
- HEIGHT: 175 cm (5ft 7.7")
- WEIGHT: 72 kg (11.3 stones) (BMI 23.5, measured 6 months ago)
- SMOKING STATUS: Never smoked
- ALCOHOL: Occasional
- LAST BLOOD PRESSURE READING: 138/82, taken 6 months ago
- PAST MEDICAL HISTORY: Irritable Bowel Syndrome (2000) - coeliac screening negative. Hypertension (2006).
- MEDICATIONS: Lisinopril 5 mg od
- OCCUPATION: Administrative officer in the local Council, now retired
- LAST CONSULTATION: 6 months ago for routine review of hypertension
- APPEARANCE: Nothing of note.

Presenting complaint

Doctor, I am having problems with my bowels again. I've been having hard stools for about a month now. I go to the loo only once in 4-5 days.

Confirm you have read the Presenting Complaint

Henry Strauss

Patient information

- AGE: 61 years old
- ETHNICITY: Caucasian
- HEIGHT: 175 cm (5ft 7.7")
- WEIGHT: 72 kg (11.3 stones) (BMI 23.5, measured 6 months ago)
- SMOKING STATUS: Never smoked
- ALCOHOL: Occasional
- LAST BLOOD PRESSURE READING: 138/82, taken 6 months ago
- PAST MEDICAL HISTORY: Irritable Bowel Syndrome (2000) - coeliac screening negative. Hypertension (2006).
- MEDICATIONS: Lisinopril 5 mg od
- OCCUPATION: Administrative officer in the local Council, now retired
- LAST CONSULTATION: 6 months ago for routine review of hypertension
- APPEARANCE: Nothing of note.

Hb 13.0 g/dL (13.5-18)*
MCV 76 fL (76-100)
WCC 8.1 x 10^9 (4.0-11.0)
Neutrophils 6.2 x 10^9 (2.0-7.5)
Lymphocytes 2.0 x 10^9 (1.3-3.5)
Monocytes 0.2 x 10^9 (0.2-0.8)
Eosinophils 0.3 x 10^9 (0-0.4)
Platelets 295 x 10^9 (150-450)

Two weeks later, Henry Strauss comes back to see you.

Patient information

- AGE: 61 years old
- ETHNICITY: Caucasian
- HEIGHT: 175 cm (5ft 7.7")
- WEIGHT: 72 kg (11.3 stones) (BMI 23.5 measured 6.5 months ago)
- SMOKING STATUS: Never smoked
- ALCOHOL: Occasional
- LAST BP: 132/78, taken 2 weeks ago
- PAST MEDICAL HISTORY: Irritable Bowel Syndrome (2000) - coeliac screening negative. Hypertension (2006).
- MEDICATION: Lisinopril 5 mg od
- OCCUPATION: Administrative officer in the local Council, now retired
- LAST CONSULTATION: 2 weeks ago with constipation for one month.
- APPEARANCE: Nothing of note.

Presenting complaint

Good morning doctor. Do you remember that I came to see you two weeks ago with constipation? I did everything you told me but it hasn't gone away. I'm not sure what else to do.

