The use of deliberate reflection to reduce confirmation bias among orthopedic surgery residents

O uso da reflexão deliberada para reduzir o viés de confirmação em residentes de ortopedia

Antônio Barbosa Chaves
orcid.org/0000-0001-8289-4084
antonioortohand@gmail.com

Alexandre Sampaio Moura
orcid.org/0000-0002-4818-5425
alexandresmoura@gmail.com

Rosa Malena Delbone de Faria
orcid.org/0000-0001-7740-8408
rmdfmedicina@gmail.com

Ligia Cayres Ribeiro
orcid.org/0000-0002-4277-3066
ligiacayres@gmail.com

Abstract

Introduction: Cognitive biases might affect decision-making processes such as clinical reasoning and confirmation bias is among the most important ones. The use of strategies that stimulate deliberate reflection during the diagnostic process seems to reduce availability bias, but its effect in reducing confirmation bias needs to be evaluated.

Aims: To examine whether deliberate reflection reduces confirmation bias and increases the diagnostic accuracy of orthopedic residents solving written clinical cases.

Methods: Experimental study comparing the diagnostic accuracy of orthopedic residents in the resolution of eight written clinical cases containing a referral diagnosis. Half of the written cases had a wrong referral diagnosis. One group of residents used deliberate reflection (RG), which stimulates comparison and contrast of clinical hypotheses in a systematic manner, and a control group (CG), was asked to provide differential diagnoses with no further instruction. The study included 55 third-year orthopedic residents, 27 allocated to the RG and 28 to the CG.

Results: Residents on the RG had higher diagnostic scores than the CG for clinical cases with a correct referral diagnosis (62.0±20.1 vs. 49.1±21.0 respectively; p = 0.021). For clinical cases with incorrect referral diagnosis, diagnostic accuracy was similar between residents on the RG and those on the CG (39.8±24.3 vs. 44.6±26.7 respectively; p = 0.662). We observed an overall confirmation bias in 26.3% of initial diagnoses (non-analytic phase) and 19.5% of final diagnoses (analytic phase) when solving clinical cases with incorrect referral diagnosis. Residents from RG showed a reduction in the confirmation of incorrect referral diagnosis when comparing the initial diagnosis given in the non-analytic phase with the one provided as the final diagnosis (25.9±17.7 vs. 17.6±18.1 respectively; Cohen d: 0.46; p = 0.003). In the CG, the reduction in the confirmation of incorrect diagnosis was not statistically significant.

Conclusions: Confirmation bias was present when residents solved written clinical cases with incorrect referral diagnoses, and deliberate reflection reduced such bias. Despite the reduction in confirmation bias, diagnostic accuracy of residents from the RG was similar to those from the CG when solving the set of clinical cases with a wrong referral diagnosis.

Keywords: Clinical reasoning, bias, deliberate reflection, medical education, medical resident.

Resumo

Introdução: Os vieses cognitivos podem afetar tanto os processos de tomada de decisão como o raciocínio clínico e o viés de confirmação está entre os mais importantes. O uso de estratégias que estimulam a reflexão deliberada durante o processo diagnóstico parece reduzir o viés de disponibilidade, mas seu efeito na redução do viés de confirmação precisa ser avaliado.

1 Universidade José do Rosário Vellano (UNIFENAS-BH), Belo Horizonte, MG, Brazil.
2 Universidade de São Paulo (FMUSP), São Paulo, SP, Brazil.
Objetivos: examinar se a reflexão deliberada reduz o viés de confirmação e aumenta a acurácia do diagnóstico de residentes de ortopedia ao resolverem casos clínicos escritos.

Métodos: estudo experimental comparando a acurácia diagnóstica de residentes de ortopedia na resolução de oito casos clínicos escritos contendo um diagnóstico de encaminhamento. Metade dos casos escritos tinha um diagnóstico de encaminhamento errado. Um grupo de residentes utilizou a reflexão deliberada (GR), que estimula a comparação e o contraste de hipóteses clínicas de maneira sistemática, e um grupo controle (GC) foi solicitado a fornecer diagnósticos diferenciais sem maiores instruções. O estudo incluiu 55 residentes de ortopedia do terceiro ano, 27 alocados no GR e 28 no GC.

