

# Clinician and computer: a study on patient perceptions of artificial intelligence in skeletal radiography

Thomas York <sup>1,2</sup>, Heloise Jenney,<sup>1</sup> Gareth Jones<sup>3</sup>

**To cite:** York T, Jenney H, Jones G. Clinician and computer: a study on patient perceptions of artificial intelligence in skeletal radiography. *BMJ Health Care Inform* 2020;**27**:e100233. doi:10.1136/bmjhci-2020-100233

Received 01 September 2020  
Revised 05 October 2020  
Accepted 15 October 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Trauma and Orthopaedics, Imperial College Healthcare NHS Trust, London, UK

<sup>2</sup>Department of Surgery and Cancer, Imperial College London, London, UK

<sup>3</sup>Clinical Senior Lecturer, Trauma and Orthopaedics, Imperial College London, London, UK

## Correspondence to

Dr Thomas York;  
thomas.york19@imperial.ac.uk

## ABSTRACT

**Background** Up to half of all musculoskeletal injuries are investigated with plain radiographs. However, high rates of image interpretation error mean that novel solutions such as artificial intelligence (AI) are being explored.

**Objectives** To determine patient confidence in clinician-led radiograph interpretation, the perception of AI-assisted interpretation and management, and to identify factors which might influence these views.

**Methods** A novel questionnaire was distributed to patients attending fracture clinic in a large inner-city teaching hospital. Categorical and Likert scale questions were used to assess participant demographics, daily electronics use, pain score and perceptions towards AI used to assist in interpretation of their radiographs, and guide management.

**Results** 216 questionnaires were included (M=126, F=90). Significantly higher confidence in clinician rather than AI-assisted interpretation was observed (clinician=9.20, SD=1.27 vs AI=7.06, SD=2.13), 95.4% reported favouring clinician over AI-performed interpretation in the event of disagreement. Small positive correlations were observed between younger age/educational achievement and confidence in AI-assistance. Students demonstrated similarly increased confidence (8.43, SD 1.80), and were over-represented in the minority who indicated a preference for AI-assessment over their clinicians (50%).

**Conclusions** Participant's held the clinician's assessment in the highest regard and expressed a clear preference for it over the hypothetical AI assessment. However, robust confidence scores for the role of AI-assistance in interpreting skeletal imaging suggest patients view the technology favourably.

Findings indicate that younger, more educated patients are potentially more comfortable with a role for AI-assistance however further research is needed to overcome the small number of responses on which these observations are based.

## INTRODUCTION

Presentations relating to the musculoskeletal system account for more than 60% of emergency department primary diagnoses in the UK,<sup>1</sup> and as many as 50% are investigated by means of a plain radiograph (X-ray).<sup>2</sup>

In the UK, this ubiquitous imaging is typically interpreted by junior doctors and nurse

practitioners. Despite the introduction of safety netting measures (eg, virtual review clinics and remote reporting<sup>3 4</sup>), concerns remain regarding diagnostic inaccuracy and its sequelae.<sup>5</sup> Fractures represent the majority of missed diagnoses,<sup>6</sup> exemplified by the mid-foot where they are not recognised in 33%–40% of cases.<sup>7</sup>

In this context it is unsurprising that there is significant interest in the development of artificial intelligence (AI) algorithms capable of delivering point of care radiographic interpretation.

Early clinical studies have demonstrated significant successes for AI driven interpretation of mammograms<sup>8</sup> and chest radiographs.<sup>9</sup> Algorithms for skeletal radiology are less mature, but effective viability and concept studies have now been reported<sup>10 11</sup> amid growing interest from healthcare providers and industry.<sup>12 13</sup>

Despite this, the integration of AI into healthcare systems has not been without controversy and research has raised concerns over how such technology may lead to a deterioration of human clinical skills as we are called on to use them less.<sup>14</sup> The interface between AI-outputs and the human response to them is also an area of concern; so called 'automation bias', the propensity to over-rely on autonomous processes, has led to multiple deaths in aviation<sup>15</sup> and automotive engineering.<sup>14</sup>

With an increasing likelihood that patients may soon have their imaging, at least in part, reviewed by AI it is crucial to understand their attitudes towards this technology so that software can be developed in a patient-centric manner. Much of the existing understanding of the public's attitude to this technology is founded in consumer research,<sup>16 17</sup> and while this has some relevance to healthcare, it does not reflect the importance individuals place on their health, nor does it address

the unique relationship of trust between patient and clinician.