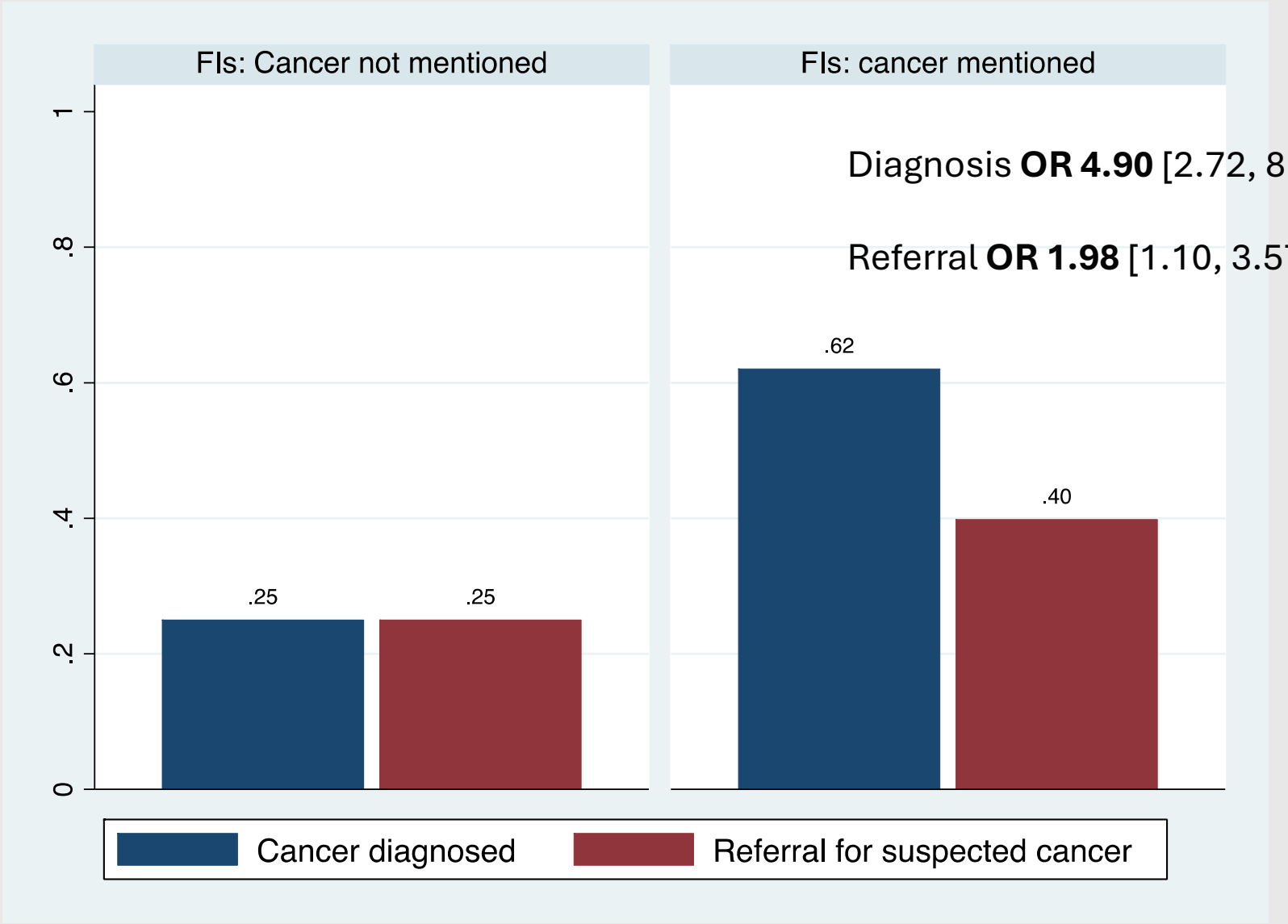
Confirm you have read the Presenting Complaint

Design

- Think aloud
- Coded and analysis of initial utterances
  - Cancer mentioned? Yes (1), No (0)

