BMJ Open

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025973
Article Type:	Research
Date Submitted by the Author:	14-Aug-2018
Complete List of Authors:	Weidenbusch, Marc; University Hospital of LMU Munich, Institute for Medical Education; University Hospital of LMU Munich, Department of Internal Medicine IV Lenzer, Benedikt; University Hospital of LMU Munich, Institute for Medical Education Sailer, Maximilian; LMU Munich, Department of Psychology Strobel, Christian; University Hospital of LMU Munich, Institute for Medical Education Kunisch, Raphael; University Hospital of LMU Munich, Institute for Medical Education Kiesewetter, Jan; University Hospital of LMU Munich, Institute for Medical Education Fischer, Martin; University Hospital of LMU Munich, Institute for Medical Education Zottmann, Jan; University Hospital of LMU Munich, Institute for Medical Education
Keywords:	Undergraduate medical education, Case-based learning, Clinical reasoning, Social interaction, Medical decision making

RESEARCH PAPER

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Dr. med. Marc Weidenbusch[1,2*], *Benedikt Lenzer*[2*], *Prof. Dr. phil. Maximilian Sailer*[3], *Dr. phil. Christian Strobel*[2], *Raphael Kunisch*[2], *Dr. phil. Jan Kiesewetter*[2], *Prof. Dr. med. Martin R. Fischer*[2] *& Dr. phil. Jan M. Zottmann*[2]

[1]Nephrologisches Zentrum, Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, LMU München

[2]Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München, LMU München

[3]Department Psychologie, Ludwig-Maximilians-Universität München

[*]equal contribution

Corresponding author:
Dr. phil. Jan M. Zottmann
Institut für Didaktik und Ausbildungsforschung in der Medizin
Klinikum der Universität München, LMU München
Pettenkoferstr. 8a, D-80336 Munich, Germany
Tel: +49 89 4400 57203
Fax: +49 89 4400 57202
E-Mail: jan.zottmann@med.uni-muenchen.de

Word count: 3070 words

Keywords: undergraduate medical education, case-based learning, clinical reasoning, social interaction, medical decision making

Abstract

Objective: Fostering clinical reasoning is a mainstay of medical education. Based on the Clinicopathological Conferences, we propose a case-based peer teaching approach called Clinical Case Discussions (CCDs) to promote the respective skills in medical students. This study compares the effectiveness of different CCD formats with varying degrees of social interaction in fostering clinical reasoning.

Design, Setting, Participants: A single-center randomised controlled trial with a parallel design was conducted at a German university. The 106 study participants were stratified (age, gender, year of study, prior CCD participation, performance in a pre-test) and tested regarding their clinical reasoning skills right after CCD participation and two weeks later.

Intervention: Participants worked either within a live discussion group (Live-CCD), a group watching recordings of the live discussions (Video-CCD), or a group working with printed cases (Paper-Cases). The presentation of case information followed an admission-, discussion-, summary-sequence.

Primary and secondary outcome measures: Clinical reasoning skills were measured with a knowledge application test addressing the students' conceptual, strategic, and conditional knowledge. Additionally, subjective learning outcomes were assessed.

Results: With respect to learning outcomes, the Live-CCD group displayed the best results, followed by Video-CCD and Paper-Cases. No difference was found between Live-CCD and Video-CCD groups in the delayed post-test; both outperformed the Paper-Cases group. Regarding subjective learning outcomes, the Live-CCD received significantly better ratings than the other formats.

Conclusions: This study demonstrates that the CCD approach is an effective and sustainable clinical reasoning teaching resource for medical students. Subjective learning outcomes underline the importance of learner (inter-)activity in the acquisition of clinical reasoning skills in the context of case-based learning. Higher efficacy of more interactive formats can be attributed to positive effects of collaborative learning. Future research should investigate how the Live-CCD format can further be improved and how video-based CCDs can be enhanced through instructional support.

