Northumbria Psychology Bulletin (ISSN 3033-4225)
https://www.northumbriajournals.co.uk/index.php/npb/article/view/1430

Published date: 29 November 2024

Research Article

Inattentional blindness in radiology: a concise checklist approach

Christopher Lavender[1*], Joanna Greer[1]

[1]Northumbria University, Department of Psychology, Northumberland Building, City Campus, Newcastle upon Tyne, NE1 8ST, UK

***: Northumbria University Psychology Student Author**

Abstract

Inattentional blindness has been identified as a partial cause for missed diagnoses among radiologists. Missed findings present a significant challenge as they can have clinical implications for patients. This study investigated the effectiveness of a four-item concise medical checklist in reducing inattentional blindness among radiologists when interpreting chest computed tomography (CT) scans. Thirty-two radiologists participated in the study: an experimental group (with the checklist, n = 18) and a control group (no checklist, n = 14). Participants were instructed to read seven chest CT stacks (one practice case and six experimental cases), and to mark all lung nodules ≥3 mm. In the final CT stack, a breast cancer mass and lymphadenopathy served as the inattentional blindness stimuli. Lung nodule detection was marginally higher in the control group (62%) than in the experimental group (55%), but this difference was not statistically significant. Almost 80% of radiologists in both groups failed to report the breast cancer mass, whilst lymphadenopathy identification was at chance level in both the control (50%) and experimental (58%) groups. Group comparisons for both analyses were also non-significant. These findings suggest that a concise medical checklist may not be an effective solution to mitigate inattentional blindness among radiologists when interpreting chest CT scans. Further research and alternative approaches are warranted to address diagnostic errors in medical imaging resulting from inattentional blindness.

Keywords: Inattentional Blindness, radiologists, diagnostic Errors, chest computed tomography (CT), medical checklist.

Corresponding Author:
Christopher Lavender, Department of Psychology, Northumberland Building, City Campus, Newcastle upon Tyne, NE1 8ST, UK; email: chris_lavender@hotmail.co.uk

Introduction

When attending to even the simplest tasks, observers commonly fail to notice a stimulus — either an object or an event — that, in hindsight, appears obvious. This is known as inattentional blindness and was famously demonstrated in an experiment where participants, who were required to count the number of ball passes between players, missed seeing a person in a gorilla costume present within the scenario (Simons & Chabris, 1999). Since then, inattentional blindness has been observed in various studies, such as static and moving objects displayed on a computer screen, walking and talking on a cell phone, and simulated assault (Bressan & Pizzighello, 2008; Chabris *et al*., 2011; Hyman *et al*., 2010; Most *et al*., 2001). This failure to perceive seemingly noticeable stimuli is thought to reflect limitations in one's attentional capacity and the selective processing of information (Mack, 2003; Simons, 2000).

Within healthcare settings, inattentional blindness can have real-world implications for clinical diagnosis, with high rates of diagnostic errors being partially attributed to this phenomenon (Garg et al., 2022; Jager *et al*., 2014). For example, Kim and Mansfield (2014) found that out of 1,269 abnormalities present in 656 radiology examinations, 42% were missed. Notably, 7% of these errors were attributed to inattentional blindness.

While human error cannot be discounted in such incidents, even experts are susceptible to inattentional blindness (Ekelund *et al*., 2022). Some studies have suggested that expertise might reduce its occurrence by freeing up attentional resources (Drew *et al*., 2013; Pammer *et al*., 2018; Simons & Schlösser, 2017), while others have claimed experts may be more susceptible due to their deep focus on specific tasks (Ho *et al*., 2017). However, a meta-analysis revealed that experts only had marginally improved performance compared with novices, and crucially, that differing stimuli (e.g., experimental manipulation such as a non-clinical image cf. clinically relevant such as a lung nodule) had minimal modulating effects, even when it was related to the expert's domain (Ekelund *et al*., 2022). In one of the included studies, Drew *et al*., (2013) found that 83% of expert radiologists failed to identify an image of a gorilla located in computed tomography (CT) images of the lungs, despite the gorilla being 48 times larger than the average lung nodule. Expectedly, expert radiologists outperformed naïve participants with no medical training in regard to the mean lung nodule detection rate (55% vs. 12%), but only slightly outperformed them when it came to inattentional blindness (83% vs. 100%).

Building on previous work (Drew *et al*., 2013), Williams *et al*., (2021) used clinically relevant stimuli in the form of an incidental finding, which is the discovery of unexpected or unrelated abnormalities (Lumbreras *et al*., 2010). In their first study, when searching for lung nodules, 66% of radiologists did not report an incidental breast cancer mass. However, in their second study, with a different sample of radiologists, only one out of 30 radiologists did not report it when they were asked to check all the abnormalities present from a list of six options (Williams *et al*., 2021). These studies emphasize the robustness of the inattentional blindness phenomenon across stimuli and underscore the need to mitigate its impact in clinical practice, given the harm that could arise from missed incidental findings for both patients and clinicians (Berlin, 2007; Morris *et al*., 2009).

To reduce the high incidence of diagnostic errors in radiology generally, several strategies are used with varying success. For example, double-reading, where two or more radiologists examine the same images, is grounded in the belief that multiple clinicians reviewing the same images improves accuracy. Yet the rates of discrepancy between clinicians’ diagnoses are relatively low and therefore need to be balanced against demanding workloads (Geijer & Geijer, 2018). Additionally, a considerable number of second-opinion radiology reports go unread, suggesting the potential for more efficient resource allocation (Heinz *et al*., 2020). Thus, there is a need for a low-resource-dependent solution that can be easily integrated into radiologists' workflows.