Resultados: residentes no GR tiveram escores diagnósticos mais altos do que o GC para casos clínicos com um diagnóstico de encaminhamento correto (62,0±20,1 vs. 49,1±21,0 respectivamente; p = 0,021). Para os casos clínicos com diagnóstico de encaminhamento incorreto, a acurácia diagnóstica foi semelhante entre os residentes do GR e do GC (39,8±24,3 vs. 44,6±26,7 respectivamente; p = 0,662). Observamos viés geral de confirmação em 26,3% dos diagnósticos iniciais (fase não analítica) e 19,5% dos diagnósticos finais (fase analítica) na resolução de casos clínicos com diagnóstico de encaminhamento incorreto. Os residentes do GR mostraram uma redução na confirmação do diagnóstico de encaminhamento incorreto ao comparar o diagnóstico inicial dado na fase não analítica com aquele fornecido como diagnóstico final (25,9±17,7 vs. 17,6±18,1 respectivamente; Cohen d: 0,46; p = 0,003). No GC, a redução na confirmação do diagnóstico incorreto não foi estatisticamente significativa.

Conclusões: o viés de confirmação esteve presente quando os residentes resolveram casos clínicos escritos com diagnósticos de encaminhamento incorretos e a reflexão deliberada reduziu esse viés. Apesar da redução do viés de confirmação, a acurácia diagnóstica dos residentes do GR foi semelhante à do GC na solução do conjunto de casos clínicos com diagnóstico de encaminhamento incorreto.

Palavras-chave: raciocínio clínico, viés, reflexão deliberada, educação médica, residente de medicina.

Introduction

Clinical reasoning is a critical skill for practicing physicians and is related to their ability to generate diagnoses and make decisions. Evidence shows that there are two different forms of reasoning, one being non-analytical, based on pattern recognition (type 1), and the other being analytical dependent on the application of rules (type 2) (1). Non-analytical reasoning is fast and intuitive and used for generating initial hypotheses and solving routine clinical cases. Analytical reasoning is more conscious, slow, and logical, and therefore more time consuming and more demanding on working memory for its execution. The latter is used in hypothesis checking and seems fundamental in resolving complex and atypical clinical cases, more susceptible to “diagnostic pitfalls” (2). The non-analytical and analytical reasoning systems seem to act simultaneously, with the non-analytical system generating the initial diagnostic hypotheses and the analytical system monitoring them, and even being able to correct erroneous diagnostic hypotheses that have been (3-4).

Initial wrong diagnostic hypotheses may occur more frequently in emergency care due to the need for the rapid definition of the diagnosis and appropriate management. In a study conducted in an emergency department from a general hospital in Italy, misdiagnosis was the main reason for lawsuits involving orthopedic surgeons, especially those dealing with hand, wrist, and shoulder injuries (5). Reinforcing the diagnostic challenge of upper limb injuries, a study conducted in Hong Kong, based on a review of undiagnosed fractures by orthopedists in an emergency department, noted that 25.2% of diagnostic errors were related to upper limb fractures (6).

Diagnostic errors can be classified into three types: (a) non-fault errors, which relate to the complexity of each clinical case, which can be silent, atypical, or mimicking a more typical case; (b) system-related errors, in which the diagnosis is not made or made late due to latent imperfections in the health system and (c) cognitive errors, directly linked to the physician and his training, failures in the collection or interpretation of clinical history, incorrect reasoning or lack of knowledge (7).

Cognitive errors appear to be the leading cause of diagnostic errors and are often the result of cognitive biases (8-10). In a study analyzing 100 diagnostic error cases, cognitive errors were observed in 74% of cases (11). Among such biases is the confirmation bias, which can be defined as a clear tendency to maintain the initially established clinical hypothesis, disregarding existing contrary evidence (13). In other words, confirmation bias is present when clinical reasoning is based only on the initial clinical hypothesis and, in a selective
and biased manner, the information initially provided is accepted, disregarding clinical data that contradicts it (8, 13). It also involves a tendency to selectively search for confirmatory evidence that supports the initial diagnosis, rather than searching for evidence that might refute (14-15).

