**Abstract**

Objective

We reviewed experimental studies of confidence in medical diagnosis: what factors determine clinicians’ confidence in diagnostic decisions, and how confidence affects diagnosis and treatment. The results of 79 empirical studies were synthesised to establish broad findings across medical subdisciplines.

Design

A scoping review of both medical and psychological literature was conducted. We categorised the articles identified by the review according to their methodology and research findings.

Data sources

We systemically searched SCOPUS, MEDLINE, PsycINFO and Global Health, and then performed citation tracking within the references of these initial papers to identify missed articles.

Eligibility criteria

Studies were included if they reported quantitative results from an empirical study in which clinical participants reported their confidence or certainty during a diagnostic decision. Studies comprised a broad set of medical subdisciplines.

Results

3,829 potential articles were identified. 79 articles met the inclusion criteria. Across these articles, confidence was not found to be well-calibrated to true diagnostic accuracy regardless of clinician experience. We organised the articles under two main themes: the determinants of confidence (as opposed to objective diagnostic accuracy) and the uses of confidence later in the patient’s care pathway. On the former, confidence has been found to be affected by a number of factors, including patient case complexity, early diagnostic differentials and contextual factors within the healthcare environment. Factors that affect confidence but not accuracy demonstrate how the two can be decoupled (resulting in overconfidence or underconfidence). On the latter theme, confidence is found to affect further patient testing, medication administration and referral rates.

Conclusions

Results from this review has implications for medical education and practice around diagnostic uncertainty and considerations of work from cognitive psychology. This review culminates in a model based on the literature that demonstrates a differing set of factors that affect diagnostic confidence/certainty and diagnostic accuracy. Such a model informs future work on diagnosis, especially within medical education, on how appropriate confidence can be prompted and communicated amongst clinicians.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Past work notes the prevalence of diagnostic errors and posits a link between such errors and cognitive biases, with one such bias being overconfidence. This study aims to synthesise past work on diagnostic confidence to understand the factors studied that contribute to clinicians’ confidence in their diagnoses, as well as how confidence is used by clinicians during patient care.

WHAT THIS STUDY ADDS

As a result of this scoping review, we have identified that diagnostic confidence and accuracy have separate contributing factors related to the patient, the clinician making the diagnosis and the environmental context. We have also identified the pathways by which confidence, and misplaced confidence, can affect diagnosis and patient care. As a result of these findings, we introduce an integrative model of confidence throughout the patient care process.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

This study prompts future clinical work to focus not only on eliciting calibrated confidence but also on appropriate communication of uncertainty, which would have implications for medical education around how diagnostic reasoning is taught and for the workplace culture of healthcare environments. We also demonstrate the need to better model how confidence evolves over the course of a diagnostic decision rather than as a static quantity.

**INTRODUCTION**

Medical diagnoses are core to a doctor’s job, as their accuracy is crucial to high quality and safe patient treatment and care. As the accuracy of diagnoses are important, research on diagnosis has been grounded in the incidence of errors. A report from the US Institute of Medicine1 concluded that most patients will experience a diagnostic error within their lifetime. Around 32% of clinical errors have been found to be caused by clinician assessment, particularly the clinician’s failure to weigh up competing diagnoses2. Studies have also investigated the downstream consequences of diagnostic errors. Unnecessary treatments (or ‘overtreatment’) was estimated to cost the US healthcare system between 158 and 226 billion dollars in 20113. Diagnostic errors have also been found to lead to longer hospital stays and increased patient mortality4.

Heuristics are commonly used with diagnostic decisions. For example, making a diagnosis may involve considering a hypothesis as likely because the displayed symptoms seem to correspond with a prototypical case of a particular condition5. A clinician may have recently experienced a patient with a particular condition and, upon seeing another patient with what are perceived to be similar symptoms, is then more likely to diagnose that patient with the same condition again6. While these heuristics are often effective, associated cognitive biases are thought to be linked to diagnostic error. One cognitive bias that has been posited as a contributor of diagnostic error is overconfidence7, which leads to a failure to consider alternative diagnoses and inappropriate weighting of clinicians within group settings.

Confidence is viewed within the cognitive psychology literature as subjective assessments of a decision’s quality or accuracy8. Confidence can track objective accuracy quite well but this is often imperfect, as it manifests in overconfidence and underconfidence. We refer to confidence as being ‘calibrated’ if it closely predicts objective accuracy (i.e. such that the individual is neither overconfident nor underconfident). There is a rich literature within the field of cognitive psychology on the mechanisms of confidence. Confidence has been thought of as the relative strength of evidence in favour of one decision alternative compared to others. However, there is also evidence of a ‘positive evidence bias’, in that more evidence in favour of a chosen option increases confidence regardless of the evidence against that option.

Confidence is especially important in healthcare, as overconfidence can lead to insufficient consideration of diagnostic alternatives and inadequate care. In the absence of objective feedback, confidence can be used as a marker of how likely someone is to be correct9. In the case of medicine, a lack of clearly communicated feedback can cause clinicians to proceed as if they have received positive feedback. This means that they do not adequately update their internal model of the patient and hence they increase their confidence inappropriately10. Confidence is especially important, therefore, in social settings where it has significant sway on group discussions and deliberations. In a group setting, confident members tend to be listened to more, which can lead others astray if they are overconfident, but groups with more calibrated confidence are higher in accuracy and more effective at deliberation.

We conducted a scoping review in order to collate and synthesise the current work on studying diagnosis as a cognitive process, both in terms of what contributes to appropriate (or ‘calibrated) confidence/certainty and how confidence/certainty is utilised within the wider medical decision making process. Our full research questions can be found in Box 1.

**BOX 1: Scoping Review Research Questions (Preregistered)**

Primary questions:

* **RQ1:** How calibrated are the confidence/certainty judgements made during diagnostic decisions by clinicians relative to their actual accuracy?
* **RQ2:** How are confidence/certainty judgements utilised within the wider diagnostic decision process?