The work conducted into patient's views of AI indicates significant variation by its application. One study showed that wearable biometric monitoring devices (that integrate AI technology) were seen as greatly beneficial by 20% of patients; however, 35% would refuse to use them.<sup>18</sup> In screening for diabetic retinopathy, much higher favourability was observed with 96% being satisfied with an AI led assessment, and 78% preferring it to the manual alternative.<sup>19</sup>

This study set-out to explore patient attitudes to the potential use of AI in assisting clinicians with the interpretation, and subsequent management, of injuries identified through skeletal radiography. It also sought to investigate whether patient's would prioritise their

clinician's opinion over that of an AI in the event of disagreement. Demographic factors, pain and technology-use were explored as factors potentially influencing these attitudes.

## METHODS

### Questionnaire

The study was prospectively approved by Imperial College Healthcare National Health Service (NHS) Trust. A questionnaire was developed (figure 1), based on a scoping review of the literature<sup>20-22</sup> and according to guidance issued by NHS England with regards to question order, structure and response formats.<sup>23</sup> Plain English was used throughout as per the National Council for Voluntary

## Fracture Clinic Survey Artificial Intelligence in Bone and Joint Injuries



Artificial Intelligence: This is when a computer is used to think or learn, letting it make decisions in a similar way that a person might do.

A common example of this could be the spam filter on your email, deciding whether or not a message should come through to you or not.

This survey will ask simple questions about who you are, how you use technology, and your thoughts on using artificial intelligence to look at injuries of bones and joints.

### Privacy and Consent Statement

As a patient at the fracture clinic in St Mary's or Charing Cross hospital, you are invited to take part in this survey on the possible use of artificial intelligence in looking at X-rays of bone injuries.

If you chose to take part in this survey, you will be helping to decide how we can improve the care of patients with similar conditions to your own in the future. Your own treatment/management will not be affected in any way by choosing to, or choosing not to, take part.

The information collected in this survey will not identify you individually and will be held and processed in accordance with GDPR(2018) regulations. Your answers may form part of published work which summarises the opinions of all those who take part.

Should you wish to receive a copy of any work published as a result of this survey, you can contact {STUDY EMAIL ADDRESS}.

Please mark your choice with X

### Q1. What is your gender?

Male

Female

Non-binary/other

### Q2. What is your age in years? (e.g. 45)

### Q3. What is your current employment status?

Employed full time

Employed part time

Self-employed

Homemaker

Unemployed

Unable to work

Student

Retired

### Q4. What is your highest level of education?

No school qualifications

GCSEs/O-Levels

A-Levels/I-Bacc

BTEC/Vocational award

Bachelor's degree

Master's degree

Doctorate

**Figure 1** Participant questionnaire (page 1 of 2). NHS, National Health Service.

**Q5. Do you own or use a smart phone?**

Yes

No

If you answered No, please go to Q.7

**Q6. How frequently do you use a smart phone?**

Less than 30 minutes each day

30 minutes to an hour each day

Between 1 and 2 hours each day

Greater than 2 hours each day

**Q7. Do you own or use a laptop computer?**

Yes

No

If you answered No, please go to Q.8

**Q8. How frequently do you use a laptop computer?**

Less than 30 minutes each day

30 minutes to an hour each day

Between 1 and 2 hours each day

Greater than 2 hours each day

**Q9. How would you rate the pain of your injury at its worst? Please circle your answer.**

1      2      3      4      5      6      7      8      9      10

← No pain at all

Worst pain imaginable →

**Q10. How would you rate the pain of your injury now?**

1      2      3      4      5      6      7      8      9      10

← No pain at all

Worst pain imaginable →

**Q11. How confident do you feel in your doctor/surgeon's ability to assess your X-rays?**

1      2      3      4      5      6      7      8      9      10

← Not confident at all

Extremely confident →

**Q12. How would you feel about artificial intelligence helping your doctor/surgeon to look for injuries on your X-rays?**