Several instructional methods have been proposed for reducing cognitive biases in clinical reasoning (16). One such method involves using an instrument that promotes deliberate reflection, guiding the learner to analyze the initial clinical diagnosis searching for evidence that confirms or contradicts it, and contrast it with alternative (4). The use of deliberate reflection, i.e., stimulating analytical reasoning in a structured manner, has led to improved diagnostic accuracy among internal medicine residents in studies that analyzed cognitive error related to availability bias (17, 18).

To the best of our knowledge, there are no studies analysing the effect of deliberate reflection on reducing diagnostic confirmation bias, nor its effect on diagnostic accuracy during the resolution of orthopedic clinical cases. This experimental study aims to evaluate the effect of deliberate reflection on orthopedic residents’ diagnostic accuracy during the resolution of orthopedic upper limb injuries. We sought to assess whether the use of deliberate reflection could improve diagnostic accuracy and reduce confirmation bias.

Methods

Study design

Controlled experimental study with third-year orthopedic residents conducted in a single session where they were asked to solve eight written clinical cases of upper limb injuries. The clinical cases contained referral diagnoses to induce confirmation bias. The residents were not informed that in four cases, this referral diagnosis was plausible but incorrect. Participants were randomly assigned to one of two experimental groups (deliberate reflection vs differential diagnosis) and solved one clinical case at a time.

Participants

Seventy-nine third-year residents from 24 accredited medical residency programs in orthopedics and traumatology in Minas Gerais, Brazil, participate in this study. The medical residency in orthopedics and traumatology in Brazil has a 3-year duration, so the intervention included residents in their final training year. The experiment was conducted in October 2019, in the city of Belo Horizonte, during a preparatory class for board certification. The Brazilian Society of Orthopedics and Traumatology gave full support to the study that was conducted in line with the principles of the Declaration of Helsinki and have been approved by the ethics committee of UNIFENAS (approval number: 18880619.2.0000.5143).

Procedures

• Construction of the written clinical cases

The eight clinical cases contained clinical history data, physical examination, and high quality radiological images. The cases used in this experiment had a set of conditions that represented the most frequent diagnostic errors found in a retrospective analysis of 432 referrals made by orthopedists from emergency care units to an orthopedic hospital (Hospital São Francisco de Assis, Belo Horizonte, Brazil) from April 2018 to March 2019.

A panel of 15 hand surgery specialists were involved in the assessment of the referrals and found eight conditions that were more prone to diagnostic errors: 1) lesion of the scapholunate ligament of the carpus; 2) distal radius fracture; 3) “Terrible triad” of the elbow (elbow fracture-dislocation); 4) elbow radial head fracture; 5) dorsal perilunate dislocation of the carpus (wrist dislocation); 6) fracture of the proximal phalanx of the finger of the hand; 7) scaphoid bone fracture of the wrist; and 8) injury to the radial collateral ligament of the metacarpophalangeal joint of the thumb.

Once the diagnoses chosen were defined, the cases were prepared by one of the study authors (ABC) and validated by two other hand surgery specialists and two orthopaedic residents.
**Intervention**

The eight written clinical cases were presented in a sequential and controlled manner. Participants received written directions on how to solve the cases, which differed according to the experimental group to which they were allocated.

In the first step of clinical cases analysis, participants in both groups were instructed to do a quick reading of the case (1 minute) and provide an initial diagnostic hypothesis (non-analytical step). The directions for the second step, which lasted 5 minutes for each case, differed by experimental group. Residents in the reflection group (RG) were instructed to reanalyse the case using an instrument to promote deliberate reflection as used in previous studies (17, 19-20). Briefly, deliberate reflection consisted of describing the clinical findings that favour or contradict the initial diagnostic hypothesis, as well as expected findings in the case description if this hypothesis held true that were, however, absent. Then the residents were asked to list two other differential diagnoses and repeat the above procedure for each of the hypotheses. Finally, they were asked to contrast the diagnoses and select which one would be the most likely for the case. Residents in the control group (CG) were asked to reread the case and list two differential diagnoses before choosing a final diagnosis with no further instructions.