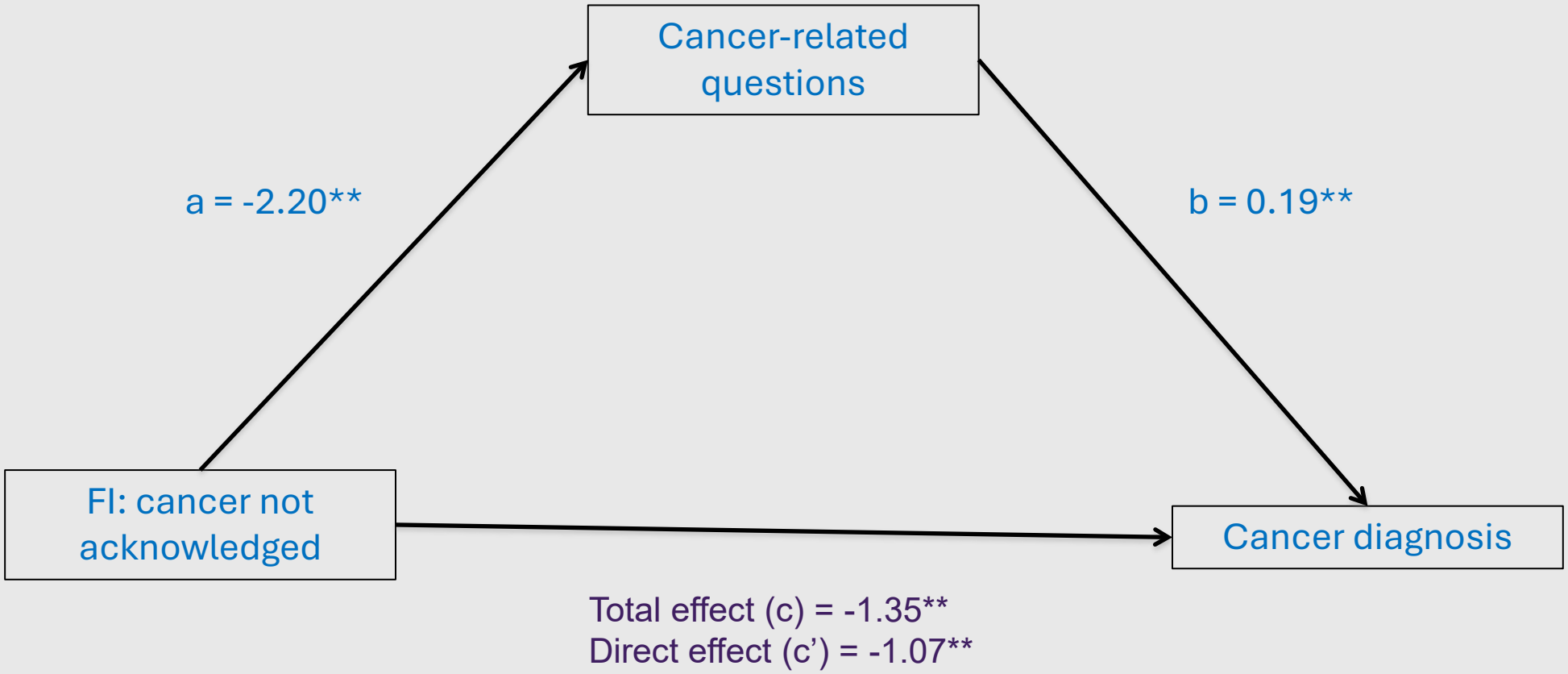
Early, explicit acknowledgement of cancer possibility

Significantly increased cancer diagnosis and referral odds



First impressions drive information gathering

A mediation model



explained 28% of the total effect

First impressions drive interpretation

Predecisional Information Distortion (PID)

- We tend to evaluate information in a biased way in order to:

1. Consolidate a decision already made (post-decision)
2. Strengthen an emerging decision (pre-decision)

- Incoming information is evaluated as supportive of an emerging judgement, preference or currently leading hypothesis (Russo et al. 1998; DeKay, 2015)
- Innate drive to attain and maintain ‘cognitive consistency’ (Russo et al. 2008; Simon et al. 2004).

PID in clinical diagnosis

Medical Decision Making (2012) **32**(6), 831-839

Information Distortion in Physicians' Diagnostic Judgments

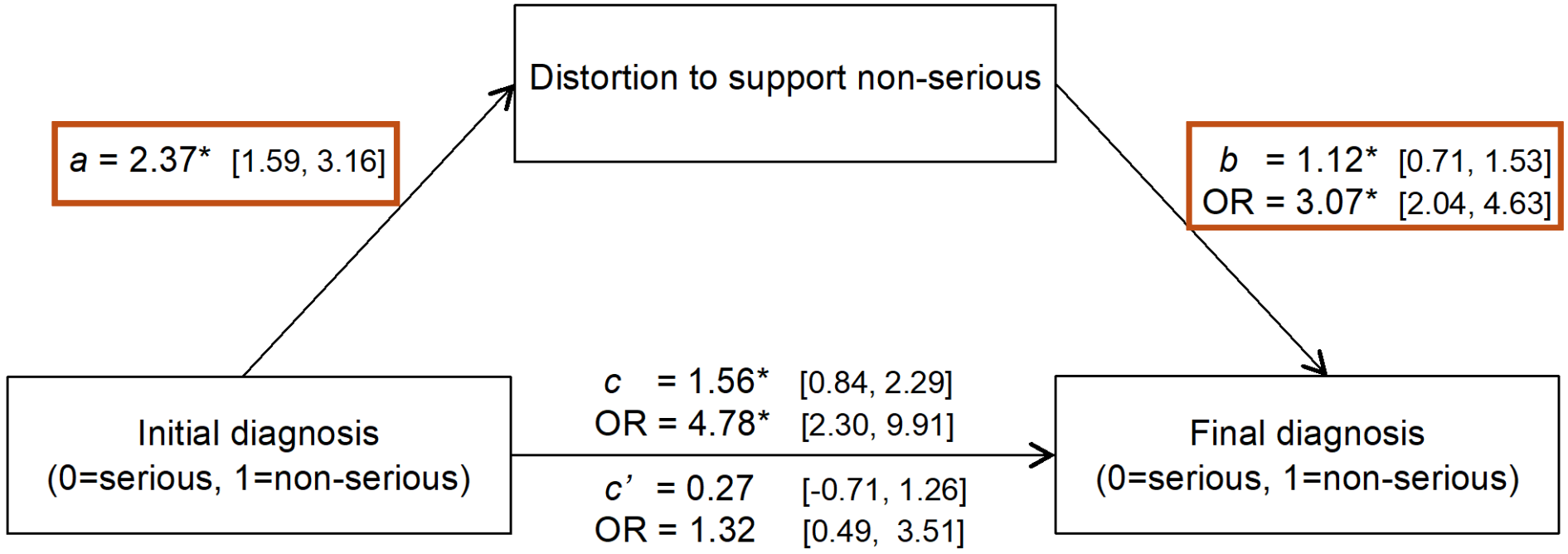
Olga Kostopoulou, PhD, J. Edward Russo, PhD, Greg Keenan, MSc, Brendan C. Delaney, MD, Abdel Douiri, PhD