Article summary

Strengths and limitations of this study:

- Fist empirical study on the implementation of Clinical Case Discussions in undergraduate medical education.
- Comparison of Clinical Case Discussions with differing grades of social interaction to determine their effectiveness on medical students' acquisition of clinical reasoning skills by between-group analyses.
- Implementation of multidimensional and multilayered test instruments in a pre-, post- and delayed post-test design to measure clinical reasoning skills by a knowledge application test and self-assessment.
- The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e., a distinction of conceptual, strategic, and conditional knowledge).
- Despite the large sample size and strict randomisation the ubiquitous selection bias in medical education when predominantly motivated students register voluntarily for trials could have influenced the results of this study.

Introduction

Curriculum developers face the challenge of implementing competence-oriented frameworks such as CanMEDS (Canada), the NKLM (Germany) or PROFILES (Switzerland), including the need to train clinical reasoning skills as a medical doctor's key competence.[1-3] As such, clinical reasoning skills are crucial not only for appropriate medical decision making, but also to avoid diagnostic errors and the associated harm for both patients and healthcare systems.[4]

Case-based learning has been proposed to foster clinical reasoning skills[5] and is well accepted amongst students.[6] Case-based learning found an early representation in Clinicopathological Conferences (CPC, first introduced by Cannon in 1900[7]) which are practiced until today. The Clinicopathological Conferences conducted at the Massachusetts General Hospital are published on a regular basis known as the *Case Records* series of the New England Journal of Medicine. In those CPCs the "medical mystery"[8] presented by the case under discussion calls readers to think about the possible diagnosis themselves, before it is finally disclosed at the last part of the CPC. Despite the absence of definitive evidence for efficacy as a teaching method, CPCs have widely been used in medical education since the early 20th century to foster clinical reasoning.[9-11] While these CPC-Case Records reaches lots of medical readers around the world, it has been criticised as being anachronistic with a diagnosing "star (i.e. the discussant), performing, acutely aware of being the center of attention".[12]

Case-based learning formats are embedded in a context, which is known to promote learning better than providing facts in an abstract, non-contextual form.[13] Merseth proposed three essential elements for cases: "They are real, they rely on careful research and study; and they provide data for consideration and discussion by users".[14]

Compared to case vignettes, elaborated and authentic cases provide increased diagnostic challenge, comprising an additional value for medical training.[15]

However, due to their setup, CPCs are often a passive learning situation for participants, as they listen to the discussant laying out his or her clinical reasoning on the case under discussion. According to the ICAP framework by Chi et al.,[16] teaching formats increase their efficacy from passive < active < constructive < interactive learning environments. Based on the ICAP model, any intervention that would lead to more effective cognitive (i.e. constructive or interactive) learner activities should improve the learning outcomes of that format. Especially when students interactively engage in discussions among each other, learning is enhanced. Accordingly, case-based learning has been found to be particularly beneficial in collaborative settings.[17] However, another important aspect to consider in collaborative learning environments is the potential for social loafing, i.e. mostly passive participation of students.[18] To foster optimal learning effects, students should thus be encouraged to be interactively engaged. One prerequisite to achieve self-guided learning in groups is a low threshold for students to come forward with their questions and participate in ensuing discussions.[19] To this end, peer teaching has been established as an effective tool to stimulate discussions.[20] To make sure peer tutors are not overwhelmed in moderating these discussions, the presence of an experienced clinician appears to be warranted[21] in addition to a specific training of the tutors.

Taken together, while traditional CPCs encompass some important dimensions of effective case-based learning environments, they are not systematically aiming at constructive or interactive learner activities that are known features of effective teaching formats.[16,22] Therefore, we introduced Clinical Case Discussions (CCD) in

undergraduate medical education to account for these features. We still use the Case Records of the Massachusetts General Hospital,[9] as these cases exemplify realistic patient encounters and fulfill the criteria for an interactive collaborative learning process as explained above. In the CCD approach, cases are presented only until the hospital admission of the patient, followed by an interactive discussion about possible diagnoses and diagnostic strategies. After all test results have been discussed, the actual diagnosis is disclosed and the pitfalls and take-home messages of the case are summarised.