One such tool might be a medical checklist. Medical checklists have been proposed by several researchers over the years to try and address the challenge of inattentional blindness in medical imaging (Gefter & Hatabu, 2023; Williams *et al*., 2021). They are common within healthcare settings and encouraged in diagnostic radiology generally (Iyer *et al*., 2013), typically consisting of a series of steps or questions to help guide medical professionals during tasks such as surgical procedures, medication administration, and patient admissions and discharges (Winters *et al*., 2009). Studies have demonstrated their effectiveness in reducing prescription, and surgical errors (Alagha *et al*., 2011; Haugen *et al*., 2015; Haynes *et al*., 2009; Thomassen *et al*., 2014). To date however, limited research has assessed the efficacy of a checklist in medical imaging, and no study has comprehensively tested the use of a checklist in relation to the incidence rate of inattentional blindness among radiologists when interpreting chest CT scans. Whilst the study discussed above (Williams *et al*., 2021) suggests a checklist might have some efficacy, its scope is restricted by the analysis of only two scans and no control group.

One study, however, does provide some support for its use. The study found that among 40 medical students who were tasked with reading 18 chest X-rays, those who used a systematic medical checklist detected more abnormalities compared with those who did not (Kok *et al.,* 2017). In contrast, research using a checklist-style structured report for maxillofacial CT scans did not find any increase in the reporting accuracy rates for undetected pathology issues (Powell *et al*., 2014). The main drawback identified in this previous study (Kok *et al*., 2017) was that the systematic checklist was too time-consuming as it consisted of anatomical areas, potential pitfalls, and commonly missed diagnoses. This poses significant implementation challenges in a clinical setting where accuracy and efficiency are key.

Perhaps more suited to real-life clinical settings would be a concise medical checklist, consisting of short, simple questions that can encompass a range of possible findings and abnormalities, without the need to specify all of them. The underlying rationale is that a checklist might disrupt the observer's cognitive process, effectively slowing them down and potentially enhancing their overall visual perception and decision-making accuracy (Croskerry *et al*., 2012a, 2012b; Ely *et al*., 2011).

The aim of the current study was to determine if a concise medical checklist could reduce inattentional blindness among radiologists, thereby resulting in greater detection of clinical abnormalities. It was hypothesised that 1): a concise checklist would facilitate a more vigilant search strategy, leading to improved detection of lung nodules, and 2): using the checklist would increase the detection rate of incidental findings.

Method

A between-group quasi-experimental design was used comparing two independent groups (experimental (checklist) or control (no checklist)) on the detection of lung cancer nodules (hypothesis 1) and breast cancer symptoms (hypothesis 2).

Participants

An *a priori* power analysis, conducted using G*Power 3.1 (Faul *et al.,* 2007), using an independent samples *t*-test (α = .05 and effect size of d = 0.5 at 80% power determined that a sample size of 128 participants (64 participants per group) was required. A group allocation matrix was created based on the desired sample size (n = 124) using an online randomisation tool (https://www.graphpad.com/quickcalcs/randomize1/; GraphPad, Boston, MA). A total of 64 participant numbers were randomly assigned to either Group A or Group B. Participants assigned to Group A were allocated to the experimental group (with the checklist), while those assigned to Group B were allocated to the control group (without the checklist).

Participants had to be aged 18+ years old, and an attending radiologist or a resident in a radiology training program. Participants could not participate if they had self-reported abnormal or non-corrected vision. The study was approved by the Faculty of Health and Life Sciences Ethics Committee at Northumbria University (ref: 4953). All participants provided electronic informed consent.

Recruitment began by contacting radiology line managers at three cancer centres in Guangzhou, China, who advertised the study on *WeChat*, and by word of mouth. The line manager communicated when and where the researcher (C.L) was on-site conducting the study. A total of 32 participants were recruited. Upon signing up for the study, each participant was sequentially assigned a participant number and allocated to either the experimental or control group accordingly.

Materials and Measures

Pre-task survey

Participants completed a self-report pre-task survey that gathered demographic information, including professional experience (the number of years that they had been working in the field of radiology), position, specialty, the number of CT scans interpreted by each radiologist per week, and vision-related impairments.

Computed tomography (CT) stacks

The same anonymised seven chest CT stacks used by Williams *et al.,* (2021) were used in this study, six of which are freely available from the Lung Image Database Consortium (LIDC) for research purposes (Armato *et al.,* 2011).

Three of the CT stacks (CT1, CT4, CT6 and CT7) contained lung nodules, and three did not (CT2, CT3 and CT5). CT stack 1 (the practice case), had eight lung nodules; CT stack 4 had 23; CT stack 6 had 10; and CT stack 7 (the inattentional blindness stack) had eight lung nodules. Williams *et al.,* (2021) reported the inattentional blindness stack as having nine lung nodules. However, one specified location did not appear to show a visible nodule, resulting in a total of eight lung nodules in this study. The seventh CT stack also included a large (9.1 cm) breast cancer mass, visible on 17 of 66 image slices, along with lymphadenopathy. These incidental findings served as the inattentional blindness stimuli and were chosen for their clear visibility within typical lung window settings (*Figure 1*). All stacks had a resolution of 512x512 pixels.

Lung nodule detection accuracy was measured by comparing participants' annotations with the LIDC reference data (Armato *et al.,* 2011) based on screen-recorded videos. Annotations were deemed accurate if they were within 30 pixels and two slices of the nodule's centre of mass. For the inattentional blindness stack, the location of the lung nodules was obtained from a previous study (Williams *et al*., 2021) and the same process applied.