Judgment and Decision Making, Vol. 9, No. 6, November 2014, pp. 572–585

Predecisional information distortion in physicians' diagnostic judgments: Strengthening a leading hypothesis or weakening its competitor?

Martine Nurek* Olga Kostopoulou† York Hagmayer‡

PID explained 91% of the total effect *c*



* *P*<0.001

Confidence can increase bias

- The more certain we are about an initial hypothesis, the less likely we are to test it sufficiently and rigorously

Kourtidis *et al.*
Cognitive Research: Principles and Implications (2022) 7:103
https://doi.org/10.1186/s41235-022-00453-y

Cognitive Research: Principles and Implications

ORIGINAL ARTICLE Open Access

Influences of early diagnostic suggestions on clinical reasoning

Ploutarchos Kourtidis, Martine Nurek, Brendan Delaney and Olga Kostopoulou*

"**Initial certainty was the main driver of behaviour**: high certainty led to significantly fewer information requests, more biased information evaluations, and fewer changes in diagnosis when encountering new information that either suggested additional possibilities or did not entirely fit with one's leading hypothesis. The negative relationship between initial certainty and change in diagnosis was **consistent across the three experiments**" (*N*s=194, 248, 190)

Individual differences in thinking style

Actively Open-minded Thinking (AOT)

- Are you open to alternative explanations that oppose an initial judgement?
- Do you actively search for evidence to disconfirm pre-established beliefs (Baron, 2006, 2019).

- Example AOT questions:
  - People should revise their conclusions in response to relevant new information.
  - Changing your mind is a sign of weakness.
  - People should search actively for reasons why they might be wrong.

AOT and diagnosis

- Active open-mindedness was associated with more requests for information, more investigations, and more differential diagnoses (Kourtidis et al. 2022).

Debiasing thinking

Good thinking strategies

To increase AOT

- Facilitate consideration of alternatives
- Direct attention to evidence that would otherwise not be considered.

1. Consider the opposite (What if…?)
2. Self-generate alternatives
3. Counteract unjustified confidence (Why might I be wrong? Consider reasons why my initial judgement might be wrong)
4. Designate a devil's advocate (human or AI)
5. Adopt another stakeholder's perspective (patient, carer, medical director)

(e.g., Larrick, *Debiasing,* 2004)

Do these work in medicine?

- Scarce evidence, mostly negative (no impact)
  - Medical students or junior doctors
  - Stand-alone, one-off interventions
  - Written vignettes

- Can they be deployed spontaneously?

Jess's Rule

'Three strikes and we RETHINK'

R Review the case thoroughly after three unresolved or escalating presentations.

E Elevate the concern for a 'fresh eyes' approach and clinical reassessment.

T Think again—especially if the original diagnosis was unsubstantiated.

H Hear the patient fully and consider what's changed or been missed.

I Invite a second opinion or peer discussion within Primary Care.

N Navigate referral to secondary care where appropriate.

K Keep continuity with the same GP and ensure in-person assessment if previous care was remote.