To investigate the effectiveness of the CCD approach in undergraduate medical education, we designed an intervention trial and assessed clinical reasoning skills in medical students before and after *participating in* live CCDs or *being exposed* to video recordings of live CCDs. We compared these formats and its effects on clinical reasoning with the more traditional approach of working through written cases. When carrying out this randomised trial, we hypothesised that participation in live CCD sessions would lead to a higher increase of clinical reasoning skills than simply reading the cases. To better understand possible effects of the CCD learning environment with its social components on learning outcomes, participation in live CCDs as outlined above was additionally compared to the effects of watching videos of CCDs online. This comparison also seemed relevant from an economic point of view as video-streaming of lectures and seminars are prevalent at many instituitions in higher education allowing for flexible and scalable access to learning materials.[23]

Methods

Study participants / Ethics

Initially, we recruited 106 volunteer medical students of XXXXX Medical Faculty. Participants were stratified by age, gender, year of study, prior CCD participation and performance in a knowledge application pre-test at T_0. They were then randomly assigned to one of the experimental groups and a total of 90 participants eventually completed the study, 31 of them were male and 59 female. They were 20 to 41 years old (M = 23; SD = 2.97) and in their first to eighth clinical semester (M = 3.5; SD = 1.78). The protocol for the trial was approved by the Institutional Ethics Review Board. Written informed consent was obtained from all study participants and they received a financial reimbursement of 50 Euros upon completion of the trial.

Patient and public involvement

No patients or public were involved in this research.

Clinical Case Discussions

In all experimental groups the intervention was based on the same three, independent internal medicine cases[24-26] which were worked through in an iterative approach in different formats: (a) peer-moderated live case discussions in an interactive setting (Live-CCD, n = 30), (b) a single-learner format utilizing an interactive multimedia platform displaying video recordings of the live case discussions (Video-CCD, n = 27), and (c) a single-learner format in which the students worked with the original paper cases of the NEJM (Paper-Cases, n = 33). The cases were prepared in a way that participants in each format were exposed to the same case information.

In all three groups cases were presented in a specified structured manner similar to the original Clinicopathological Conferences (see Figure 1). In each format the students

("discussants") had to fill out a form after the admission in which the case had to be summarised and a list of clinical problems and working diagnoses had to be provided. Subsequently, between discussion and summary a second case-summary had to be completed in which the final diagnostic test and the most likely diagnosis had to be proposed.

****

Insert Figure 1 about here

****

In the Live-CCD group, the case presentation was prepared beforehand by a voluntary discussant ("presenter"), who presented the facts in the admission (according to the structure shown in Figure 1). Electronic slides and flipcharts were used to transport case information. Original test results were revealed by the presenter during the discussion only when requested by the group of students. Furthermore, the presenter summarised the differential diagnosis, important pathophysiological features of the case at the end of the session and provided a short take home message. A moderating medical student ("moderator") was trained in case presentation and facilitated a reasonable approach to the patient encounter in close communication with the discussants. In the discussion the moderator helped the students develop their diagnostic strategy by co-evaluating their requested findings and the reasoning employed. Supervision of the correctness of medical facts and the correct diagnostic approach were ultimately granted by a clinician who could stop the discussion at any point when faulty reasoning was evident or discussants explicitly requested the facilitation of an experienced physician. We varied the staff between each Live-CCD to minimise effects of personal teacher characteristics. Live sessions typically lasted 90 minutes and were recorded with multiple cameras.

Students in the Video-CCD format worked on a single-learner multimedia workstation on which a video recording of the Live-CCD were displayed. These recordings also contained the electronic slide presentation from the Live-CCD and enabled simultaneous observation of the discussion from multiple camera angles. Participants could pause and partially skip the videos.

In the Paper-Cases group participants received the case information of each CCD section sequentially (i.e. admission, discussion, summary) in a print format. In both single-learner formats students could choose their personal working speed with no time limit. In each of the three formats full access to the internet was permitted for additional information.