Participants reviewed each of the seven chest CT stacks sequentially (one practice and six experimental) marking all nodules ≥3 mm using MicroDicom Viewer (MicroDicom Ltd, Sofia, Bulgaria) on a Microsoft Surface 7 tablet computer. The CT stacks were preloaded, and the screen's brightness set to full (104 nits). The Window Level (WL) and Window Width (WW) were configured to standard lung settings (WL = 400, WW = 1600) and kept constant throughout the task (*Figure 2*). Participants used a preselected ellipse measuring tool to the mark lung nodules. The screen activity was recorded using Snipping Tool (version 11.2, Microsoft Corporation, Redmond, USA)

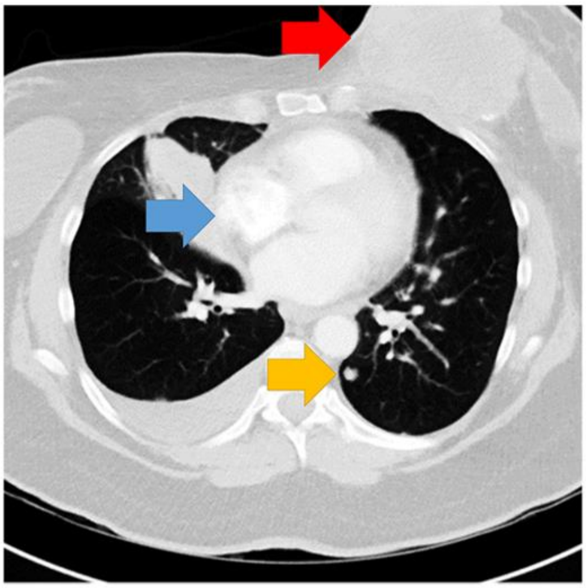

Figure 1: Annotated Image of the Inattentional Blindness CT Stack (note: the breast cancer mass is indicated by the red arrow, lymphadenopathy by the blue arrow, and a lung nodule by the yellow arrow. These arrows were not visible in the experimental display

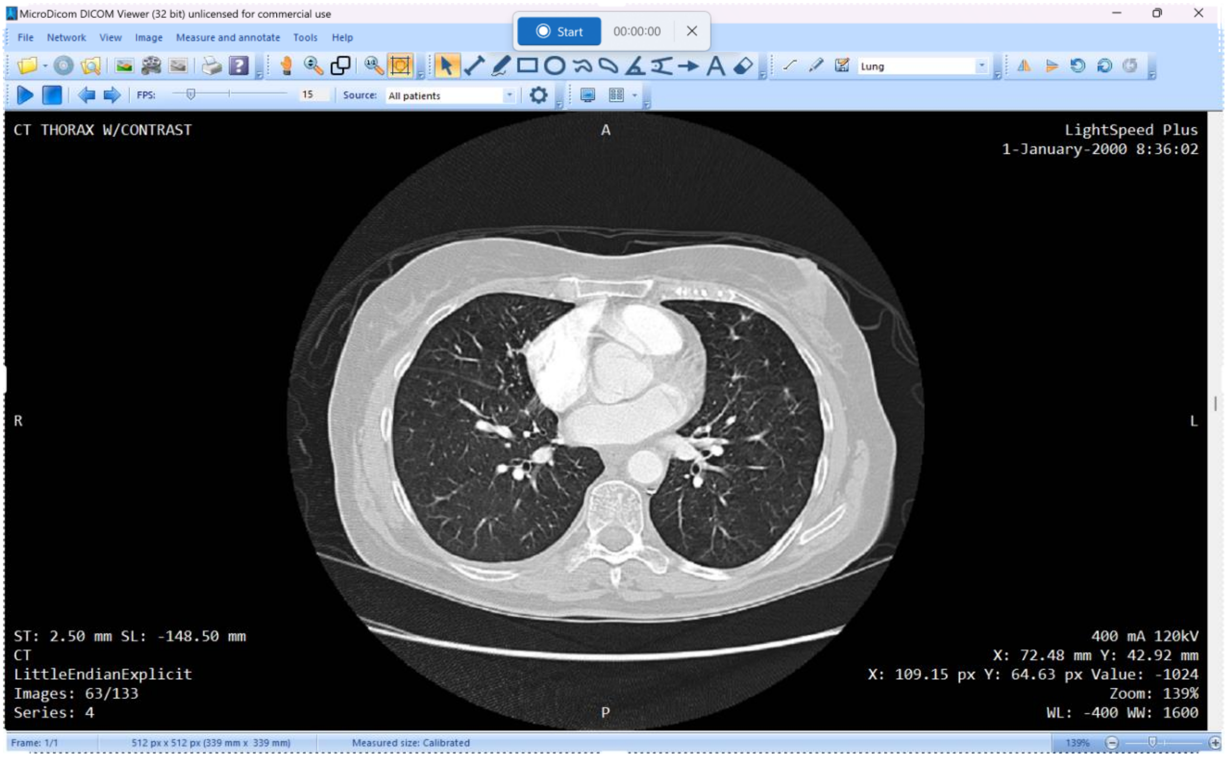

Figure 2: A screenshot from CT Stack 1 illustrating the WL and WW settings (note: The displayed image illustrates commonly used lung windowing parameters (WL = 400, WW = 1600)).

Concise medical checklist

The checklist (Table 1) was developed in consultation with an expert radiologist with over 30 years of clinical and research experience, along with strategies for minimising potential misdiagnoses (Busby *et al*., 2018). The questions were designed to be simple yet comprehensive, covering a range of possible findings without the need to enumerate all of them nor significantly increase radiologists' workloads (Cankurtaran *et al.,* 2023).

Table 1: concise medical checklist questions

Question	Response
1. Did you adhere to your primary and secondary search patterns?	Yes / No
2. Did you remember to check your blind spots?	Yes / No
3. Are you satisfied with your searching?	Yes / No

Question 1 was intended to ensure that the radiologist followed a structured and systematic approach during the scan, ensuring that no areas were missed; Question 2 served as a reminder for radiologists to check areas that might not have been immediately visible or were typically overlooked, reducing the chance of inattentional blindness; and Question 3 encouraged self-reflection, prompting radiologists to assess whether they felt confident in their diagnostic process.

Post-task survey

All participants answered a post-task survey (Table 2) that assessed their awareness of the breast cancer mass and lymphadenopathy in CT stack 7. The post-task survey used the same questions as a previous study (Williams *et al*., 2021).