RCGP Royal College of General Practitioners

Jessica Brady CEDAR Trust

NHS England

Department of Health & Social Care

External aids

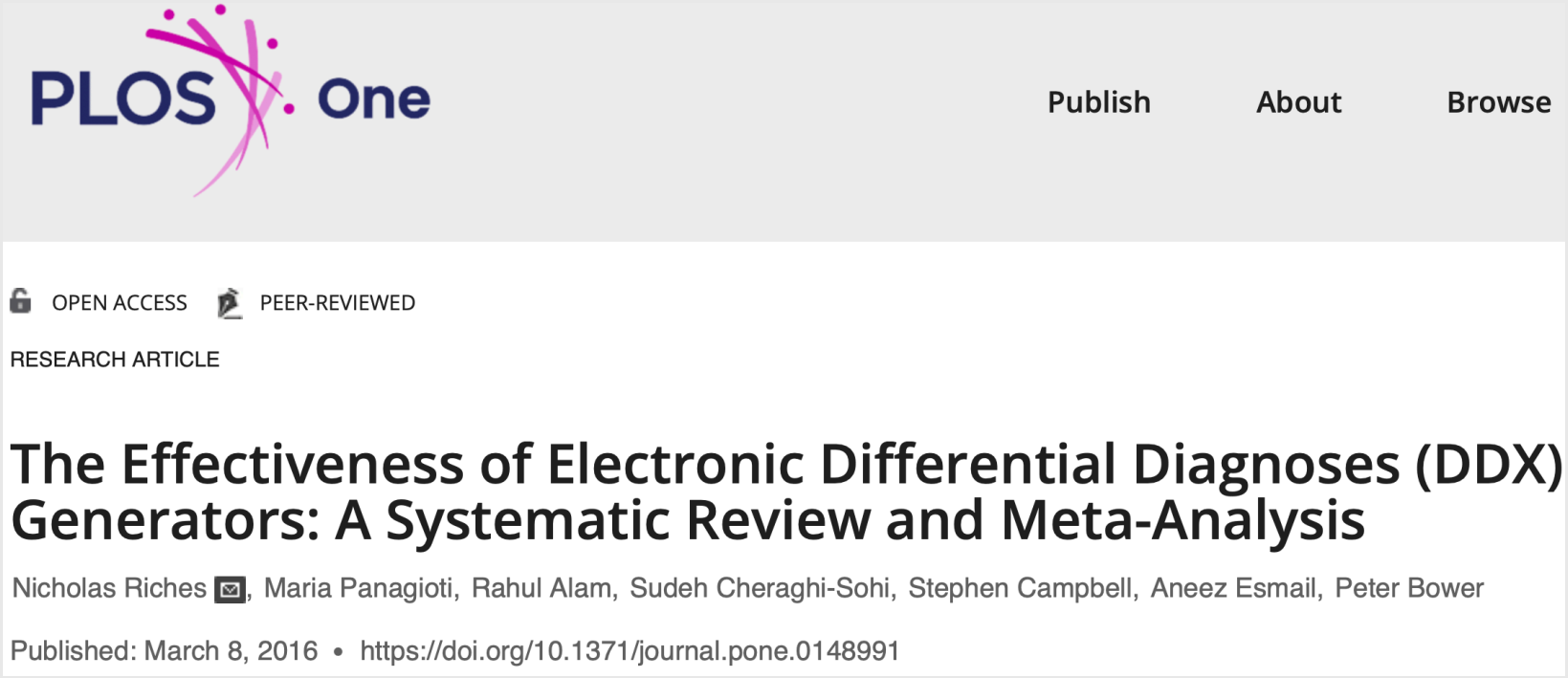
Diagnostic Decision support systems

Differential diagnosis generators

- Potentially helpful – mixed evidence
- Small but significant improvement in accuracy (before-after studies)
- Poor methodological quality of the included studies & high between-studies heterogeneity

  - Extensive lists of low relevance

  - Inconclusive evidence re: number of investigations, cost-effectiveness



Potential biasing effect

- “when the most likely diagnosis was included in the list of suggestions (vs. not included), physicians who gave that diagnosis initially, tended to request less information, evaluate it as more supportive of their diagnosis, become more certain about it, and change it less frequently when encountering new but ambiguous information; in other words, they seemed to **validate rather than question** their initial hypothesis.”