Study design

We conducted a single-center randomised controlled trial consisting of a total of five course sessions with a parallel design (see Figure 2): In an introductory session (T_0) participants were introduced to the principles of the CCD approach and the sequence of this trial. In the experimental phase, participants attended three weekly interventional course sessions of 90 minutes each in one of the three aforementioned groups with the respective CCD formats. T_1 testing was carried out at the end of the last experimental course session. Two weeks after completion of the interventional courses a delayed knowledge application post-test (T_2) was conducted.

****

Insert Figure 2 about here

****

Instruments

Learning outcomes with respect to clinical reasoning were measured with a knowledge application test that consisted of 29 items (i.e. a maximum of 29 points could be achieved). The test was to be filled out within 45 minutes and comprised multiple choice items, key feature problems and problem-solving tasks,[27] addressing the conceptual, strategic, and conditional knowledge of the participants. Overall test reliability was satisfactory (Cronbach's $\alpha = .71$).

Subjective learning outcomes were measured at T_1 with a short questionnaire consisting of 9 items (e.g. "I learned a lot during the CCD course", "The CCD course increased my learning " or "I recommend the implementation of the CCD teaching format into the curriculum"). Participants were asked to rate these items on a Likert scale ranging from 1 (I don't agree) to 5 (I fully agree). Reliability of the corresponding scale was good (Cronbach's $\alpha = .95$). Additionally, study participants were asked to share their views on positive and negative aspects of the respective training format through open items at the end of the questionnaire.

Statistical Analysis

The required sample size (N = 128) was estimated to detect medium effect sizes with a power of 80% and a significance level of $\alpha = .05$. For between-group analyses, ANOVAs were conducted with *post-hoc* Bonferroni tests for multiple comparisons.

Results

Preliminary analyses

Prior knowledge (T_0) did not differ across groups, with $M = 5.34$; $SD = 1.93$ for Live-CCD, $M = 4.76$; $SD = 1.90$ for Video-CCD, and $M = 5.76$; $SD = 2.24$ for Paper-Cases with $F(2,87) = 1.78$, $p = .174$ (n.s.). Gender distribution was skewed between the experimental groups due to drop-out, but did not affect the learning outcomes as male students ($M = 12.33$; $SD = 4.25$) and female students ($M = 10.80$; $SD = 3.50$) did not differ significantly in the knowledge application post-test, $F(2,88) = 3.37$, $p = .07$ (n.s.).

Effects of the CCD format on learning outcomes related to Clinical Reasoning

Experimental groups differed significantly with respect to the knowledge application post-test (see Table 1), $F(2,87) = 27.07$, $p = .000$, partial $\eta^2 = .384$. The Live-CCD group ($M = 14.10$; $SD = 3.32$) outperformed both the Video-CCD ($M = 11.69$; $SD = 3.34$) and the Paper-Cases group ($M = 8.5$; $SD = 2.44$). Post hoc Bonferroni tests revealed significant differences between Live-CCD and Video-CCD ($p = .011$) as well as the Paper-Cases group ($p = .000$). The difference in the knowledge application post-test between Video-CCD and the Paper-Cases group was also significant ($p = .000$).

Two weeks after course completion, the effect of the teaching format was still found in a delayed knowledge application post-test, $F(2,87) = 30.91$, $p = .000$, partial $\eta^2 = .415$. Both Live-CCD ($M = 13.36$; $SD = 3.23$) and the Video-CCD ($M = 11.84$; $SD = 2.92$) outperformed the Paper-Cases group ($M = 7.89$; $SD = 2.41$). Post hoc Bonferroni tests revealed significant differences between the Live-CCD and Paper-Cases group ($p = .000$) as well as between the Video-CCD and Paper-Cases group ($p = .000$). However, the difference between Live-CCD and Video-CCD was not significant in the delayed knowledge application post-test ($p = .146$).

Table 1. Overview of the findings of the study.