Table 2: post-task survey questions

Question	Response
1. Did the final case seem any different than any of the other trials?	Yes / No
2. Did you notice any other medically relevant findings on the final case	Yes / No
3. Did the final case show signs of breast cancer?	Yes / No
4. Did the final case show signs of lymphadenopathy?	Yes / No

Questions 1 and 2 were presented in the same order for all participants, as in a previous study (Williams *et al.,* 2021). Questions 3 and 4, could not be randomized without displaying the questions on the same page concurrently, which could have impacted the truthfulness of the participants' responses. Instead, they were presented one per page, in the order above.

All materials were originally written in English and translated into Mandarin by a certified translator. A back translation was also conducted using an online platform. This was performed by a member of the research team (C.L.) whose native language is English and who was not fluent in Mandarin, allowing for an unbiased comparison.

Procedure

Data collection took place at participants' workplaces. All study measures, including the pre-task survey, concise medical checklist and post-task survey, were completed electronically on participant smartphones using Qualtrics XM (Qualtrics, Provo, UT). After providing consent, participants completed the pre-task survey. Next, participants reviewed each of the seven chest CT stacks sequentially (one practice and six experimental) marking all nodules. Participants in the experimental group completed the concise medical checklist after reviewing each CT stack before moving on to the next one. After completing the CT task, all participants answered the post-task survey. Finally, participants received a debrief sheet, and the researcher (C.L.) was available to answer any questions.

Data analysis

To assess if the concise medical checklist would improve lung nodule detection by facilitating a more vigilant search strategy (hypothesis 1), continuous variables, including lung nodule detection accuracy and task duration, were examined. The Mann-Whitney U test was applied to these variables, as the data distribution did not meet the assumptions for parametric tests. To assess if the checklist increased the detection rate of incidental findings (hypothesis 2), categorical variables were analysed. These included participants’ responses to the checklist questions (‘yes/’no’) and whether they noticed signs of breast cancer or lymphadenopathy in the final case (‘yes/’no’). A chi-square test was used to analyse these categorical data. Between-groups effect sizes were interpreted as $d = .1$ for a small effect size, $d = .3$ for a medium effect size and $d = .5$ for a large effect size (Fritz *et al.,* 2012). All statistical analyses were carried out using SPSS (version 28).

Results

Participant demographics are summarised in Table 3.

Table 3: participant demographics

	Control (n = 14)	**Experimental (n = 18)**
Gender (male / female; n (%)	6 male (43) / 8 female (57)	5 male (28) / 13 female (72)
Profession (n (%))	10 Residents (71) / 4 Attendings (29)	16 Residents (89) / 2 Attendings (11)
Specialisation (%)	Abdominal Radiology (21) / Neuroradiology (21) / Thoracic Radiology (14) / Musculoskeletal Radiology (7) / Other (36)	Abdominal Radiology (39) / Breast Imaging (17) / Neuroradiology (11) / Interventional Radiology (6) / Other (28%)
Age (years; M / SD)	27.79 (8.92)	28.00 (7.30)
Experience (years; M / SD)	4.64 (9.06)	3.39 (6.51)
Scans/week (years; M / SD)	181.43 (105.01)	180.00 (105.01)

Abbreviations: F: female; M: mean; SD: standard deviation

Task duration

The task duration was equivalent across the groups (experimental: M = 22.82 mins, SD = 8.00 mins; control: M = 22.17 mins, SD = 4.90 mins; p > .05).

Lung nodule detection accuracy

When including the practice case (CT1) with CT stacks 4, 6, and 7, the mean nodule identification score for the control group was 30.57 (SD = 6.13) out of 49 (62%), while the experimental group scored M = 27.16 (SD = 7.53) out of 49 (55%). The Mann-Whitney U test indicated no significant between-group difference (U = 82.50, Z = -1.65, p > .05), accompanied by a medium effect size (r = -.29). Excluding the practice case yielded similar results (control group: M = 25.50, SD = 5.04) out of 41 (62%), experimental group: (M = 22.88, SD = 6.03) out of 41 (55%), U = 80.50, Z = -1.73, p > .05, r = -.30). Performance on CT stacks 1 and 4 was lower than on stacks 6 and 7, with both groups scoring similarly on the inattentional blindness stack (CT7), as shown in Table 4.

Table 4: control and experimental group lung nodule detection scores

	CT Stack 1	**CT Stack 4**	**CT Stack 6**	**CT Stack 7**
Control (M / SD)	5.07 (1.43)	10.43 (3.00)	8.07 (2.01)	7.00 (0.87)
Control (%)	63	45	80	88
Experimental (M / SD)	4.28 (2.13)	8.28 (3.83)	7.61 (2.20)	7.00 (1.02)
Experimental (%)	54	35	76	88

Abbreviations: CT: computed tomography; SD: standard deviation

Additional markings

All participants made additional markings on the CT stacks that were not classified as lung nodules according to the reference data, including duplicate markings, visibility of lung nodules across multiple image slices, lung nodules ≤ 3mm, and other suspected abnormalities. The experimental group made numerically more additional markings compared with the control group across five CT stacks, although all group differences were not statistically significant (all *p*-values >.05; Table 5).

Table 5: control and experimental group additional markings (lung nodule) detection scores

	Control (n = 14)		**Experimental (n = 18)**			
Stack	Mean	SD	Mean	SD	*p*-value	Effect size (*r*)
CT1 (8)	3.50	1.74	4.78	3.82	.47	.12
CT2 (0)	0.64	1.15	1.06	2.10	.93	.01
CT3 (0)	1.86	1.70	2.28	5.37	.24	.20
CT4 (23)	5.93	3.73	5.00	5.41	.15	.25
CT5 (0)	1.14	1.87	0.28	0.75	.09	.29
CT6 (10)	13.07	7.45	17.39	11.84	.31	.17
CT7 (8)	2.36	2.56	2.39	3.97	.66	.07

Note: the number of lung nodules for each CT stack is provided in brackets
Abbreviations: CT: computed tomography; SD: standard deviation

Concise medical checklist

The experimental group's frequency of and percentage of responses to the medical checklist questions are shown in Table 6.