1      2      3      4      5      6      7      8      9      10

← Very unhappy

Completely happy →

**Q13. How would you feel about artificial intelligence helping your doctor/surgeon in deciding how to manage your injuries?**

1      2      3      4      5      6      7      8      9      10

← Very unhappy

Completely happy →

**Q14. If the artificial intelligence gave a different answer to your doctor/surgeon, which opinion do you think would be more likely to be correct?**

Artificial intelligence

Doctor/surgeon

**Figure 2** Participant questionnaire (page 2 of 2).

Organisations' guidelines.<sup>24</sup> Data were handled in accordance with General Data Protection Regulations.

The questionnaire included a total of 14 questions, with patients asked to select the most appropriate answer from a list (figure 2). Given the complexity of AI and the variability of its understanding, the questionnaire began with a definition of the concept, including an example designed to be relevant to the widest possible audience (figure 1). Demographic data, employment status, educational level and smart phone/laptop computer use were recorded. Participants were also asked to rate their pain, using a 10-point Likert scale, both at its worst and at the time of questionnaire completion. The final four questions assessed participant's confidence in a clinician's

capacity to interpret their radiographs, their feelings towards the use of AI to assist with diagnosis or management, and who they would trust in the event of a disagreement between clinician and algorithm. Pilot testing of the questionnaire was performed with subsequent review and finalisation by the authors.

### Setting and population

The questionnaire was distributed in the fracture clinics of a large London university hospital (~1.3 million annual patient contacts<sup>25</sup>) from 01 July 2020 to 01 August 2020. On arrival, reception staff issued a single paper copy of the questionnaire to every patient over the age of 16, with instructions to place completed forms in a secure

ballot box prior to attending their scheduled appointment. Completed questionnaires were collected at the end of the study period and the responses collated. IBM SPSS Version 26 software was used for statistical analysis. T-tests were used to detect differences between groups, and correlation assessed using Pearson correlation coefficient. A p value of  $\leq 0.05$  was considered significant.

## RESULTS

A total of 300 questionnaires were produced, 218 were completed and returned (72.2%). Two incomplete forms were excluded, leaving 216 questionnaires with a legible, interpretable answer to each question.

### Confidence in clinician and AI-assisted interpretation

The mean confidence of participants in their clinician's ability to correctly interpret their radiographs was 9.20, where 10=extremely confident, 1=not confident at all. Mean confidence in AI-assisted interpretation was 7.06, significantly lower ( $t(215)=-14.34$ ,  $p<0.001$ ). Across all participants, a positive correlation was observed between confidence in the clinician's interpretation and confidence in AI-assisted interpretation ( $r(215)=0.21$ ,  $p=0.002$ ). When asked to identify the opinion they would favour in the event of disagreement between the clinician and AI, 95.4% (206) of participants selected the clinician, and only 4.6% (10) the AI.

### AI-assisted interpretation and AI-assisted management

Participants rated their confidence in AI-assisted interpretation (7.06) significantly higher than their confidence in AI-assisted management (4.86),  $t(215)=11.03$ ,  $p<0.001$ . A strong positive correlation was observed between confidence in AI-assisted interpretation and confidence in AI-assisted management ( $r(215)=0.693$ ,  $p<0.001$ ).

### Age and gender

The mean age for participants was 40.20 years. More men (126) than women (90) completed the questionnaire (M:F=1.4:1), and on average female participants were 10.3 years older.

Small negative correlations were observed between age and both confidence in AI-assisted radiographic interpretation ( $r(215)=-0.170$ ,  $p=0.0123$ ) and confidence in AI-assisted patient management ( $r(215)=-0.244$ ,  $p<0.001$ ).