![Figure 1](image-url)  
*Figure 1.* The flowchart of the study.

**Statistical analysis**

Diagnostic accuracy was computed for each clinical case and a binary system (0 – incorrect; 1 - correct) was used to score the diagnosis given by the participants. Three orthopedic hand surgery scored the diagnoses given for each of the eight clinical cases. Differences in scores between the evaluators were resolved by consensus. A final mean score was calculated and later converted proportionally to a scale of 0 to 100 points. Comparative analyses of diagnostic scores from different experimental condition (RG vs. CG) overall and stratified by case type (correct and vs. incorrect referral diagnosis) were performed using the Mann-Whitney test. The confirmation bias score was calculated based on the sum of confirmations for each group, considering only the cases with wrong referral diagnosis. Intragroup comparison of the confirmation bias score at the two diagnostic moments (initial vs. final) was made using the Wilcoxon test. Between group comparison of confirmation bias scores was performed using the Mann-Whitney test. All results were considered significant if the p value was ≤ 0.05.
Results
Fifty-five (69.6%) of the seventy-nine residents that were invited to participate volunteered to the study, with 27 allocated to the RG and 28 to the CG. The groups were similar regarding age (31.0±4.2 vs. 30.8±3.4 years), respectively; p = 0.854, and male gender (89.2% vs. 96.3%, respectively; p = 0.611).

Overall initial diagnostic accuracy was not statistically different between the RG and CG (50.9±13.4 vs. 46.9±17.9, respectively; p = 0.194). Residents from the RG showed higher final diagnostic accuracy when compared to those in the CG in cases that contained a correct referral diagnosis (62.0±20.1 vs. 49.1±21.0, respectively; p = 0.021). In cases with incorrect referral diagnoses, diagnostic accuracy was similar between the RG and CG (39.8±24.3 vs. 44.6±26.7, respectively, p = 0.662) (Table 1).

TABLE 1 – Total final diagnostic accuracy and by diagnoses of correct and incorrect referrals.

<table>
<thead>
<tr>
<th>Diagnostic accuracy</th>
<th>RG</th>
<th>CG</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases with incorrect referral diagnosis, n (%)</td>
<td>39.8</td>
<td>44.6</td>
<td>0.662</td>
</tr>
<tr>
<td>Cases with correct referral diagnosis, n (%)</td>
<td>62.0</td>
<td>49.1</td>
<td>0.021</td>
</tr>
<tr>
<td>Total, n (%)</td>
<td>50.9</td>
<td>46.9</td>
<td>0.194</td>
</tr>
</tbody>
</table>

*p = Mann–Whitney test.

Analyzing the diagnostic confirmation score in the four cases with incorrect referral diagnosis, we observed that confirmation bias occurred in 26.4% of the diagnoses given after the first step (non-analytical step), and 19.5% of the final diagnoses (after the analytical step). The reduction in confirmation of the incorrect referral diagnosis between the first and second step occurred in both groups but was statistically significant only in the RG (25.9±17.7 vs. 17.6±18.1). Cohen d: 0.46; p = 0.003) (Table 2).

TABLE 2 – Percentage of confirmation of referral diagnosis at different diagnostic moments (initial vs. final), stratified by experimental group (reflection vs. control) and by type of clues (correct vs. incorrect).