Table 6: Concise Medical Checklist responses

	Question 1		Question 2		Question 3	
Stack	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
CT1 (*n* / %)	17 (94)	1 (6)	12 (67)	6 (33)	16 (89)	2 (11)
CT2 (*n* / %)	17 (94)	1 (6)	14 (78)	4 (22)	13 (72)	5 (28)
CT3 (*n* / %)	15 (83)	3 (16)	14 (78)	4 (22)	13 (72)	5 (28)
CT4 (*n* / %)	17 (94)	1 (6)	16 (89)	2 (11)	17 (94)	1 (6)
CT5 (*n* / %)	16 (89)	1 (6)	15 (83)	2 (11)	11 (61)	6 (33)
CT6 (*n* / %)	16 (89)	2 (11)	15 (83)	3 (17)	17 (94)	1 (6)
CT7 (*n* / %)	18 (100)	0 (0)	16 (89)	2 (11)	17 (94)	1 (6)

Notes: Question 1: Did you adhere to your primary and secondary search patterns?; Question 2: Did you remember to check your blind spots?; Question 3: Are you satisfied with your searching? One participant did not complete the medical checklist for CT stack 5.
Abbreviations: CT: computed tomography

Post-task survey

Question 1: Did the final case seem any different than any of the other trials?

In the experimental group, 94% of participants reported that the last case appeared different, while it was 86% in the control group. A chi-square test revealed no statistically significant difference between the groups, with a small effect size ($\chi^2(1) = .70$, $p > .05$, $V = .14$).

Question 2: Did you notice any other medically relevant findings on the final case?

In the experimental group, 78% of participants did not report seeing signs of breast cancer, while in the control group, 79% did not report seeing signs. A chi-square test revealed no statistically significant difference between the groups, with a negligible effect size ($\chi^2(1) = .00$, $p > .05$, $V = .01$).

Question 3: Did the final case show signs of breast cancer?

In the experimental group, 78% of participants did not report seeing signs of breast cancer, while in the control group, 79% did not report seeing signs. A chi-square test analysis revealed no statistically significant difference between the groups, with a negligible effect size ($\chi^2(1) = .00$, $p > .05$, $V = .01$).

Question 4: Did the final case show signs of lymphadenopathy?

Fifty-eight percent of participants in the experimental group and 50% of the control group noticed signs of lymphadenopathy. One response was missing for this question in the experimental group. A chi-square test revealed no statistically significant difference between the groups, with a negligible effect size ($\chi^2(1) = .24$, $p > .05$, $V = .08$).

Discussion

This study aimed to assess whether a concise medical checklist could mitigate inattentional blindness among radiologists when interpreting chest CT stacks. The findings suggest that, contrary to the hypotheses, the medical checklist did not lead to any noticeable benefits in the detection of lung nodules (hypothesis 1), nor breast cancer symptoms (hypothesis 2).

Notably, the control group had a higher mean identification score compared with the experimental group (62% vs. 55%), though the difference was not statistically significant. The performance of both groups was also similar to the findings in a previous study (Williams *et al.,* 2021). Considering the slightly higher detection rate in the control group, the checklist used by the experimental group may have potentially had a negative impact on performance. For instance, the experimental group tended to make more additional markings, but at the same time, detected numerically fewer lung nodules. This warrants further investigation to explore the checklist's influence on detection outcomes.

Equally, the checklist showed no superior performance in detecting the lymphadenopathy or breast cancer. This was despite the majority of participants reporting a perceived difference in the last CT stack (control group 85%, experimental group 94%). Nearly 80% of participants in both groups failed to notice the breast cancer mass, a rate higher than the 66% reported in a previous study (Williams *et al.,* 2021). It is possible that the checklist primed radiologists in the experimental group to anticipate something different about the final CT stack; however, it failed to yield favourable outcomes in detecting the incidental findings.

While the checklist appeared to influence the behaviour of some participants, this effect was not universally observed across the experimental group. Although the responses indicated a gradual shift from “no” to “yes” this may have been due to socially desirable responding. A small minority still reported not adhering to their primary search patterns nor checking their blind spots as they progressed through the CT stacks. This suggests the checklist may have been treated as a checkbox exercise by some participants. It could also indicate satisfaction of search bias, which refers to the tendency for some radiologists to curtail their search efforts after identifying one or several abnormalities. While this can be seen as a strategy to conserve cognitive resources, it can also leave room for oversight (Berbaum, 1990; Busby *et al.,* 2018).

A potential explanation for these non-significant findings could be that the checklist was not sensitive enough. A more explicit and systematic checklist may have resulted in improved performance, as seen in Williams *et al*.'s (2021) second study, where 97% of radiologists noticed the breast cancer mass when they were specifically prompted to check for it. Yet, more detailed checklists can be time-consuming (Kok *et al.* 2017). Given the multitude of potential abnormalities, a comprehensive systematic checklist would likely necessitate categorisation based on importance, ultimately complicating the already rocky landscape of incidental findings (see Booth *et al*., 2016). Moreover, participants in this study reported reading an exceptionally high number of scans per week, rendering a lengthy checklist impractical as a clinically useful tool. In a previous study (Williams *et al*., 2021), radiologists reported reading an average of 41 CT scans per week. In contrast, radiologists in this study read approximately 180 scans per week. This discrepancy is likely due to the excessive patient workloads in China (Li & Xie, 2013).