Subsidiary questions:

* **RQ3:** What are the prevalent ways in which diagnostic confidence and certainty are operationalised as variables?
* **RQ4:** What strategies, tools or frameworks have been used to prompt better calibration of both confidence and certainty?
* **RQ5:** What types of empirical procedures/tasks are used to study confidence and certainty in diagnostic decisions? Do they come to different conclusions?
* **RQ6:** What are the discrepancies between the concepts/research questions studied in the context of confidence and certainty in the cognitive psychology literature and the medical diagnosis literature?
* **RQ7:** What areas of research are still underexplored within the context of medical diagnosis?

**METHODS**

**Search Strategy**

This review used a protocol that was preregistered on the Open Science Framework: <https://osf.io/wz5se>. We conducted a systematic scoping review of empirical studies of confidence and certainty in medical diagnosis. We primarily consulted JBI’s PRIMSA-ScR Checklist for Scoping Reviews15. The search strategy was designed in cooperation with a subject specialist librarian at the University of Oxford’s Bodleian Libraries group. The search string comprised of keywords to capture an intersection of four elements: confidence/certainty, medical diagnoses, decision making and a study population of medical staff/students (i.e. clinicians, physicians, doctors and medics). The full search terms can be found in Box 2. The databases SCOPUS, MEDLINE, PsycINFO and Global Health were searched during February 2024. Finally, we hand-searched the citations of the included articles from these databases for further relevant articles via backward and forward scanning (Tranfield et al., 2003; Webster & Watson, 2002).

**BOX 2: Search Terms**

(clinicians OR physicians OR doctors OR medics)

AND

( confiden\* OR uncertain\* OR certain\*)

AND

( diagnosis AND medical )

AND

( decision OR ( decision AND making ) OR decision-making )

**Study Selection**

The inclusion criteria for studies were as follows: (1) original empirical studies with quantitative results, (2) written in the English language, (3) experimental paradigm uses medical diagnostic decisions, (4) confidence or certainty is measured as a dependent variable. We exclude editorials, review papers and opinion papers, though we do include dissertations. We do not exclude any medical subdisciplines and do not exclude papers based on publication date. Identified articles were uploaded onto Rayyan, which detects duplicate papers for manual checking and removal.

**Research Synthesis**

The included papers were reviewed to answer the set of research questions that can be found in Box 2. In addressing these questions, papers were first categorised by their broad research methodology (i.e., patient vignettes, in situ questionnaires, etc.) and their medical population of study (e.g., medical students, general practitioners/primary care physicians, medics etc.). We reviewed the experimental procedures to extract their key manipulations/independent variables. This included case complexity, use of a cognitive intervention and medical experience or knowledge on the part of study participants. We also extracted dependent variables as they pertain to confidence/certainty and, where relevant, recording of both diagnostic differentials and information seeking. Each of the paper’s key findings were summarised and then all findings were categorised under recurring themes.

**RESULTS**

**Findings of Scoping Review**

The initial search returned a total of 3,332 articles. Of these, 675 duplicate articles were removed after manual review, leaving 2,717 articles for screening. After exclusions based on initial review of the articles’ titles and abstracts, 165 articles remained for review of the full text. When applying the inclusion criteria described in the Study Selection section, 50 eligible articles were included. Based on these, 439 articles were retrieved for review from the included articles’ citations. After applying both exclusions of duplicates and our inclusion criteria, 29 further articles were identified. This produced a total of 79 articles for inclusion and synthesis. The PRISMA diagram for this search process can be found in Figure 1.

A diagram of a flowchart

Description automatically generated

**FIGURE 1 – PRISMA Diagram of Literature Review**

**Broad Characteristics of Studies**

A summary of the included studies are found in Table 1. Our included studies were published between 1991 and 2024. A histogram showing the distribution of publications over time can be found in Figure 2. We especially note that 36 of the 79 studies (46%) were published from 2019 onwards, indicating a recent surge of research interest in this field and the timeliness of such a scoping review. The studies were published over 60 different publications, including both medical and psychological journals. 19 studies were published in a journal related to medical education, making it the most common research area. Other research areas most represented were Primary Care/General Practice, Emergency Medicine and Nursing.

**Methodologies**

When examining the research methodologies, 44 studies (56%) involved the use of clinical patient text vignettes in their study design, making it the most used experimental task. Other experimental methodologies include the use of imaging (e.g. ECG, X-Rays, MRI) for diagnosis, questionnaires administered in situ to measure confidence during real patient cases as they are happening, and high-fidelity simulations (either digitally or using a mannequin). It is conceivable however that these methodologies differ in ways aside from how naturalistic they are. One included study in this review found that nurses were both less accurate and less confident in a high-fidelity simulation compared to a paper-based vignette72, hinting at limitations in generalising less naturalistic paradigms (e.g. vignettes) to how clinicians would behave in their everyday medical practice.

Another aspect of these methodologies worth noting is the operationalisation of both diagnostic confidence and accuracy (and then by extension, how calibration is operationalised). Studies are mostly aligned on the use of a self-reported scale for confidence, be it a scale from 1-10 or 1-100. The use of such scales is common within cognitive psychology and such measured confidence values are often found to predict other behavioural variables of uncertainty, such as the tendency to seek further information or to opt out of making a decision. 24 studies (30%) allowed participants to input multiple diagnostic differentials rather than a single diagnosis. Confidence is then either measured for each differential or in the set of differentials as a whole. Accuracy is easier to operationalise as a single diagnosis (whether it is correct or not), but it is less naturalistic to how clinicians may consider competing diagnoses in their everyday practice. Hence, measures of how calibrated confidence judgements are to true diagnostic accuracy are heavily contingent on how diagnoses are recorded, which has a bearing on how reliable findings on overconfidence/underconfidence are.

|  |  |  |  |
| --- | --- | --- | --- |
| **Publication Year** | | **Subdiscipline / Population** | |
| 1991-2000 | 10 | Primary Care / General Practice | 26 |
| 2001-2010 | 11 | Medical Students | 15 |
| 2011-2020 | 29 | Emergency Medicine | 10 |
| 2021- | 29 | Nursing | 6 |
|  |  | Pathology | 4 |
| **Methodology** | | Radiology | 4 |
| Textual Vignette | 44 | Other | 14 |
| Imaging Interpretation (e.g. ECG) | 20 |  |  |
| In Situ Questionnaires/Surveys | 13 |  |  |
| High-Fidelity Simulation | 2 | **Total** | **79** |