<table>
<thead>
<tr>
<th></th>
<th>RG</th>
<th>CG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases with incorrect referral, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial diagnostic confirmation</td>
<td>63.9</td>
<td>49.1</td>
</tr>
<tr>
<td>Final diagnostic confirmation</td>
<td>62.0</td>
<td>47.6</td>
</tr>
<tr>
<td>Cases with correct referral, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial diagnostic confirmation</td>
<td>25.9</td>
<td>26.8</td>
</tr>
<tr>
<td>Final diagnostic confirmation</td>
<td>17.6</td>
<td>21.4</td>
</tr>
</tbody>
</table>

Variables expressed as mean and standard deviation. CG, control group; RG, reflection group. Teste de Wilcoxon. *p = 0.003.
Discussion

In this study we analysed if confirmation bias would affect final-year orthopaedics’ residents diagnostic accuracy regarding diagnosis of upper limb injuries. We also compared the effect of two different diagnostic strategies, deliberate reflection and differential diagnosis, on residents’ accuracy when confirmation bias was present. To that end, we randomly assigned residents to diagnose eight clinical cases with referral diagnostic hypothesis, four of which were plausible but incorrect diagnoses, either through deliberate reflection or differential diagnosis. Confirmation bias emerged as about one-quarter of wrong referral diagnostic hypotheses were confirmed by the residents. Although the final diagnostic scores for these cases did not differ between intervention groups, we observed a statistically significant reduction in confirmation bias among residents that engaged on deliberate reflection relative to those who did differential diagnoses.

Other studies that evaluated the effects of strategies intended to reduce students’ susceptibility to cognitive biases indicated minimal or no effects. A Canadian controlled trial study with senior medical students did not find a significant effect of cognitive strategies that stimulate metacognition on cognitive biases of students when trying to solve clinical cases in various emergency medical specialties, including orthopedics (21). On the other hand, an experimental study conducted by Mamede et al. with first - and second - year internal medicine residents observed a significant effect of deliberate reflection on reducing availability bias (17). More recently, Mamede et al. (18) showed that an intervention directed at increasing knowledge about discriminating clinical features, i.e. clinical findings that help differentiate diseases that have similar clinical presentation, can also decrease susceptibility to availability bias, reducing diagnostic error among second-year internal medicine residents. We found no studies, however, examining the use of deliberate reflection specifically to reduce diagnostic confirmation bias.

The use of a tool with different characteristics than the one used in our study but with a similar purpose, called SLOW (mnemonic for the terms “Sure about that, Look at the data, Opposite, Worst case scenario”) was analyzed in a study conducted with medical students, residents, and experienced physicians. Participants solved cases with different cognitive biases, and the SLOW tool did not improve diagnostic (22). However, it is worth noting that in a sub-analysis, the SLOW tool showed more significant potential to help minimize the effect of confirmation bias compared to other biases. In our study, the confirmation of incorrect diagnoses decreased when residents deliberately reflected upon the cases, suggesting a moderate positive effect of deliberate reflection in reducing confirmation bias. These results suggest that more analytical reasoning strategies could indeed help prevent confirmation bias.

In our experiment, we observed that the diagnostic accuracy of third year orthopaedic residents in solving clinical cases constructed in a way that induces confirmation bias was similar between those who used deliberate reflection and who did differential diagnoses. However, a sub analysis revealed that diagnostic accuracy was higher among the residents in the deliberate reflection group when solving cases with correct referral diagnosis. Thus, for cases with clues in the correct direction, reflection may indeed have a beneficial effect. We cannot rule out the possibility that the difference in this subgroup of cases occurred simply by chance, arising from random error, but it should be noted that the overall beneficial effect of deliberate reflection in resolving cases without diagnostic bias has been previously demonstrated in studies with medical students with different levels of expertise (23, 19, 24, 25).

The lack of overall differences in diagnostic accuracy between the experimental groups could also be explained by the fact that both groups had enough time to reflect upon the cases. The difference between groups in our study was the type of reflection, with one group reflecting in a deliberate and structured way, being encouraged to contrast differential diagnoses systematically, and the other group reflecting in a less structured
way since they were asked only to provide a list of differential diagnoses. It may also be that this structuring of reflection is not as crucial for more advanced learners as it is for undergraduate medical students where the effect of deliberate reflection has mostly been shown (23, 25).