	Teaching format					
	Live-CCD		Video-CCD		Paper-Cases	
	M	(SD)	M	(SD)	M	(SD)
Knowledge application pre-test	5.34	(1.92)	4.76	(1.90)	5.76	(2.24)
	$n = 30$		$n = 27$		$n = 33$	
Knowledge application post-test	14.10	(3.32)	11.68	(3.34)	8.50	(2.44)
	$n = 30$		$n = 27$		$n = 33$	
Delayed knowledge application post-test	13.36	(3.23)	11.84	(2.92)	7.89	(2.41)
	$n = 30$		$n = 27$		$n = 33$	
Subjective learning outcomes	4.20	(.63)	3.18	(1.24)	3.00	(.99)
	$n = 30$		$n = 27$		$n = 31$	

Effects of the CCD format on subjective learning outcomes

Experimental groups differed significantly with respect to subjective learning outcomes (see Table 1), $F(2,85) = 13.16$, $p = .000$, partial $\eta^2 = .236$. Participants of the Live-CCD group ($M = 4.20$; $SD = .63$) assigned better ratings to their course format than participants in the Video-CCD group ($M = 3.18$; $SD = 1.24$) and the Paper-Cases group ($M = 3.0$; $SD = .99$). Post hoc Bonferroni tests showed that the Live-CCD differed from the Video-CCD ($p = .001$) and the Paper-Cases group ($p = .000$) in this regard. An additional Duncan post-hoc test confirmed that the Video-CCD and the Paper-Cases group did not differ from each other in this regard ($p = .48$).

To investigate the relations between the subjective assessment and the knowledge application tests applied at the end and two weeks after the course, we calculated correlations between the different outcome measures. Subjective learning outcomes correlated on a medium level with both the knowledge application post-test ($r = .343$, n

= 88, p = .001) and the delayed knowledge application post-test (r = .339, n = 88, p = .001).

In the Live-CCD group, 83% of the students were in favour of implementing routine Live-CCD into the medical curriculum. Only 45% and 31% of students from the Video-CCD and Paper-Cases groups voted for an implementation of their respective course in the curriculum. With respect to the open items from the subjective learning outcomes questionnaire, participants from all groups praised the quality of the cases. Participants from the Live-CCD group particularly valued their course format for providing an opportunity to practice "diagnostic thinking" and the "focus on practice elements". They also mentioned that "you can look up theoretical knowledge, but you can't look up applied knowledge". Students in the Video-CCD group, on the other hand, praised features of the digital learning environment as they could "pause, reflect, or quickly do a Google search" when watching the case discussions. However, they also criticised it was not possible for them to "participate in a more active way".

Discussion

This randomised controlled study shows that even relatively short CCD interventions can lead to improved and sustainable learning outcomes with respect to clinical reasoning. This provides evidence that the CCD approach, which is based on Clinicopathological Conferences, is an effective teaching resource to foster clinical reasoning skills in medical students. We had hypothesised that a more interactive course format would result in an improvement of clinical reasoning skills when compared with less interactive formats. Results show that the Live-CCD indeed leads to the highest learning outcomes in medical students compared to less interactive formats. Consistent

with our hypothesis, clinical reasoning skills, as measured with our knowledge application test, had the highest gain in the Live-CCD group. These positive effects of the CCD teaching format on clinical reasoning skills proved sustainable as shown by the results in the delayed knowledge application post-test. Overall, these results are in line with a recently published study on diagnostic reasoning[28] where students who worked in pairs were more accurate in their diagnosis than individual students despite having comparable knowledge. Collaborative clinical reasoning has thus far been underrepresented in the literature, yet seems to solve many of the educational problems regarding diagnostic errors.[29]

The significant difference between the Live-CCD and the Video-CCD group can be explained by the findings of a meta-analysis that showed technology-assisted single-person learning to be inferior to group learning because of the decreased social interaction.[30] However, it is important to note that two weeks after the course, participants of the Live-CCD and Video-CCD groups did not differ significantly anymore while both groups still clearly outperformed the Paper-Case group. In other words, watching a video of the live case discussion was found to be more beneficial for learners regarding their clinical reasoning skills than just reading the printed cases. Subjective learning outcomes suggest that students prefer the live discussion over the other formats and were linked to their performance in the knowledge application test. Additional qualitative data from the open item answers suggests that the Live-CCD format supported students in performing clinical reasoning and that the active discussion of cases was particularly valued by the students.