**TABLE 1: Broad Characteristics of Included Studies**

A graph of a number of years

Description automatically generated with medium confidence

**FIGURE 2 – Distribution of Papers by Publication Year**

34 studies (43%) looking at medical experience or training’s effect on confidence, either measured as a dependent variable or by recruiting participants in either a ‘novice’ and ‘experienced’ group. 24 studies (30%) investigated how participants evaluated diagnostic differentials by allowing participants to record multiple differentials in their diagnosis. 19 studies (24%) manipulated the complexity or difficulty of the patient case (excluding the studies that made use of real patient cases in situ). Finally, 10 studies (13%) measured or manipulated information seeking to study the extent to which the presentation of information affected diagnostic confidence.

**Methodologies**

**Research Themes**

Miscalibration of Confidence and Certainty Judgements to Objective Accuracy

On the whole, there is limited evidence of calibrated confidence judgements, either in terms of underconfidence16-18 or overconfidence19-21. We note that calibration was investigated in studies using vignettes in the vast majority of cases, as there is an established ground truth in each case (unlike in situ studies involving real patients) to compare the participants’ responses to in order to gauge accuracy. Calibration is then calculated in a similar manner for a large amount of these studies: participants report their confidence/certainty in their diagnosis and this value is compared to their true accuracy. As vignettes are quick and simple to administer, participants can complete several diagnoses during a single study such that both their confidence and accuracy can be averaged across cases.

Calibration is affected by a number of factors. Studies have found that calibration is affected by the complexity or difficulty of the presented case22-24. When confidence judgements are not sensitive to the difficulty or complexity of the case, confidence stays fairly constant for difficult cases whilst accuracy decreases, leading to increased overconfidence (and decreased calibration). In past studies, complexity is manipulated by either presenting patient cases with more comorbid conditions23 or by showing more conflicting information about the patient to indicate multiple possible conditions25. Calibration may also be improved by the presence of feedback during a training period26-27.

Contextual factors that pertain to the situated medical environment can also affect confidence, as found using more naturalistic paradigms. For example, clinicians may be constantly interrupted for other tasks28, which can especially happen during busier shifts where they have to manage more patients29 and are not present on rounds when a patient is previously discussed by staff30. Studies that emulated each of these instances found lower diagnostic confidence, though because the researchers for these studies do not obtain measures of accuracy due to working in situ, we can only determine how these contextual factors affect confidence rather than calibration.

Both calibration and case complexity seems to heavily interact with experience such that more experienced clinicians are better able to pick up on when a case is more complex/difficult and adjust their confidence accordingly25,40. We note that calibration interacts with experience such that underconfidence tended to be exhibited by less experienced clinicians (or students), whilst overconfidence was observed more for experienced clinicians. However, a difference in calibration across experience was not always observed in the results41-42. Looking at the link between calibration and experience alone may be too simplistic. There are other aspects of experience that influence diagnoses. Experienced clinicians were found to be less likely to ‘distort’ neutral information to be confirmatory of existing beliefs43. One study found that whilst experienced clinicians were not more accurate in their diagnoses, they were more willing to switch and request more information44 (see below section on Uses of Confidence).

A subset of past work has also manipulated the manner in which information presented to clinicians during the diagnostic process affected confidence. Higher confidence was found when clinicians were presented with redundant patient information60 and when given all available patient information rather than having to gather information themselves61. Clinicians were also found to be more confident when presented with an Electronic Health Record of the patient alongside other information62 and when presented patient history first rather than out of order63, indicating a positive effect of a complete patient history available early on in a case on confidence. An erroneous patient history has also been found to cue both novice and experienced clinicians to incorrect diagnoses whilst confidence remained relatively high, resulting in overall overconfidence64.

A minor theme of past work has been a distinction between experience (operationalised as years of experience or role seniority) and knowledge (measured using standardised tests of medical knowledge). In medical students, the calibration of confidence judgements were found to improve with years of education but not with medical45. The aforementioned information ‘distortion’ was found to affect novice clinicians more43 and lower knowledge was found to be related to higher susceptibility to irrelevant, distracting features of a patient46. However, the latter study found that medical knowledge was not directly associated with confidence.

**BOX 3: Papers on Imaging and Confidence**

There were also a subset of papers that find an increase in confidence when providing clinicians with specialised imaging for a patient when making diagnoses, be they MRI scans31-32, CT scans33, evacuation proctography34 or photos of wounds35. Meanwhile, another subset of papers use various forms of computer-aided decision support systems with the goal of improving confidence, with mixed results36-39. These results are perhaps not surprising, but do warrant addressing as per our inclusion criteria.

Interventions at the Point of Generating Differentials

A subset of past work asked clinicians to generate diagnostic differentials based on patient information. This process of generating diagnostic differentials has then been subject to experimental manipulations and interventions (such as early diagnostic suggestions) to investigate their effect on accuracy and confidence. One can imagine the applicability of this work, for instance when a clinician is transitioning care over a patient to another clinician whilst giving an introduction to the case. A general theme of this work is that there is a tendency toward higher weighting of early information, indicating a primacy effect. Early diagnostic suggestions have been found to be highly influential in the subsequent decision process by clinicians finding these suggestions difficult to ignore and increasing their confidence with them50-51. This also affects the breadth of differentials considered, with fewer differentials considered when provided these suggestions52 and an underweighting of differentials if they were not considered early on in the diagnostic process53. This is where interventions aim to mitigate this tendency by asking clinicians to explicitly consider alternatives, which increases their accuracy and calibration54, or prompt the consideration of the patient’s ‘red flags’ to consider in diagnoses, which was found to increase confidence on simpler cases but not accuracy55. These interventions seemingly have to be explicit however, as simply asking clinicians to reflect on their decision without guidance56-57 or participate in an educational training course58-59 does not seem to suffice in improving diagnostic accuracy and calibration.