Another explanation for the lack of an overall effect of deliberate reflection may be related to the fact that the learner’s knowledge is more important than the way they reason (26-28). Because our study was randomized, we would not expect overall differences in knowledge between the two groups, and neither intervention provided content that could increase residents’ knowledge. Strategies aimed at increasing specific knowledge through multiple examples seem to be critical in helping medical doctors to differentiate between diseases that appear similar and are promising strategies to “immunize” them against biases in reasoning, such as availability bias (18).

Unfortunately, definitive solutions for cognitive errors do not exist. Further studies are needed to assess the effect of combined educational strategies and tools that encourage students and residents to become aware of these biases, to consider alternative diagnoses, to improve metacognition, to calibrate their diagnosis from immediate feedback and to strengthen their knowledge of discriminating features (14, 16, 26).

Conclusions

In clinical cases in which an incorrect referral diagnosis was provided, the group of orthopedic residents who solved them using deliberate reflection significantly reduced confirmation bias. However, no improvement in overall diagnostic accuracy was observed using deliberate reflection compared to a control group, guided only to provide differential diagnoses. Cognitive biases remain a crucial topic in clinical reasoning and still require future research to identify the best strategies to avoid their occurrence in clinical decision-making.

Notes

This study is part of the result of a Professional Master’s Degree in Health Education conclusion in 2019/2020 at the José Rosário Vellano University (UNIFENAS), by one of the authors (ABC), entitled “Use of structured reflection as an instructional method to reduce diagnostic confirmation bias in orthopedic residents.”

English language version has been made by Sabrina Ramos Gomes, graduated in Letters – Edition Technologies, Doctoral student in Letters by the program in Language Studies at the Federal Center for Technological Education of Minas Gerais (CEFET-MG).

Funding

This study did not receive financial support from external sources.

Conflicts of interest disclosure

The authors declare no competing interests relevant to the content of this study.

Authors’ contributions

All the authors declare to have made substantial contributions to the conception, or design, or acquisition, or analysis, or interpretation of data; and drafting the work or revising it critically for important intellectual content; and to approve the version to be published.

Availability of data and responsibility for the results

All the authors declare to have had full access to the available data and they assume full responsibility for the integrity of these results.

References


Antônio Barbosa Chaves
Orthopedic physician and hand surgeon. Master in Health Education, José do Rosário Vellano University (UNIFENAS), Belo Horizonte, MG, Brazil. Preceptor in Hand Surgery at the Hand Surgery Residency Service at São Francisco Hospital – Santa Lucia Unit, Belo Horizonte, MG, Brazil. Head of the Orthopedics Residency Service at Hospital Belo Horizonte, Belo Horizonte, MG, Brazil.

Alexandre Sampaio Moura
Infectious Diseases physician. Master in Public Health, Columbia University, NY, USA. PhD in Infectious Diseases, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil. Research professor at the Masters in Health Education Program, José do Rosário Vellano University (UNIFENAS), Belo Horizonte, MG, Brazil. Research professor at the Stricto Sensu Postgraduate Course, Faculdade Santa Casa, Belo Horizonte, MG, Brazil.

Rosa Malena Delbone de Faria
Hematologist physician; Master and PhD in Hematology, Escola Paulista de Medicina (EPM), Universidade de São Paulo (UNIFESP), São Paulo, SP. Specialist in Education for Health Professions, Federal University of Ceará (UFCE), Fortaleza, CE, Brazil. Professor at the School of Medicine, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil; Academic Director, Faculdade Santa Casa, Belo Horizonte, MG, Brazil; Professor of the Specialization Course in Health Education, Faculdade de Medicina, Universidade de São Paulo (FMUSP), São Paulo, SP, Brazil.

Ligia Cayres Ribeiro
Physician, School of Medicine, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil. Master in health professions education, Maastricht University (UM), Maastricht, Netherlands. Research professor at the Masters in Health Education Program, José do Rosário Vellano University (UNIFENAS), Belo Horizonte, MG, Brazil.

Mailing address
Antônio Barbosa Chaves
Ortopédico BH
Rua Professor Otávio Coelho Magalhães, 111
Mangabeiras, 30210300
Belo Horizonte, MG, Brasil