Kourtidis *et al.*
Cognitive Research: Principles and Implications (2022) 7:103
https://doi.org/10.1186/s41235-022-00453-y

Cognitive Research: Principles and Implications

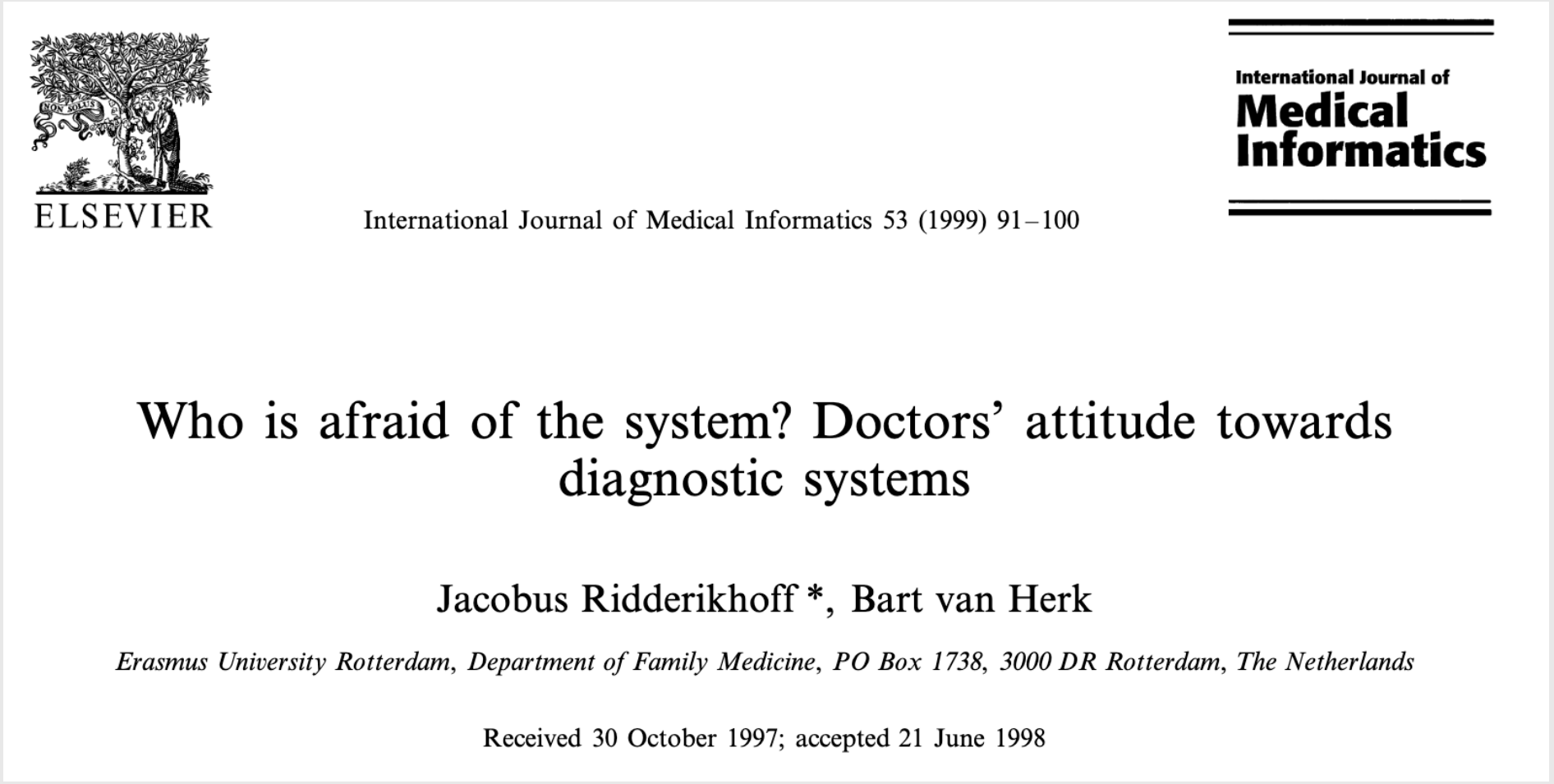
ORIGINAL ARTICLE Open Access

Influences of early diagnostic suggestions on clinical reasoning

Ploutarchos Kourtidis, Martine Nurek, Brendan Delaney and Olga Kostopoulou*

Potential biasing effect

- “The look at the ddx list seems to serve only one purpose: the verification of the diagnostic assumption” (p. 98).