Uses of Confidence

With more naturalistic studies, it is possible to isolate ways in which confidence and certainty are utilised within the wider diagnostic process, especially as healthcare involves transition of care between multiple clinicians and departments. Past work has attempted to establish a link between confidence and further seeking of patient information and tests with mixed results. Doctors with lower confidence were found to be linked to more test orders29 whilst more metacognitively aware pathologists (i.e. who tended to report confidence judgements that were closer to their true accuracy) were found to be more likely to request further tests when they were unsure42. Confidence has also been linked to prescribing medication, though overtreatment of unnecessary medications was found to be linked to both underconfidence47 and overconfidence21. Higher confidence has also been linked to referral rates to other specialists in other departments48 and to a lower willingness to admit mistakes18. Lower confidence has also been found to result in less specific diagnoses for patients in situ49. We note that whilst past work on confidence has examined its role within groups (as discussed in the Introduction), only one included article looked at confidence in group decisions and found that a multidisciplinary panel was more confident and better calibrated than a single clinician. We return to this later as a theme for future work to investigate further.

Framework for Diagnostic Decisions

We synthesise the included findings into a theoretical framework to illustrate how various factors distinctly impact diagnostic confidence and accuracy. This framework is shown below in Figure 3. We especially note that the framework both summarises the existing research as described here and proposes directions for future research that has been relatively untapped by the included studies. Namely, the diagnostic decisions that are emulated experimentally view diagnosis as linear processes whilst future work could focus on better simulating the cyclical nature of diagnosis, such as how feedback is integrated into clinicians’ existing knowledge and how confidence and information seeking interact with each other. On the latter, past work has focused on how confidence is related to further testing and information seeking, but not how information seeking itself is related to confidence. Whilst past research has, as a whole, also hinted at a distinction between medical experience and knowledge as they pertain to confidence, there has been little work comparing and contrasting the two directly. This is especially pertinent given the social influence that seniority can have within a group, reducing the likelihood of more junior clinicians speaking up about potential errors in the presence of more experienced clinicians65.

A diagram of a medical procedure

Description automatically generated

**FIGURE 3: Theoretical Framework. The dark boxes show stages of the diagnostic decision process as they proceed over time. The black arrows indicate when a factor impacts the target. The green dashed arrows show links between factors that have been identified as recommendations for future work.**

**DISCUSSION**

**Strengths and Limitations**

The strength of this review is in its scope. Whilst similar past reviews have focused on either mapping instances of cognitive biases within medical errors66-68 or on medical uncertainty more broadly69-70, our review is the first to comprehensively map out the literature that links confidence and certainty to medical diagnoses. Our work demonstrates that both errors and confidence are fruitful areas to study within diagnosis, as they both determine how calibrated a clinician is when expressing certainty/uncertainty. Our review also highlights the recent interest in this field of work, which may be related to the increased focus on artificial intelligence in healthcare, particular for diagnosis, and how uncertainty is also important for any such tools to communicate as well. As such, there is interest in understanding where and how diagnostic uncertainty arises. The breadth of the literature reviewed here, in terms of the large number of involved medical subdisciplines, demonstrates how important confidence is as an area of study in terms of its broad applicability across healthcare. This review is, to our knowledge, the first of its kind in terms of its broad remit to underscore the importance of studying diagnostic confidence across medicine. Given the wide range of subdisciplines that have studied confidence, there is then a need for a wider focus on confidence and metacognition within medical education. Findings from metacognition are already being used to inform educational practices outside of medicine to improve students’ memory retention.

Our forward and backward citation search identified 37% of our sources, which is rather high but is related to the aforementioned broad applicability of diagnostic confidence. A challenge with systematically reviewing the literature is that certain medical subpopulations were missed from the initial search, including residents, interns, dermatologists, surgeons, pathologists and medical students. As mentioned, we do not limit this review to any medical subdisciplines in order to capture a more broad set of diagnostic decisions but the search query was not comprehensive enough for this. In addition, several studies found during the reference search did measure confidence/certainty but not as a primary variable of interest. This meant that confidence was not prominently mentioned, instead focusing on diagnostic accuracy or calibration. Finally, some studies framed the experimental task as diagnostic ‘interpretations’ or ‘assessments’, meaning that our search query based on ‘decisions’ did not capture these studies. With these findings in mind, future searches should modify the query to include more medical subpopulations and the use of ‘self-monitoring’ and ‘calibration’ as dependent variables whilst excluding the narrow focus on decision making.

**Implications and Future Clinical Research**

This scoping review shows the importance and the (particularly recent) surge in interest in diagnostic confidence. Whilst confidence has been linked to diagnostic error in the past7, studying it requires insights from cognitive psychology to inform medical education and practice71. How clinicians evaluate their decisions contributes to their effectiveness, as an overconfident clinician may overlook diagnostic possibilities, delay treatment or ignore crucial information. Conversely, an underconfident clinician may be less likely to speak up in a group about potential errors. Our review finds that confidence and accuracy are rarely aligned during diagnoses. Whilst cognitive interventions such as considering alternative diagnoses and guided reflections have been tested, there is yet to be a standardised cognitive framework to teach non-technical skills such as expression of uncertainty. Notably from these papers, miscalibration of confidence is not only a function of social and environmental factors, as such miscalibration was also observed for vignette studies performed by individual participants. Such factors only serve to amplify systematic tendencies toward misaligned confidence/certainty. When coupled with the aforementioned uses of confidence within the patient’s care pathway, a misplaced sense of confidence can have a large impact on patients, including the increased likelihood of insufficient care or longer hospital stays due to missed diagnoses. It is perhaps from these downstream effects of confidence that errors are more likely to arise, rather than simply from diagnoses themselves. Studying errors has tended to the focus of past work, but studying confidence is similarly fruitful, as both lines of work are important to understand calibration within diagnosis.