Of the 10 participants who indicated a preference for the assessment of an AI over that of their clinician, mean age was found to be significantly younger than the

wider cohort at 24.5 ( $t(215)=-3.05$ ,  $p=0.00125$ ). Of these 10 participants, 70% were men and 30% were women (table 1).

Female participants showed significantly higher confidence in their clinician's assessment than their male counterparts,  $t(215)=3.42$ ,  $p<0.001$ . Conversely female participants displayed a lower confidence in the idea of AI-assisted management,  $t(215)=-4.51$ ,  $p<0.001$ . No significant gender difference was observed with regard to AI-assisted interpretation.

### Employment and educational achievement

A similar distribution of employment status was observed across both genders. The majority of the participants were in some form of work or study (69.4%), with the largest single demographic being those in full time employment (41.2%).

Those participants who were unable to work indicated a significantly lower confidence in their clinician's assessment ( $t(215)=-3.22$ ,  $p<0.001$ ), AI-assisted interpretation ( $t(215)=-2.49$ ,  $p=0.0067$ ) and AI-assisted management ( $t(215)=-1.98$ ,  $p=0.024$ ).

Students reported significantly higher confidence in AI-assisted interpretation ( $t(215)=3.36$ ,  $p<0.005$ ) and in AI-assisted management ( $t(215)=3.68$ ,  $p<0.005$ ). Despite accounting for only 10.6% of the total, 50.0% of participants identifying a preference for an AI assessment over that of their clinician were students (table 2).

A-levels/I-Bacc (30.5%) and bachelor's degrees (30.1%) were the most widely held level of highest educational achievement (table 3). A small positive correlation was observed between increasing educational achievement and confidence in AI-assisted interpretation of radiographs ( $r(215)=0.137$ ,  $p=0.045$ ).

### Laptop computer and smartphone use

One hundred and eighty-seven (86.6%) participants owned or used a smart phone, with 110 (58.8%) using it for greater than 2 hours each day (table 4). A small positive correlation was observed between duration of smartphone use and confidence in AI-assisted interpretation ( $r(215)=0.156$ ,  $p=0.0201$ ).

One hundred and sixty-five participants (76.4%) had access to a laptop, with 84 (38.9%) using it for greater than 2 hours each day (table 5). No significant correlation was found between duration of laptop computer use and confidence in AI assistance for either interpretation, or management.

**Table 1** Gender

Gender	Participants n (%)	Age M (SD)	Clinician confidence M (SD)	AI-interpretation confidence M (SD)	AI-management confidence M (SD)	AI preference n (%)
Male	126 (58.3)	35.9 (15.44)	8.96 (1.39)	7.24 (2.11)	5.42 (2.10)	7 (70)
Female	90 (41.7)	46.18 (17.24)	9.54 (0.98)	6.8 (2.15)	4.06 (2.34)	3 (30)
Total	216(100)	40.2 (16.95)	9.2 (1.27)	7.06 (2.13)	4.86 (2.30)	10 (100)

**Table 2** Employment

Employment	Participants n (%)	Age M (SD)	Clinician confidence M (SD)	AI-interpretation confidence M (SD)	AI-management confidence M (SD)	AI preference n (%)
Full time	89 (41.2)	37.28 (11.70)	9.32 (0.96)	7.37 (1.60)	5.32 (1.61)	2 (20)
Part time	19 (8.8)	31.67 (10.23)	9.55 (0.86)	5.28 (2.11)	2.83 (2.18)	1 (10)
Self-employed	14 (6.5)	39.54 (11.16)	9.54 (0.78)	7.46 (1.71)	4.85 (2.67)	1 (10)
Homemaker	5 (2.3)	49.2 (2.04)	9.8 (0.45)	7 (2.55)	1.4 (0.55)	0 (0)
Unemployed	24 (11.1)	36.59 (9.40)	8.92 (1.38)	6.46 (1.71)	4.58 (1.84)	0 (0)
Unable to work	12 (5.5)	47.8 (12.3)	8.08 (2.68)	5.58 (3.08)	3.58 (2.97)	1 (10)
Student	23 (10.6)	19.04 (2.18)	8.65 (1.67)	8.43 (1.80)	6.48 (2.47)	5 (50)
Retired	30 (13.9)	70.67 (4.51)	9.6 (0.72)	6.67 (2.54)	4.23 (2.01)	0 (0)
Total	216 (100)	40.2 (16.95)	9.2 (1.27)	7.06 (2.13)	4.86 (2.30)	10 (100)