Generalisability

The conclusions of this study are applicable to a broader audience of medical students. The CCD approach and its respective formats can easily be implemented in routine medical education. Peer teaching courses hold the promise of being more easy to install and more easy to staff then courses led by faculty. The study population consisting of students with heterogeneous levels of clinical experience implies that the CCD is an effective teaching format not only for students at the beginning of their clinical career but also for intermediate students. On the other hand, generalisability is potentially limited as only students from one medical school participated in our study.

Limitations of the study

There are certain limitations of this study that have to be addressed: One important limitation is the single-centre nature of this study and the relatively small sample size. Before the CCD approach can be implemented on a larger scale, a validation of our findings is therefore required. Caution is clearly warranted with the effect sizes shown in this trial, as it has been shown that effect sizes of learning intervention trials tend to be inflated compared to the effectiveness of the intervention when used in routine education.[31] Against this backdrop, we suggest replication to further validate the results found in this study and strengthen the outlined implications.

Implications for policy makers / Future research questions

Based on our findings, the CCD approach is a useful asset for medical educators to widen the range of clinical reasoning teaching tools. Live-CCD can thus be seen as a prime candidate for routine implementation in clinical reasoning curricula. Future research should aim to identify which Live-CCD elements (i.e. the roles, case contents,

or the course structure) contribute in which way to the improvement of clinical reasoning skills in medical students. Regarding the Video-CCD, means of instructional support to increase the effectiveness and interactivity of the video-based format should be investigated in an attempt to exploit its full potential.

Acknowledgements

The authors thank Johanna Huber and her team for technical support with the evaluation, Thomas Brendel and Thomas Bischoff for help with the video production and Mark S. Pecker for critical reading of our manuscript and valuable suggestions. The authors also thank the CCD student discussants and moderators for their contributions. We wish to sincerely address our gratitude to the CCD team for organisational support with the study: Nora Koenemann, Simone Reichert, Sandra Petrenz, Fabian Haak, Bjoern Stolte, Simon Berhe, Bastian Brandt, and Thomas Lautz. Marc Weidenbusch wishes to express special thanks to Bernd Gansbacher for introduction to CCDs.

Funding statement

This work was supported by the German Federal Ministry of Education and Research (grant no. 01PL12016) and an intramural grant of the Medical Faculty of the University of Munich (Lehre@LMU).

Competing Interests

Marc Weidenbusch declares to have no conflict of interest.

Benedikt Lenzer declares to have no conflict of interest.

Maximilian Sailer declares to have no conflict of interest.

Christian Strobel declares to have no conflict of interest.

Raphael Kunisch declares to have no conflict of interest.

Jan Kiesewetter declares to have no conflict of interest.

Martin R. Fischer declares to have no conflict of interest.

Jan M. Zottmann declares to have no conflict of interest.

Author contributions

MW, BL, MF and JZ planned the study.

MW, BL and CS were responsible for data acquisition.

MW, BL, RK, JK, JZ and MS analysed and interpreted the data.

MW, BL and JZ drafted the manuscript, all authors contributed significant intellectual content and all authors gave final approval of the version to be published.

Data sharing statement

Dataset and detailed information about the CCD formats is available upon request.

References

1. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. Medical teacher. 2010;32(8):638-45.

2. V. MMFdBDe. Nationaler Kompetenzbasierter Lernzielkatalog Medizin (NKLM). Germany: MFT Medizinischer Fakultätentag der Bundesrepublik Deutschland e. V.; 2015.

3. Harasym PH, Tsai TC, Hemmati P. Current trends in developing medical students' critical thinking abilities. The Kaohsiung journal of medical sciences. 2008;24(7):341-55.