Whilst we have reviewed a wide breadth of the available literature on confidence/certainty in diagnoses, there are still many avenues for future work. Namely, the majority of the studies presented here do not study diagnosis as a constantly ongoing and evolving process. In reality, diagnoses feature a back and forth between seeking information and evaluating that information in the context of currently considered diagnostic possibilities. This is likely a function of the methodologies used where participants are asked for diagnoses at a particular point in time, usually after reading all the available information on a patient. Vignette studies are of course much easier for researchers to control for confounding variables in order to standardise the diagnoses made for all participants. More naturalistic, in situ methods can be especially fruitful. However, for in situ studies, interrupting clinicians to report their diagnostic thinking can be intrusive on their work. Hence, future research should first focus on utilising other methodologies for capturing the diagnostic reasoning of clinicians as it evolves with time and the receipt of new information. This could include getting clinicians to think aloud as they make diagnoses73 or using a visual representation of clinicians’ thought processes to capture paths and sources of diagnoses54.

We also note that the included studies have not sufficiently studied how confidence or uncertainty is communicated in groups, even though group discussion is an essential part of healthcare. For instance, past literature in psychology has identified a tendency for individuals within a group to not share information that is unknown to the rest of the group (‘hidden information’), and focusing more on information that is already widely known (referred to ‘shared information bias’). A group’s tendency towards this bias is moderated by the confidence of its constituent members. In addition, future work could seek to identify how the certainty that one has in their mind differs from the confidence that one has to communicate to others (be they patients or other clinicians) in their diagnoses. Clinicians may modify how they communicate certainty with others, especially given the collaborative nature of healthcare and social benefits of communicating opinions with confidence in order to be listened to in a group12. The included papers have also not looked at individual differences in expressions of confidence, where past work from cognitive psychology has found individual systematic tendencies toward higher or lower confidence. Hence, whilst task-level or environmental factors affect confidence and calibration, individual clinicians may also have trait-level factors that are predictive too.

While in situ paradigms can be difficult to administer, there is however considerable value in understanding how contextual and environmental factors can have an impact on confidence and to what extent they are more or less impactful than factors pertaining to the patient or the clinician themselves. The clinician is unlikely to be working alone in any given instance. They are part of a wider team, department and healthcare system, all of which can impact confidence on an individual level. Understanding the cognition of diagnosis for the individual can then be applied further to make changes to the wider medical context to support accurate and calibrated decisions.

**CONCLUSIONS**

Through this scoping review, we found that confidence is frequently not calibrated to accuracy during diagnostic decisions. We also found across the literature that different factors affect confidence and accuracy separately, which may help to explain why such instances of overconfidence or underconfidence are observed. Finally we identified several papers that underscore how confidence affects the subsequent care pathway of patients. Taken together, these findings have implications for how diagnostic certainty and confidence should be studied in future clinical work, including the role that information gathering and interpretation has on diagnoses and usage of naturalistic paradigms.