Based on the performance results above, it raises the question of whether the concise medical checklist was actually competing for the radiologists' attention instead of aiding it. It may have inadvertently added to their cognitive demands by introducing an additional task. This diversion from the primary task may have obscured not only their ability to find lung nodules but also their capacity to detect other possible abnormalities. This aligns with Grissinger's perspective on medication errors and inattentional blindness, in which the author dismisses the use of "*error reduction strategies*" (Grissinger, 2012, p. 542), which would include the use of an intervention like a checklist. Instead, Grissinger (2012) advocates two methods: those that seek to minimize potential distractions and those that seek to enhance the visibility of important information. This is because splitting our attention may increase errors, irrespective of whether one of the things dividing our attention is actually a reminder to remain visually attentive, as was the case in this study.

Regarding Grissinger's (2012) first point on minimising distractions, it is well-established that multitasking often compromises task performance because when we attempt to multitask, we merely end up switching between tasks. This overwhelms the demands placed on the neurocognitive systems that support and control sustained attention, making tasks longer to complete (Madore & Wagner, 2019). Similarly, it may be that the introduction of the medical checklist increased both the perceptual load — the number of items that needed to be attended to — and the cognitive load, increasing the difficulty of the primary task (Matias *et al*., 2022). It may have also heightened uncertainty, as reflected in the higher number of additional markings in the experimental group. This may have made it more challenging for observers to detect the unexpected stimuli, as they juggled lung nodule detection and completing the checklist, which may have been further compounded by having to jump between electronic devices (computer and smartphone). This is supported by evidence indicating that distractions from a primary task — whether walking, driving, or using a mobile phone — can increase inattentional blindness (Chen *et al*., 2016; Strayer *et al*., 2011). Additionally, the implementation of distraction-free practices and zones in the administration of medication among nurses, has shown a significant reduction in errors of up to almost 80% in some instances (Connor *et al*., 2016; Westbrook *et al*., 2017). Thus, the cognitive and perceptual demands of managing multiple tasks, such as marking lung nodules and completing the checklist, may not yield benefits in performance when compared with not using the checklist. If this is the case, it could render the concise checklist defunct, or worse, unintentionally lead to reduced awareness and decreased sensitivity to unexpected stimuli, making it detrimental to performance.

To address this, one approach to sustaining attention on a primary task while minimizing internal noise and remaining receptive to unexpected stimuli is said to be mindfulness. This method shifts the focus from relying on external aids to fostering an internal strategy for combating inattentional blindness (see Schofield *et al.,* 2015; Burton *et al*., 2016). Some support for this comes from a controlled pilot study among young neurosurgeons, which revealed that those who underwent an eight-week mindfulness-based intervention programme had a lower incidence of inattentional blindness errors compared with a control group (Pandit *et al*., 2022).

Regarding to Grissinger's (2012) second point on enhancing the visibility of potentially important stimuli, success in the field of medical imaging inevitably depends on developments in artificial intelligence and machine learning. This is because, while changes in medical imaging technology have become more advanced, the shift from X-rays to more advanced imaging like CT scans, the human brain and eye remain unchanged (Robinson, 1997). Artificial intelligence is already being utilised in some hospitals to assist radiologists. These systems typically highlight potential regions of interest in a bright colour for the radiologists to review, with some detecting lung nodules even years in advance (Lovelace Jr *et al*., 2023). The performance of such tools in the field of lung nodule detection is often equivalent to or better than human observers (Gu *et al*., 2021a, 2021b; Li *et al*., 2022). However, future studies are essential to explore their full potential with respect to inattentional blindness, as increased luminescence is unlikely to eliminate inattentional blindness because it still requires human verification, ultimately leaving it at risk of oversight despite brightly coloured potential abnormalities appearing unmissable.

This study has several limitations, most notably the small sample size. A larger sample is essential to thoroughly assess the efficacy of various medical checklists among radiologists when examining CT stacks. Some scholars have highlighted the need for larger sample sizes to ensure the generalizability and reliability of findings in radiological research (Blackmore, 2001). Additionally, the absence of eye-tracking data limited the ability to correlate participants' observations of the breast cancer mass with their gaze patterns, as in a previous study (Williams *et al.,* 2021). The translation of the written materials is also a further limitation. Although they were translated by a certified translator, the back translation was conducted by a member of the research team rather than by someone unfamiliar with the study's purpose. Lastly, despite the considered methodology used for calculating lung nodule detection accuracy, it was done manually, introducing the potential for human error, namely inattentional blindness. The impact of manual error on diagnostic accuracy is well-documented in the literature (Berbaum *et al*., 1990; Kim & Mansfield, 2014).

Overall, the findings suggest that a concise medical checklist may not effectively reduce inattentional blindness among radiologists evaluating chest CT scans. The checklist showed no clear benefits in marking lung nodules and may even lead to similar or worse performance than not using it. Instead, efforts might be better focused on helping radiologists reduce internal noise and finding ways to enhance the importance of unexpected stimuli. However, further research is needed before fully dismissing the potential of checklists in addressing inattentional blindness in medical imaging.

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and/or its supplementary materials.