### Pain

The mean maximum pain experienced post-injury was 6.81 (SD=1.70) where 10=worst pain imaginable, 1=no pain at all. Pain was found to be significantly lower at the time of questionnaire completion (3.88, SD=2.01) (mean difference=2.93,  $p<0.001$ ). No correlation was observed between pain and confidence in the clinician or AI.

## DISCUSSION

### Key findings

The most striking finding of this study is the widely held esteem for the reporting skills of clinicians. With an average 10-point Likert rating of 9.2, participants demonstrated extreme confidence in their capacity to correctly interpret radiographs. Although objective evidence of reporting errors suggests that this confidence may be over-valued,<sup>10 26</sup> these observations are consistent with the enduring strength of the doctor–patient relationship, a rapport which continually places doctors among the most trusted of professions both in the UK<sup>27</sup> and internationally.<sup>28 29</sup> Consequently, the significantly lower rating for AI interpretation may reflect the robustness of this relationship rather than any inherent distrust in AI. Indeed, a

mean rating of 7.06 suggests substantial confidence does exist in the technology, even at this early stage.

More broadly though, evidence indicates that people possess a strong preference for human advice over that received from an automated process. This inclination is only overcome when the perceived expertise (pedigree) of the automated adviser is made significantly greater than that of its human counterpart.<sup>30</sup> The evidence of this study strongly supports such a conclusion, with all but 10 participants indicating a preference for the assessment of their clinician over that of an AI.

The great majority of participants owned or regularly used smart phones (86.6%) and laptop computers (74.5%), reflecting a broad societal awareness of novel technologies and potentially their application to new domains.<sup>31</sup> Polling in the UK found 63% of people reported that ‘they knew something about AI’<sup>32</sup> and that, after improvements in home energy efficiency, healthcare was the area to which they felt it would be of most benefit.<sup>33</sup> With younger—more technologically minded—patients becoming a greater proportion of the patient population as they age, it would be reasonable to predict that confidence in the diagnostic use of AI will only continue to grow with time. Such a trend would be

**Table 3** Educational achievement

Educational achievement	Participants n (%)	Age M (SD)	Clinician confidence M (SD)	AI-interpretation confidence M (SD)	AI-management confidence M (SD)	AI preference n (%)
No school qualifications	16 (7.41)	50.9 (18.06)	8.31 (2.24)	5.3 (2.21)	3.6 (2.22)	1 (10)
GCSEs/O-levels	45 (20.8)	39.6 (18.82)	9.16 (1.09)	7.18 (2.19)	4.93 (2.26)	2 (20)
A-levels/I-Bacc	66 (30.5)	40.1 (18.24)	9.39 (1.04)	7.12 (2.12)	4.79 (2.33)	3 (30)
BTEC/vocational award	10 (4.6)	38.6 (14.21)	9.3 (1.34)	7.5 (2.55)	5.6 (2.80)	1 (10)
Bachelor’s degree	65 (30.1)	37.2 (13.62)	9.18 (1.29)	7.12 (1.96)	5.06 (2.30)	3 (30)
Master’s degree	12 (5.55)	43.33 (16.12)	9.42 (0.90)	7.75 (1.86)	5 (1.86)	0 (0)
Doctorate	2 (0.9)	56 (4.24)	10 (0.00)	7.5 (0.71)	4 (1.41)	0 (0)
Total	216 (100)	40.2 (16.95)	9.2 (1.27)	7.06 (2.13)	4.86 (2.30)	10 (100)