**REFERENCES**

1. McGlynn EA, McDonald KM, Cassel CK. Measurement is essential for improving diagnosis and reducing diagnostic error: a report from the Institute of Medicine. [Jama](https://pubmed.ncbi.nlm.nih.gov/26571126/). 2015 Dec 15;314(23):2501-2.
2. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, Elstein AS, Hasler S, Kabongo ML, Krosnjar N, Odwazny R. Diagnostic error in medicine: analysis of 583 physician-reported errors. [Archives of internal medicine](https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/1108559). 2009 Nov 9;169(20):1881-7.
3. Berwick DM, Hackbarth AD. Eliminating waste in US health care. [Jama](https://jamanetwork.com/journals/jama/article-abstract/1148376). 2012 Apr 11;307(14):1513-6.
4. Hautz WE, Kämmer JE, Hautz SC, Sauter TC, Zwaan L, Exadaktylos AK, Birrenbach T, Maier V, Müller M, Schauber SK. Diagnostic error increases mortality and length of hospital stay in patients presenting through the emergency room. [Scandinavian journal of trauma, resuscitation and emergency medicine](https://link.springer.com/article/10.1186/s13049-019-0629-z). 2019 Dec;27:1-2.
5. Restrepo D, Armstrong KA, Metlay JP. Annals clinical decision making: avoiding cognitive errors in clinical decision making. [Annals of internal medicine](https://www.acpjournals.org/doi/abs/10.7326/M19-3692). 2020 Jun 2;172(11):747-51.
6. Mamede S, van Gog T, van den Berge K, Rikers RM, van Saase JL, van Guldener C, Schmidt HG. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. [Jama](https://jamanetwork.com/journals/jama/article-abstract/186585). 2010 Sep 15;304(11):1198-203.
7. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. [The American journal of medicine](https://www.sciencedirect.com/science/article/pii/S0002934308000405). 2008 May 1;121(5):S2-3.
8. Fleming SM, Daw ND. Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. [Psychological review](https://psycnet.apa.org/fulltext/2016-60724-003.html). 2017 Jan;124(1):91.
9. Price PC, Stone ER. Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. [Journal of Behavioral Decision Making](https://onlinelibrary.wiley.com/doi/abs/10.1002/bdm.460). 2004 Jan;17(1):39-57.
10. Jaspan O, Wysocka A, Sanchez C, Schweitzer AD. Improving the relationship between confidence and competence: implications for diagnostic radiology training from the psychology and medical literature. [Academic Radiology](https://www.sciencedirect.com/science/article/pii/S1076633220306991). 2022 Mar 1;29(3):428-38.
11. Bach DR, Dolan RJ. Knowing how much you don't know: a neural organization of uncertainty estimates. [Nature reviews neuroscience](https://www.nature.com/articles/nrn3289). 2012 Aug;13(8):572-86.
12. Zarnoth P, Sniezek JA. The social influence of confidence in group decision making. [Journal of Experimental Social Psychology](https://www.sciencedirect.com/science/article/pii/S0022103197913263). 1997 Jul 1;33(4):345-66.
13. Bang D, Ershadmanesh S, Nili H, Fleming SM. Private–public mappings in human prefrontal cortex. [Elife](https://elifesciences.org/articles/56477). 2020 Jul 23;9:e56477.
14. Pouget A, Drugowitsch J, Kepecs A. Confidence and certainty: distinct probabilistic quantities for different goals. [Nature neuroscience](https://www.nature.com/articles/nn.4240). 2016 Mar;19(3):366-74.
15. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil, H. Chapter 11: Scoping Reviews (2020 version). In: Aromataris E, Munn Z (Editors). , JBI, 2020. Available JBI Manual for Evidence Synthesis from <https://synthesismanual.jbi.glo>
16. Mann D. The Relationship between Diagnostic Accuracy and Confidence in Medical Students. [ERIC](https://eric.ed.gov/?id=ED358110)
17. Yang H, Thompson C, Bland M. The effect of clinical experience, judgment task difficulty and time pressure on nurses’ confidence calibration in a high fidelity clinical simulation. [BMC medical informatics and decision making](https://link.springer.com/article/10.1186/1472-6947-12-113). 2012 Dec;12:1-9.
18. Brezis M, Orkin-Bedolach Y, Fink D, Kiderman A. Does Physician's Training Induce Overconfidence That Hampers Disclosing Errors?. [Journal of Patient Safety](https://journals.lww.com/journalpatientsafety/abstract/2019/12000/does_physician_s_training_induce_overconfidence.10.aspx). 2019 Dec 1;15(4):296-8.
19. Friedman C, Gatti G, Elstein A, Franz T, Murphy G, Wolf F. Are clinicians correct when they believe they are correct? Implications for medical decision support. In [MEDINFO 2001](https://ebooks.iospress.nl/volumearticle/19530) (pp. 454-458). IOS Press.
20. Fernández‐Aguilar C, Martín‐Martín JJ, Minué Lorenzo S, Fernández Ajuria A. Use of heuristics during the clinical decision process from family care physicians in real conditions. [Journal of Evaluation in Clinical Practice](https://onlinelibrary.wiley.com/doi/full/10.1111/jep.13608). 2022 Feb;28(1):135-41.
21. Garbayo LS, Harris DM, Fiore SM, Robinson M, Kibble JD. A metacognitive confidence calibration (MCC) tool to help medical students scaffold diagnostic reasoning in decision-making during high-fidelity patient simulations. [Advances in Physiology Education](https://journals.physiology.org/doi/full/10.1152/advan.00156.2021). 2023 Mar 1;47(1):71-81.
22. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians’ diagnostic accuracy, confidence, and resource requests: a vignette study. [JAMA internal medicine](https://jamanetwork.com/journals/jamainternalmedicine/article-abstract/1731967). 2013 Nov 25;173(21):1952-8.
23. Hausmann D, Kiesel V, Zimmerli L, Schlatter N, von Gunten A, Wattinger N, Rosemann T. Sensitivity for multimorbidity: The role of diagnostic uncertainty of physicians when evaluating multimorbid video case-based vignettes. [PloS one](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0215049). 2019 Apr 10;14(4):e0215049.
24. Li S, Zheng J, Lajoie SP. The relationship between cognitive engagement and students’ performance in a simulation-based training environment: an information-processing perspective. [Interactive Learning Environments](https://www.tandfonline.com/doi/abs/10.1080/10494820.2020.1848879). 2023 Apr 3;31(3):1532-45.
25. Brannon LA, Carson KL. Nursing expertise and information structure influence medical decision making. [Applied Nursing Research](https://www.sciencedirect.com/science/article/pii/S0897189703000788). 2003 Nov 1;16(4):287-90.
26. Kuhn J, van den Berg P, Mamede S, Zwaan L, Bindels P, van Gog T. Improving medical residents’ self-assessment of their diagnostic accuracy: does feedback help?. [Advances in Health Sciences Education](https://link.springer.com/article/10.1007/s10459-021-10080-9). 2022 Mar;27(1):189-200.