References

Alagha, H. Z., Badary, O. A., Ibrahim, H. M., & Sabri, N. A. (2011). Reducing prescribing errors in the paediatric intensive care unit: an experience from Egypt. *Acta Paediatrica, 100*(10), 169–174. https://doi.org/10.1111/j.1651-2227.2011.02270.x

Armato, S. G., 3rd, McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., Van Beeke, E. J., Yankelevitz, D., Biancardi, A. M., Bland, P. H., Brown, M. S., Engelmann, R. M., Laderach, G. E., Max, D., … Croft, B. Y. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical Physics, 38*(2), 915–931. https://doi.org/10.1118/1.3528204

Berbaum, K. S., Franken, E. A., Jr, Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., Barloon, T. J., Behlke, F. M., Sato, Y., Lu, C. H., & el-Khoury, G. Y. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology, 25*(2), 133–140. https://doi.org/10.1097/00004424-199002000-00006

Berlin, L. (2007). Radiologic errors and malpractice: a blurry distinction. *American Journal of Roentgenology, 189*(3), 517–522. https://doi.org/10.2214/ajr.07.2209

Blackmore, C. C. (2001). The challenge of clinical radiology research. *American Journal of Roentgenology,* 176(2), 327-328. https://doi.org/10.2214/ajr.176.2.1760327

Bressan, P., & Pizzighello, S. (2008). The attentional cost of inattentional blindness. *Cognition, 106*(1), 370–383. https://doi.org/10.1016/j.cognition.2007.03.001

Burton, A., Burgess, C., Dean, S., Koutsopoulou, G. Z., & Hugh-Jones, S. (2016). How effective are mindfulness-based interventions for reducing stress among healthcare professionals? a systematic review and meta-analysis. *Stress and Health, 33*(1), 3–13. https://doi.org/10.1002/smi.2673

Busby, L. P., Courtier, J., & Glastonbury, C. M. (2018). Bias in radiology: the how and why of misses and misinterpretations. *Radiographics, 38*(1), 236–247. https://doi.org/10.1148/rg.2018170107

Cankurtaran, C. Z., Reddy, S., Cen, S. Y., Lei, X., & Walker, D. K. (2023). work-life experience of academic radiologists: food for thought. *Academic Radiology, 30*(4), 579–584. https://doi.org/10.1016/j.acra.2023.01.011

Chabris, C. F., Weinberger, A., Fontaine, M., & Simons, D. J. (2011). You do not talk about Fight Club if you do not notice Fight Club: inattentional blindness for a simulated real-world assault. *I-Perception, 2*(2), 150–153. https://doi.org/10.1068/i0436

Chen, P., Saleh, W., & Pai, C. W. (2016). Texting and walking: a controlled field study of crossing behaviours and inattentional blindness in Taiwan. *Behaviour & Information Technology, 36*(4), 435–445. https://doi.org/10.1080/0144929x.2016.1240234

Connor, J. A., Ahern, J. P., Cuccovia, B., Porter, C. L., Arnold, A., Dionne, R. E., & Hickey, P. A. (2016). Implementing a distraction-free practice with the red zone medication safety initiative. *Dimensions of Critical Care Nursing, 35*(3), 116–124. https://doi.org/10.1097/DCC.0000000000000179

Coolican, H. (2013). *Research methods and statistics in Psychology* (3rd ed.). Routledge. https://doi.org/10.4324/9780203769669

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Quality & Safety, 22* (2), 58–64. https://doi.org/10.1136/bmjqs-2012-001712

Drew, T., Võ, M. L., & Wolfe, J. M. (2013). The invisible gorilla strikes again: sustained inattentional blindness in expert observers. *Psychological Science, 24*(9), 1848–1853. https://doi.org/10.1177/0956797613479386

Ekelund, M., Fernsund, H., Karlsson, S., & Mac Giolla, E. (2022). Does expertise reduce rates of inattentional blindness? a meta-analysis. *Perception, 51*(2), 131 - 147. https://doi.org/10.1177/03010066211072466

Ely, J. W., Graber, M. L., & Croskerry, P. (2011). Checklists to reduce diagnostic errors. *Academic Medicine*, *86*(3), 307–313. https://doi.org/10.1097/ACM.0b013e31820824cd

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behaviour Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141(*1), 2–18. https://doi.org/10.1037/a0024338

Garg, R. K., Ouyang, B., Kocak, M., Bhabad, S., Bleck, T. P., & Jhaveri, M. D. (2022). Inattentional blindness to DWI lesions in spontaneous intracerebral hemorrhage. *Neurological Sciences, 43*(7), 4355–4361. https://doi.org/10.1007/s10072-022-05992-2

Geijer, H., & Geijer, M. (2018). Added value of double reading in diagnostic radiology: a systematic review. *Insights Imaging, 9*, 287–301. https://doi.org/10.1007/s13244-018-0599-0

Grissinger, M. (2012). 'Inattentional blindness': what captures your attention? *Pharmacy and Therapeutics, 37*(10), 542–555.

Gu, D., Liu, G., & Zhang, X. (2021a). On the performance of lung nodule detection, segmentation and classification. *Computerized Medical Imaging and Graphics, 89*(101886), 1-15. https://doi.org/10.1016/j.compmedimag.2021.101886

Gu, Y., Chi, J., Liu, J., Yang, L., Zhang, B., Yu, D., & Zhao, Y. (2021b). A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning. *Computers in Biology and Medicine*, *137* (104806), 1-27. https://doi.org/10.1016/j.compbiomed.2021.104806

Haugen, A. S., Søfteland, E., Almeland, S. K., Sevdalis, N., Vonen, B., Eide, G. E., Nortvedt, M. W., & Harthug, S. (2015). Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Annals of Surgery, 261*(5), 821–828. https://doi.org/10.1097/SLA.0000000000000716

Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A. H., Dellinger, E. P., Herbosa, T., Joseph, S., Kibatala, P. L., Lapitan, M. C., Merry, A. F., Moorthy, K., Reznick, R. K., Taylor, B., Gawande, A. A., & Safe Surgery Saves Lives Study Group (2009). A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England Journal of Medicine, 360*(5), 491–499. https://doi.org/10.1056/NEJMsa0810119

Heinz, S. A., Kwee, T. C., & Yakar, D. (2020). Unread second-opinion radiology reports: a potential waste of health care resources. *American Journal of Roentgenology, 215*(4), 934–939. https://doi.org/10.2214/AJR.19.22662

Ho, A. M., Leung, J., Mizubuti, G. B., Contardi, L. H., Chan, M. T. V., Lo, T., & Lee, A. K. (2017). Inattentional blindness in anesthesiology: a simulation study. *Journal of Clinical Anesthesia, 42*, 36–39. https://doi.org/10.1016/j.jclinane.2017.07.015