ELSEVIER

International Journal of Medical Informatics 53 (1999) 91–100

International Journal of Medical Informatics

Who is afraid of the system? Doctors’ attitude towards diagnostic systems

Jacobus Ridderikhoff *, Bart van Herk

Erasmus University Rotterdam, Department of Family Medicine, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Received 30 October 1997; accepted 21 June 1998

DSS: Intervene early

BJGP British Journal of General Practice
bringing research to clinical practice

Research

Early diagnostic suggestions improve accuracy of GPs: a randomised controlled trial using computer-simulated patients

Olga Kostopoulou, Andrea Rosen, Thomas Round, Ellen Wright, Abdel Douiri and Brendan Delaney

British Journal of General Practice 2015; 65 (630): e49-e54. **DOI:** https://doi.org/10.3399/bjgp15X683161

N=198 GPs (UK)

N=100 GPs (Greece)

9 rich clinical scenarios

Range of difficulty

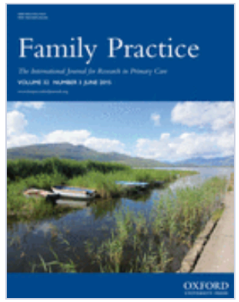
Online

Interactive

	Control	Early support
UK	63%	69%
GR	60%	71%

Pooled OR 1.40 [1.13, 1.67]

Family Practice



Volume 32, Issue 3
June 2015

JOURNAL ARTICLE

Early diagnostic suggestions improve accuracy of family physicians: a randomized controlled trial in Greece

Olga Kostopoulou ✉, Christos Lionis, Agapi Angelaki, Salma Ayis, Stevo Durbaba, Brendan C Delaney

Family Practice, Volume 32, Issue 3, June 2015, Pages 323–328,

N=34 GPs

12 standardised patients (actors)

Within-participant design:

1. Baseline (6 patients): usual practice (EHR)
2. EHR+DSS (6 patients)
   - On different days
   - Patients counterbalanced

- Diagnostic accuracy: 8-9% absolute improvement (*OR* 1.41 [1.13, 1.77]; *OR* 1.50 [1.14, 1.99])
- Investigations: no significant increase
- Time: no significant increase

A recent clinical trial

THE LANCET
Digital Health

ARTICLES · Volume 7, Issue 2, E136-E144, February 2025 · *Open Access* Download Full Issue

Diagnoses supported by a computerised diagnostic decision support system versus conventional diagnoses in emergency patients (DDX-BRO): a multicentre, multiple-period, double-blind, cluster-randomised, crossover superiority trial

Wolf E Hautz, MD [a,*] · Thimo Marcin, PhD [a,*] · Stefanie C Hautz, PhD [a] · Stefan K Schauber, PhD [b] · Prof Gert Krummrey, MD [a,c] · Martin Müller, MD [a] · et al. Show more

- 4 emergency departments in Switzerland
- adults with fever of unknown origin, non-traumatic abdominal pain, syncope, or non-specific symptoms
- Isabel Pro DDx Generator, Isabel Healthcare, UK
- **No benefit**

Early enough?

THE LANCET
Digital Health

ARTICLES · Volume 7, Issue 2, E136-E144, February 2025 · *Open Access* Download Full Issue

Diagnoses supported by a computerised diagnostic decision support system versus conventional diagnoses in emergency patients (DDX-BRO): a multicentre, multiple-period, double-blind, cluster-randomised, crossover superiority trial

Wolf E Hautz, MD [a,*] · Thimo Marcin, PhD [a,*] · Stefanie C Hautz, PhD [a] · Stefan K Schauber, PhD [b] · Prof Gert Krummrey, MD [a,c] · Martin Müller, MD [a] · et al. Show more

- *"Resident physicians inputed all signs and symptoms they deemed relevant into the prepopulated CDDSS, submitted those data, and received suggestions for possible differential diagnoses to consider from the CDDSS"*
- Is consultation of a DSS immediately after taking a patient history and conducting a physical examination sufficiently early?

A single negative trial does not mean DSS is ineffective

THE LANCET
Digital Health

This journal Journals Publish Clinical Global health Multimedia Events About

COMMENT · Volume 7, Issue 2, E108-E109, February 2025 · *Open Access* Download Full Issue

AI for medical diagnosis: does a single negative trial mean it is ineffective?