27. Staal J, Katarya K, Speelman M, Brand R, Alsma J, Sloane J, Van den Broek WW, Zwaan L. Impact of performance and information feedback on medical interns' confidence–accuracy calibration. [Advances in Health Sciences Education](https://link.springer.com/article/10.1007/s10459-023-10252-9). 2023 Jun 17:1-7.
28. Soares III WE, Price LL, Prast B, Tarbox E, Mader TJ, Blanchard R. Accuracy screening for ST elevation myocardial infarction in a task-switching simulation. [Western Journal of Emergency Medicine](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6324702/). 2019 Jan;20(1):177.
29. Gupta AB, Greene MT, Fowler KE, Chopra VI. Associations Between Hospitalist Shift Busyness, Diagnostic Confidence, and Resource Utilization: A Pilot Study. [Journal of Patient Safety](https://journals.lww.com/journalpatientsafety/abstract/2023/10000/associations_between_hospitalist_shift_busyness,.5.aspx?context=latestarticles). 2023 Oct 1;19(7):447-52.
30. Bergl PA, Shukla N, Shah J, Khan M, Patel JJ, Nanchal RS. Factors influencing diagnostic accuracy among intensive care unit clinicians–an observational study. [Diagnosis](https://www.degruyter.com/document/doi/10.1515/dx-2023-0026/html). 2024 Feb 19;11(1):31-9.
31. Mackenzie R, Dixon AK, Keene GS, Hollingworth W, Lomas DJ, Villar RN. Magnetic resonance imaging of the knee: assessment of effectiveness. [Clinical radiology](https://www.sciencedirect.com/science/article/pii/S0009926096803400). 1996 Apr 1;51(4):245-50.
32. Albrechtsen SS, Riis RG, Amiri M, Tanum G, Bergdal O, Blaabjerg M, Simonsen CZ, Kondziella D. Impact of MRI on decision-making in ICU patients with disorders of consciousness. [Behavioural Brain Research](https://www.sciencedirect.com/science/article/pii/S0166432821006173). 2022 Mar 12;421:113729.
33. Abujudeh HH, Kaewlai R, McMahon PM, Binder W, Novelline RA, Gazelle GS, Thrall JH. Abdominopelvic CT increases diagnostic certainty and guides management decisions: a prospective investigation of 584 patients in a large academic medical center. [American Journal of Roentgenology](https://ajronline.org/doi/full/10.2214/AJR.10.4467). 2011 Feb;196(2):238-43.
34. Harvey CJ, Halligan S, Bartram CI, Hollings N, Sahdev A, Kingston K. Evacuation proctography: a prospective study of diagnostic and therapeutic effects. [Radiology](https://pubs.rsna.org/doi/abs/10.1148/radiology.211.1.r99mr16223). 1999 Apr;211(1):223-7.
35. Sanger PC, Simianu VV, Gaskill CE, Armstrong CA, Hartzler AL, Lordon RJ, Lober WB, Evans HL. Diagnosing surgical site infection using wound photography: a scenario-based study. [Journal of the American College of Surgeons](https://www.sciencedirect.com/science/article/pii/S1072751516315460). 2017 Jan 1;224(1):8-15.
36. Hillson SD, Connelly DP, Liu Y. The effects of computer-assisted electrocardiographic interpretation on physicians' diagnostic decisions. [Medical Decision Making](https://journals.sagepub.com/doi/abs/10.1177/0272989X9501500202). 1995 Jun;15(2):107-12.
37. Berner ES, Maisiak RS. Influence of case and physician characteristics on perceptions of decision support systems. [Journal of the American Medical Informatics Association](https://academic.oup.com/jamia/article/6/5/428/808785). 1999 Sep 1;6(5):428-34.
38. Dreiseitl S, Binder M. Do physicians value decision support? A look at the effect of decision support systems on physician opinion. [Artificial intelligence in medicine](https://www.sciencedirect.com/science/article/pii/S0933365704001071). 2005 Jan 1;33(1):25-30.
39. Neugebauer M, Ebert M, Vogelmann R. A clinical decision support system improves antibiotic therapy for upper urinary tract infection in a randomized single-blinded study. [BMC Health Services Research](https://link.springer.com/article/10.1186/s12913-020-5045-6). 2020 Dec;20:1-0.
40. Tabak N, Bar-Tal Y, Cohen-Mansfield J. Clinical decision making of experienced and novice nurses. [Western Journal of Nursing Research](https://journals.sagepub.com/doi/abs/10.1177/019394599601800505). 1996 Oct;18(5):534-47.
41. Yang H, Thompson C. Nurses’ risk assessment judgements: A confidence calibration study. [Journal of Advanced Nursing](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2648.2010.05437.x). 2010 Dec;66(12):2751-60.
42. Clayton DA, Eguchi MM, Kerr KF, Miyoshi K, Brunyé TT, Drew T, Weaver DL, Elmore JG. Are Pathologists Self-Aware of Their Diagnostic Accuracy? Metacognition and the Diagnostic Process in Pathology. [Medical Decision Making](https://journals.sagepub.com/doi/full/10.1177/0272989X221126528). 2023 Feb;43(2):164-74.
43. Kostopoulou O, Russo JE, Keenan G, Delaney BC, Douiri A. Information distortion in physicians’ diagnostic judgments. [Medical Decision Making](https://journals.sagepub.com/doi/abs/10.1177/0272989x12447241). 2012 Nov;32(6):831-9.
44. Krupat E, Wormwood J, Schwartzstein RM, Richards JB. Avoiding premature closure and reaching diagnostic accuracy: some key predictive factors. [Medical education](https://asmepublications.onlinelibrary.wiley.com/doi/abs/10.1111/medu.13382). 2017 Nov;51(11):1127-37.
45. Hautz WE, Schubert S, Schauber SK, Kunina‐Habenicht O, Hautz SC, Kämmer JE, Eva KW. Accuracy of self‐monitoring: does experience, ability or case difficulty matter?. [Medical education](https://asmepublications.onlinelibrary.wiley.com/doi/abs/10.1111/medu.13801). 2019 Jul;53(7):735-44.
46. Mamede S, Zandbergen A, de Carvalho-Filho MA, Choi G, Goeijenbier M, van Ginkel J, Zwaan L, Paas F, Schmidt HG. Role of knowledge and reasoning processes as predictors of resident physicians’ susceptibility to anchoring bias in diagnostic reasoning: a randomised controlled experiment. [BMJ Quality & Safety](https://qualitysafety.bmj.com/content/early/2024/02/16/bmjqs-2023-016621.abstract). 2024 Feb 16.
47. Levin PD, Idrees S, Sprung CL, Weissman C, Weiss Y, Moses AE, Benenson S. Antimicrobial use in the ICU: indications and accuracy—an observational trial. [Journal of hospital medicine](https://shmpublications.onlinelibrary.wiley.com/doi/abs/10.1002/jhm.1964). 2012 Nov;7(9):672-8.
48. Calman NS, Hyman RB, Licht W. Variability in consultation rates and practitioner level of diagnostic certainty. [J Fam Pract](https://www.researchgate.net/profile/Neil-Calman/publication/21858297_Variability_in_Consultation_Rates_and_Practitioner_Level_of_Diagnostic_Certainty/links/57961ea008aec89db7b84cfd/Variability-in-Consultation-Rates-and-Practitioner-Level-of-Diagnostic-Certainty.pdf). 1992 Jul 1;35(1):31-8.