**Table 4** Smartphone use

Smartphone use	Participants n (%)	Age M (SD)	Clinician confidence M (SD)	AI-interpretation confidence M (SD)	AI-management confidence M (SD)	AI preference n (%)
No use	29 (13.4)	65.29 (13.5)	9.52 (0.99)	6.55 (2.49)	4.31 (1.94)	0 (0)
<30 min	22 (10.2)	56.68 (11.44)	9.68 (0.57)	6.41 (2.24)	4.04 (1.81)	0 (0)
30 min–1 hour	12 (5.5)	44.08 (10.46)	9.33 (0.78)	7.5 (1.88)	5.58 (3.09)	1 (10)
1–2 hours	43 (19.9)	41.16 (10.6)	9.07 (1.35)	6.67 (1.84)	4.67 (2.07)	0 (0)
>2 hours	110 (50.9)	29.52 (10.07)	9.06 (1.41)	7.42 (2.09)	5.15 (2.42)	9 (90)
Total	216 (100)	40.2 (16.95)	9.2 (1.27)	7.06 (2.13)	4.86 (2.30)	10 (100)

consistent with the results of literature in consumer and employment research, which indicates a growing familiarity with AI.<sup>33 34</sup>

While this study did detect a significant correlation between younger participants and greater confidence in AI-assisted interpretation and management, the magnitude of this effect was small. As a demographic, students showed higher confidence in these same measures and were statistically over-represented in the cohort who favoured the interpretation of an AI over that of their clinician. However, these findings should be interpreted with caution due to the limited number of responses on which they are based.

### Limitations

The design of this study was dependent on patients presenting to an in-person appointment at the fracture clinic of a central London teaching hospital during a time when attendance was purposely and observably reduced due to COVID-19.<sup>35</sup> Given these measures, and strict guidelines for ‘shielding’ of vulnerable patients,<sup>36</sup> it is likely that participants numbers were reduced and skewed towards individuals with limited comorbidities and generally better health. However, the wide range of participants included in this study (age in particular) suggest this effect is unlikely to have significantly impacted the results.

While a questionnaire was central to this study, there is a scarcity of validated methods for questionnaire reporting.<sup>37</sup> To mitigate the potential for error, validated question designs (such as the Likert scale<sup>38</sup>) were used to minimise the complexity of this interpretation. The extremely positive perception of clinicians held by

participants (9.20 out of 10 on average) may be indicative of response bias whereby participants answered in a way which they perceived as desirable for the clinicians. However, various efforts were made to mitigate this bias: questions were written in readily comprehensible language, implicit communication in the questioning was avoided<sup>26–29</sup> and completed questionnaires were anonymous and returned to a locked ballot box.

### CONCLUSIONS

Ultimately, this study indicates a significant preference among its participants for the assessment of their clinician over that of an AI. This was set against a backdrop of high confidence in the capacity of clinicians to correctly interpret the skeletal radiographs of their patients. Despite this, the prospect of AI-assisted interpretation was widely supported, with significantly lower but still robust confidence scores.

Demographic factors were identified which may suggest greater patient support for the use of AI in skeletal radiography; in particular among those of younger age, higher educational achievement and students. These findings should be interpreted cautiously due to the limited number of responses from which they are drawn.

This study indicates a basal level of patient approval for the use of AI as an assistant to their clinician. When coupled with a population growing more accustomed to the technology’s role in wider society, the continued development of AI algorithms to target skeletal radiography is clearly justified. Further research into the demographics of attitudes to AI is needed to expand on these