Hyman, I. E., Jr, Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology, 24*, 597-607. https://doi.org/10.1002/acp.1638

Jager, G., Futterer, J. J., & Rutten, M. (2014). *Cognitive errors in radiology: "Thinking fast and slow*" [Conference presentation]. European Society of Radiology 2014, Vienna, Austria. https://dx.doi.org/10.1594/ecr2014/C-0899

Kim, Y. W., & Mansfield, L. T. (2014). Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors. *American Journal of Roentgenology, 202*(3), 465–470. https://doi.org/10.2214/AJR.13.11493

Kok, E.M., Abed, A., & Robben, S.G.F. (2017). Does the use of a checklist help medical students in the detection of abnormalities on a chest radiograph? *Journal of Digital Imaging, 30*(6), 726–731. https://doi.org/10.1007/s10278-017-9979-0

Li, Q., & Xie, P. (2013). Outpatient workload in China. *The Lancet, 381*(9882), 1983–1984. https://doi.org/10.1016/s0140-6736(13)61198-8

Li, R., Xiao, C., Huang, Y., Hassan, H., & Huang, B. (2022). Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: a review. *Diagnostics, 12*(2), 298. https://doi.org/10.3390/diagnostics12020298

Lovelace Jr., B., Torres, J. M.D., Kopf, M., & Martin, P. (2023, April 12). *Promising new AI can detect early signs of lung cancer that doctors can't see*. NBC News. https://www.nbcnews.com/health/health-news/promising-new-ai-can-detect-early-signs-lung-cancer-doctors-cant-see-rcna75982

Lumbreras, B., Donat, L. E. C., & Hernández-Aguado, I. (2010). Incidental findings in imaging diagnostic tests: a systematic review. *British Journal of Radiology, 83*(988), 276–289. https://doi.org/10.1259/bjr/98067945

Mack, A. (2003). Inattentional blindness: looking without seeing. *Current Directions in Psychological Science, 12*(5), 180–184. https://doi.org/10.1111/1467-8721.01256

Madore, K. P., & Wagner, A. D. (2019). Multicosts of Multitasking. *Cerebrum, cer-04-19,* 1-8.

Matias, J., Belletier, C., Izaute, M., Lutz, M., & Silvert, L. (2022). The role of perceptual and cognitive load on inattentional blindness: A systematic review and three meta-analyses. *Quarterly Journal of Experimental Psychology, 75*(10), 1844–1875. https://doi.org/10.1177/17470218211064903

Morris, Z., Whiteley, W., Longstreth, W. T., Weber, F., Lee, Y., Tsushima, Y., Alphs, H. H., Ladd, S. C., Warlow, C., Wardlaw, J. M., & Salman, R. A. (2009). Incidental findings on brain magnetic resonance imaging: systematic review and meta-analysis. *BMJ*, *339*(1), 1-7. https://doi.org/10.1136/bmj.b3016

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: the contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science, 12*(1), 9–17. https://doi.org/10.1111/1467-9280.00303

Pammer, K., Raineri, A., Beanland, V., Bell, J., & Borzycki, M. (2018). Expert drivers are better than non-expert drivers at rejecting unimportant information in static driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour, 59,* 389–400. https://doi.org/10.1016/j.trf.2018.09.020

Pandit, A., De Gouveia, M., Horsfall, H. L., Reka, A., & Marcus, H. J. (2022). Efficacy of a Mindfulness-Based Intervention in Ameliorating inattentional blindness amongst young neurosurgeons: a prospective, controlled pilot study. *Frontiers in Surgery, 9*(916228), 1-7. https://doi.org/10.3389/fsurg.2022.916228

Robinson P. J. (1997). Radiology's Achilles' heel: error and variation in the interpretation of the Röntgen image. *The British Journal of Radiology, 70*(839), 1085–1098. https://doi.org/10.1259/bjr.70.839.9536897

Schofield, T. P., Creswell, J. D., & Denson, T. F. (2015). Brief mindfulness induction reduces inattentional blindness. *Consciousness and Cognition, 37*, 63–70. https://doi.org/10.1016/j.concog.2015.08.007

Simons D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences, 4(*4), 147–155. https://doi.org/10.1016/s1364-6613(00)01455-8

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception, 28*(9), 1059–1074. https://doi.org/10.1068/p281059

Simons, D. J., & Schlösser, M. (2017). Inattentional blindness for a gun during a simulated police vehicle stop. *Cognitive Research: Principles and Implications, 2*(37), 1-8. https://doi.org/10.1186/s41235-017-0074-3

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011b). Cognitive distraction while multitasking in the automobile. In Ross, B (Ed). *Psychology of Learning and Motivation* (pp. 29–58). https://doi.org/10.1016/b978-0-12-385527-5.00002-4

Thomassen, Ø., Storesund, A., Søfteland, E., & Brattebø, G. (2014). The effects of safety checklists in medicine: a systematic review. *Acta Anaesthesiologica Scandinavica, 58*(1), 5–18. https://doi.org/10.1111/aas.12207

Westbrook, J. I., Li, L., Hooper, T. D., Raban, M. Z., Middleton, S., & Lehnbom, E. C. (2017). Effectiveness of a 'do not interrupt' bundled intervention to reduce interruptions during medication administration: a cluster randomised controlled feasibility study. *BMJ Quality & Safety, 26*(9), 734–742. https://doi.org/10.1136/bmjqs-2016-006123

Williams, L., Carrigan, A., Auffermann, W., Mills, M., Rich, A., Elmore, J., & Drew, T. (2021). The invisible breast cancer: Experience does not protect against inattentional blindness to clinically relevant findings in radiology. *Psychonomic Bulletin & Review, 28(*2), 503–511. https://doi.org/10.3758/s13423-020-01826-4