49. Hageman MG, Bossen JK, King JD, Ring D. Surgeon confidence in an outpatient setting. [Hand](https://journals.sagepub.com/doi/abs/10.1007/s11552-013-9533-6). 2013 Dec;8(4):430-3.
50. Kämmer JE, Schauber SK, Hautz SC, Stroben F, Hautz WE. Differential diagnosis checklists reduce diagnostic error differentially: a randomised experiment. [Medical education](https://asmepublications.onlinelibrary.wiley.com/doi/pdf/10.1111/medu.14596). 2021 Oct;55(10):1172-82.
51. Kourtidis P, Nurek M, Delaney B, Kostopoulou O. Influences of early diagnostic suggestions on clinical reasoning. [Cognitive Research: Principles and Implications](https://link.springer.com/article/10.1186/s41235-022-00453-y). 2022 Dec 15;7(1):103.
52. Staal J, Speelman M, Brand R, Alsma J, Zwaan L. Does a suggested diagnosis in a general practitioners’ referral question impact diagnostic reasoning: an experimental study. [BMC Medical Education](https://link.springer.com/article/10.1186/s12909-022-03325-7). 2022 Apr 8;22(1):256.
53. Eva WK. The influence of differentially processing evidence on diagnostic decision-making ([Doctoral dissertation](https://macsphere.mcmaster.ca/handle/11375/7187)).
54. Feyzi-Behnagh R, Azevedo R, Legowski E, Reitmeyer K, Tseytlin E, Crowley RS. Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. [Instructional science](https://link.springer.com/article/10.1007/s11251-013-9275-4). 2014 Mar;42:159-81.
55. Chartan C, Singh H, Krishnamurthy P, Sur M, Meyer A, Lutfi R, Stark J, Thammasitboon S. Isolating red flags to enhance diagnosis (I-RED): an experimental vignette study. [International Journal for Quality in Health Care](https://academic.oup.com/intqhc/article/31/8/G97/5607831). 2019 Oct 31;31(8):G97-102.
56. Lambe KA, Hevey D, Kelly BD. Guided reflection interventions show no effect on diagnostic accuracy in medical students. [Frontiers in psychology](https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.02297/full). 2018 Nov 23;9:285916.
57. Costa Filho GB, Moura AS, Brandão PR, Schmidt HG, Mamede S. Effects of deliberate reflection on diagnostic accuracy, confidence and diagnostic calibration in dermatology. [Perspectives on Medical Education](https://link.springer.com/article/10.1007/s40037-019-0522-5). 2019 Aug 1;8:230-6.
58. Benvenuto-Andrade C, Dusza SW, Hay JL, Agero AL, Halpern AC, Kopf AW, Marghoob AA. Level of confidence in diagnosis: clinical examination versus dermoscopy examination. [Dermatologic surgery](https://journals.lww.com/dermatologicsurgery/abstract/2006/05000/level_of_confidence_in_diagnosis__clinical.34.aspx). 2006 May 1;32(5):738-44.
59. Kuhn J, Mamede S, van den Berg P, Zwaan L, van Peet P, Bindels P, van Gog T. Learning deliberate reflection in medical diagnosis: does learning-by-teaching help?. [Advances in Health Sciences Education](https://link.springer.com/article/10.1007/s10459-022-10138-2). 2023 Mar;28(1):13-26.
60. Heller RF, Saltzstein HD, Caspe WB. Heuristics in medical and non-medical decision-making. [The Quarterly Journal of Experimental Psychology Section A](https://journals.sagepub.com/doi/abs/10.1080/02724989243000019). 1992 Feb;44(2):211-35.
61. Gruppen LD, Wolf FM, Billi JE. Information gathering and integration as sources of error in diagnostic decision making. [Medical Decision Making](https://journals.sagepub.com/doi/abs/10.1177/0272989x9101100401). 1991 Dec;11(4):233-9.
62. Ben-Assuli O, Sagi D, Leshno M, Ironi A, Ziv A. Improving diagnostic accuracy using EHR in emergency departments: A simulation-based study. [Journal of biomedical informatics](https://www.sciencedirect.com/science/article/pii/S1532046415000477). 2015 Jun 1;55:31-40.
63. Tio RA, Carvalho Filho MA, de Menezes Mota MF, Santanche A, Mamede S. The Effect of Information Presentation Order on Residents' Diagnostic Accuracy of Online Simulated Patients With Chest Pain. [Journal of graduate medical education](https://meridian.allenpress.com/jgme/article/14/4/475/484936/The-Effect-of-Information-Presentation-Order-on). 2022 Aug 1;14(4):475-81.
64. Fawver B, Thomas JL, Drew T, Mills MK, Auffermann WF, Lohse KR, Williams AM. Seeing isn’t necessarily believing: Misleading contextual information influences perceptual-cognitive bias in radiologists. [Journal of Experimental Psychology: Applied](https://psycnet.apa.org/record/2020-28446-001). 2020 Dec;26(4):579.
65. Hémon B, Michinov E, Guy D, Mancheron P, Scipion A. Speaking up about errors in routine clinical practice: a simulation-based intervention with nursing students. Clinical Simulation in Nursing. 2020 Aug 1;45:32-41.
66. Kostopoulou O, Delaney BC, Munro CW. Diagnostic difficulty and error in primary care—a systematic review. [Family practice](https://academic.oup.com/fampra/article/25/6/400/481603). 2008 Dec 1;25(6):400-13.
67. Graber ML, Kissam S, Payne VL, Meyer AN, Sorensen A, Lenfestey N, Tant E, Henriksen K, LaBresh K, Singh H. Cognitive interventions to reduce diagnostic error: a narrative review. [BMJ quality & safety](https://qualitysafety.bmj.com/content/21/7/535.short). 2012 Jul 1;21(7):535-57.
68. Saposnik G, Redelmeier D, Ruff CC, Tobler PN. Cognitive biases associated with medical decisions: a systematic review. [BMC medical informatics and decision making](https://link.springer.com/article/10.1186/s12911-016-0377-1). 2016 Dec;16:1-4.
69. Hall KH. Reviewing intuitive decision‐making and uncertainty: the implications for medical education. [Medical education](https://asmepublications.onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2923.2002.01140.x). 2002 Mar;36(3):216-24.
70. Bhise V, Rajan SS, Sittig DF, Morgan RO, Chaudhary P, Singh H. Defining and measuring diagnostic uncertainty in medicine: a systematic review. [Journal of general internal medicine](https://link.springer.com/article/10.1007/s11606-017-4164-1). 2018 Jan;33:103-15.
71. Wears RL. Diagnosing diagnosis. [Annals of Emergency Medicine](https://room9er.com/wp-content/uploads/2014/12/diagnosing-diagnosis.pdf). 2014 Dec 1;64(6):586-7.
72. Yang H, Thompson C, Bland M. The effect of clinical experience, judgment task difficulty and time pressure on nurses’ confidence calibration in a high fidelity clinical simulation. [BMC medical informatics and decision making](https://link.springer.com/article/10.1186/1472-6947-12-113). 2012 Dec;12:1-9.
73. Coderre S, Mandin HH, Harasym PH, Fick GH. Diagnostic reasoning strategies and diagnostic success. [Medical education](https://asmepublications.onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2923.2003.01577.x). 2003 Aug;37(8):695-703.