**Table 5** Laptop computer use

Laptop use	Participants n (%)	Age M (SD)	Clinician confidence M (SD)	AI-interpretation confidence M (SD)	AI-management confidence M (SD)	AI preference n (%)
No use	51 (23.6)	53.78 (19.96)	9.09 (1.55)	6.67 (2.39)	4.59 (2.31)	2 (20)
<30 min	24 (11.1)	44.08 (14.00)	9.54 (0.93)	6.83 (1.86)	5.21 (2.17)	2 (20)
30 min–1 hour	15 (6.94)	46.87 (13.64)	9.67 (0.62)	7.34 (2.57)	5.13 (2.1)	0 (0)
1–2 hours	42 (19.4)	37.48 (13.07)	9.21 (1.12)	7.05 (2.21)	5 (2.51)	2 (20)
>2 hours	84 (38.9)	31.02 (10.92)	9.1 (1.30)	7.31 (1.97)	4.8 (2.28)	4 (40)
Total	216 (100)	40.2 (16.95)	9.2 (1.27)	7.06 (2.13)	4.86 (2.30)	10 (100)

findings; in particular to clarify the role of patient knowledgeability in their willingness for such technology to be involved in their care.

Although not yet ready to prioritise its assessment above that of their clinician's, patients do appear content for clinicians to look to AI for support in the diagnosis and management of skeletal injuries.

**Contributors** The manuscript has been read and approved by all authors, the ICMJE requirements for authorship are satisfied, and each author believes the manuscript to be an accurate and honest work. Conception or design of the work: TY, HJ, GJ. Data collection: TY, HJ. Data analysis and interpretation: TY, HJ, GJ. Drafting the article: TY. Critical revision of the article: TY, HJ, GJ. Final approval of the version to be published: TY, HJ, GJ.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Datasets are available upon reasonable request.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Thomas York <http://orcid.org/0000-0002-3946-561X>