4. Donaldson MS, Corrigan JM, Kohn LT. To err is human: building a safer health system. Washington: National Academies Press; 2000.

5. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing: let's preach what we practice. New England Journal of Medicine. 1983;309(15):921-3.

6. Hege I, Ropp V, Adler M, Radon K, Mäsch G, Lyon h, et al. Experiences with different integration strategies of case-based e-learning. Medical teacher. 2007;29(8):791-7.

7. Cannon WB. The Case Method of Teaching Systematic Medicine. The Boston Medical and Surgical Journal. 1900;142(2):31-6.

8. Eva KW. What every teacher needs to know about clinical reasoning. Medical education. 2005;39(1):98-106.

9. Harris NL. Case Records of the Massachusetts General Hospital — Continuing to Learn from the Patient. New England Journal of Medicine. 2003;348(22):2252-4.

10. Cabot RC. Case teaching in medicine: A series of graduated exercises in the differential diagnosis, prognosis and treatment of actual cases of disease: Heath; 1906.

11. Sturdy S. Knowing Cases:Biomedicine in Edinburgh, 1887—1920. Social studies of science. 2007;37(5):659-89.

12. Relman AS. Two Views. New England Journal of Medicine. 1979;301(20):1112-3.

13. Ertmer PA, Newby TJ. Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective. Performance Improvement Quarterly. 1993;6(4):50-72.

14. Merseth KK. Cases and case methods in teacher education. Handbook of research on teacher education. 1996;2:722-44.

15. Powers BW, Navathe AS, Jain SH. Medical education's authenticity problem. BMJ. 2014;348.

16. Chi MTH, Wylie R. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. Educational Psychologist. 2014;49(4):219-43.

17. Zottmann JM, Stegmann K, Strijbos J-W, Vogel F, Wecker C, Fischer F. Computer-supported collaborative learning with digital video cases in teacher education: The impact of teaching experience on knowledge convergence. Computers in Human Behavior. 2013;29(5):2100-8.

18. Hall D, Buzwell S. The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution. Active Learning in Higher Education. 2013;14(1):37-49.

19. Duncan RG, Rivet AE. Science Learning Progressions. Science. 2013;339(6118):396-7.

20. De Menezes S, Premnath D. Near-peer education: a novel teaching program. International journal of medical education. 2016;7:160.

21. Ince-Cushman D, Rudkin T, Rosenberg E. Supervised near-peer clinical teaching in the ambulatory clinic: an exploratory study of family medicine residents' perspectives. Perspectives on medical education. 2015;4(1):8-13.

22. Chi MT. Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science. 2009;1(1):73-105.

23. Ruiz JG, Mintzer MJ, Leipzig RM. The Impact of E-Learning in Medical Education. Academic Medicine. 2006;81(3):207-12.

24. Kotton DN, Muse VV, Nishino M. Case 2-2012. New England Journal of Medicine. 2012;366(3):259-69.

25. Marks PW, Zukerberg LR. Case 30-2004. New England Journal of Medicine. 2004;351(13):1333-41.

26. Uyeki TM, Sharma A, Branda JA. Case 40-2009. New England Journal of Medicine. 2009;361(26):2558-69.

27. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, et al. Representation scaffolds improve diagnostic efficiency in medical students. Medical education. 2017; 51(11):1118-1126.

28. Hautz WE, Kämmer JE, Schauber SK, Spies CD, Gaissmaier W. DIagnostic performance by medical students working individually or in teams. Jama. 2015;313(3):303-4.

29. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Medical education. 2015;49(10):961-73.

30. Lou Y, Abrami PC, d'Apollonia S. Small Group and Individual Learning with Technology: A Meta-Analysis. Review of Educational Research. 2001;71(3):449-521.

31. Springer L, Stanne ME, Donovan SS. Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. Review of Educational Research. 1999;69(1):21-51.