Olga Kostopoulou [a] · Brendan Delaney [b]

- Implementation – need to understand exactly how AI tools are implemented and how they change cognition

My takeaway

Enabling good thinking strategies

Increasing AOT

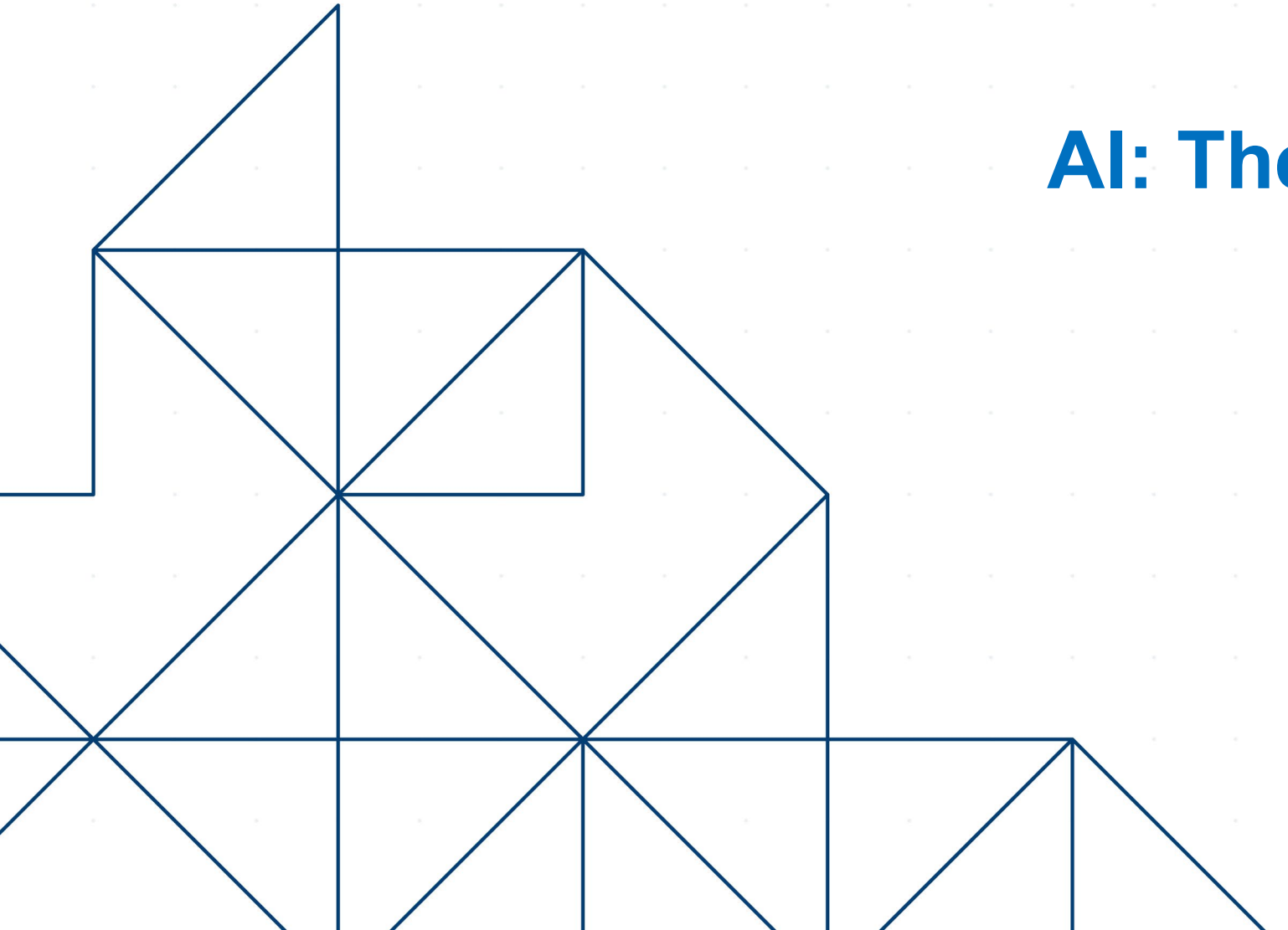
ANNUAL REVIEW OF PSYCHOLOGY Volume 76, 2025

Review Article | Open Access

Boosting: Empowering Citizens with Behavioral Science

Stefan M. Herzog[1] and Ralph Hertwig[1]

- Necessitates cognitive ability & motivation
- Self-nudging, habit formation, anticipating obstacles, changing the environment, reminders, apps & algorithms
- Long-term education and consistent training
- Safeguards autonomy and professional identity



AI: The Present & Future

- LLMs
- Explainable AI
- AI-powered, Ambient Voice Technologies (digital scribes)
- AI virtual doctors that collect data and suggest diagnoses before a clinician has met the patient.