#### REFERENCES

- NHS Digital. Accident and Emergency Attendances in England - 2009-2010, Experimental statistics - NHS Digital, 2020. Available: <https://digital.nhs.uk/data-and-information/publications/statistical/hospital-accident-emergency-activity/2009-2010> [Accessed 26 Aug 2020].
- de Lacey G. Number of casualty attenders referred for x-ray examination. *Br J Radiol* 1979;52:332-4.
- Holgate J, Kirmani S, Anand B. Virtual fracture clinic delivers British orthopaedic association compliance. *Ann R Coll Surg Engl* 2017;99:51-4.
- Brady A, Laoide Risteárd Ó, McCarthy P, et al. Discrepancy and error in radiology: concepts, causes and consequences. *Ulster Med J* 2012;81:3-9.
- Pakdemirli E. Teleradiology in the UK: friend or foe? A radiologist's perspective. *Acta Radiol Open* 2019;8:205846011986131.
- Pinto A, Berritto D, Russo A, et al. Traumatic fractures in adults: missed diagnosis on plain radiographs in the emergency department. *Acta Biomed* 2018;89:111-23.
- York TJ, Jenkins PJ, Ireland AJ. Reporting discrepancy resolved by findings and time in 2947 emergency department ankle x-rays. *Skeletal Radiol* 2020;49:601-11.
- McKinney SM, Sieniek M, Godbole V, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020;577:89-94.
- Qin ZZ, Sander MS, Rai B, et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: a multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci Rep* 2019;9:15000.
- Gan K, Xu D, Lin Y, et al. Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments. *Acta Orthop* 2019;90:394-400.
- Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians. *Proc Natl Acad Sci U S A* 2018;115:11591-6.
- NHS Accelerated Access Collaborative. Artificial intelligence in health and care Award, 2020. Available: <https://www.england.nhs.uk/aac/what-we-do/how-can-the-aac-help-me/ai-award/> [Accessed 26 Aug 2020].
- NHS Long Term Plan. Case studies, 2020. Available: <https://www.longtermplan.nhs.uk/case-studies/> [Accessed 26 Aug 2020].
- Sujan M, Furniss D, Grundy K, et al. Human factors challenges for the safe use of artificial intelligence in patient care. *BMJ Health Care Inform* 2019;26:e100081.
- Mosier KL, Skitka LJ, Heers S, et al. Automation bias: decision making and performance in high-tech Cockpits. *Int J Aviat Psychol* 1998;8:47-63.
- Gursoy D, Chi OH, Lu L, et al. Consumers acceptance of artificially intelligent (AI) device use in service delivery. *Int J Inf Manage* 2019;49:157-69.
- Schepman A, Rodway P. Initial validation of the general attitudes towards artificial intelligence scale. *Computers in Human Behavior Reports* 2020;1:100014.
- Tran V-T, Riveros C, Ravaud P. Patients' views of wearable devices and AI in healthcare: findings from the compare e-cohort. *NPJ Digit Med* 2019;2:53.
- Keel S, Lee PY, Scheetz J, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Sci Rep* 2018;8:4330.
- Boynton PM, Greenhalgh T. Selecting, designing, and developing your questionnaire. *BMJ* 2004;328:1312-5.
- Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010;11:2.
- Dell-Kuster S, Sanjuan E, Todorov A, et al. Designing questionnaires: healthcare survey to compare two different response scales. *BMC Med Res Methodol* 2014;14:96.
- NHS England. Writing an effective questionnaire. Available: <https://www.england.nhs.uk/wp-content/uploads/2018/01/bitesize-guide-writing-an-effective-questionnaire.pdf> [Accessed 26 Aug 2020].
- NCVO Knowhow. Available: <https://knowhow.ncvo.org.uk/sitemap> [Accessed 26 Aug 2020].
- Annual report 2019-2020, 2020. Available: <https://indd.adobe.com/view/d3f9b7c4-170f-4cf9-8e13-3689985b9d65> [Accessed 26 Aug 2020].
- Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging* 2017;8:171-82.
- Ipsos MORI. Public trust in doctors remains high, 2020. Available: <https://www.ipsos.com/ipsos-mori/en-uk/public-trust-doctors-remains-high> [Accessed 26 Aug 2020].
- Advisory. For the 18th year in a row, nurses are the most-trusted profession, according to Gallup, 2020. Available: <https://www.advisory.com/daily-briefing/2020/01/10/nurse-trusted> [Accessed 26 Aug 2020].
- Ipsos. Who do you have faith in? The world's most trusted professions. Available: <https://www.ipsos.com/en/who-do-you-have-faith-worlds-most-trusted-professions> [Accessed 26 Aug 2020].
- Pearson CJ, Geden M, Mayhorn CB. Who's the real expert here? pedigree's unique bias on trust between human and automated advisers. *Appl Ergon* 2019;81:102907.
- Ritchie H, Roser M. Technology adoption, 2020. Available: <https://ourworldindata.org/technology-adoption> [Accessed 27 Aug 2020].
- KANTAR. AI PR survey - summary of findings, 2020. Available: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/802548/BEIS\\_AI\\_PR\\_Survey\\_40309009\\_Topline\\_summary\\_V1\\_1\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/802548/BEIS_AI_PR_Survey_40309009_Topline_summary_V1_1_.pdf)
- PWC. Accelerating innovation how to build trust and confidence in AI, 2019. Available: <https://www.pwc.co.uk/audit-assurance/assets/pdf/responsible-artificial-intelligence.pdf>
- Blumberg Capital. Artificial intelligence in 2019: getting past the adoption tipping point, 2020. Available: <https://blumbergcapital.com/ai-in-2019/>
- Wallace CN, Kontoghiorghis C, Kayani B, et al. The impact of COVID-19 on trauma and orthopaedic surgery in the United Kingdom. *Bone Jt Open* 2020;1:420-3.
- GOV.UK. COVID-19: guidance on shielding and protecting people defined on medical grounds as extremely vulnerable, 2020. Available: <https://www.gov.uk/government/publications/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19>
- Bennett C, Khangura S, Brehaut JC, et al. Reporting guidelines for survey research: an analysis of published guidance and reporting practices. *PLoS Med* 2010;8:e1001069.
- Arnetz B, Hasson D. Validation and findings comparing vas vs. Likert scales for psychosocial measurements. *International Electronic Journal of Health Education* 2020;8:178-92.