Figure legends

Figure 1: Live-CCD Structure. CCD sessions are divided into three parts: In the *admission* part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive *discussion* part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

Figure 2: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

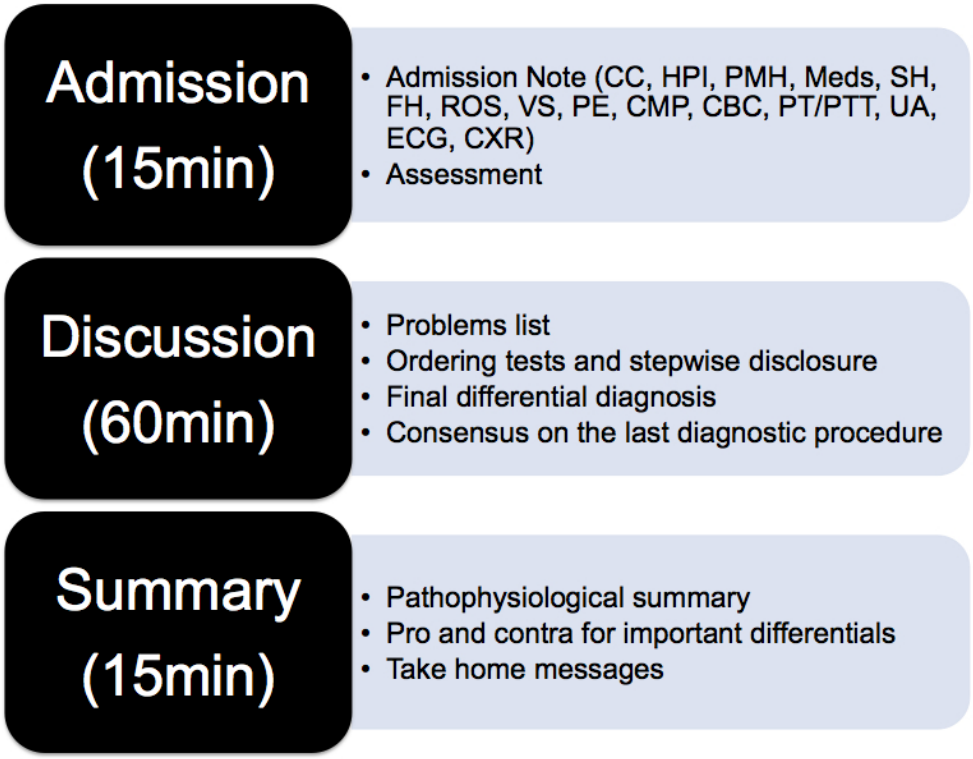


Figure 1: Live-CCD Structure. CCD sessions are divided into three parts: In the admission part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive discussion part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

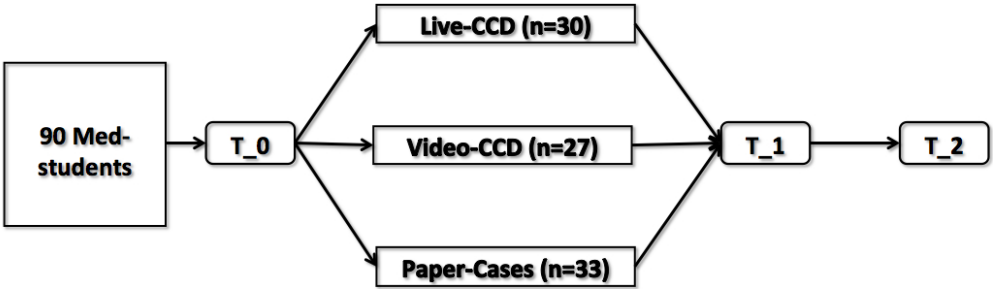


Figure 2: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-6
	2b	Specific objectives or hypotheses	6
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	7-9
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	not applicable
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	7-9
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	9-10
	6b	Any changes to trial outcomes after the trial commenced, with reasons	not applicable
Sample size	7a	How sample size was determined	10
	7b	When applicable, explanation of any interim analyses and stopping guidelines	not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	7
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	7
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	not applicable
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	not applicable
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	not applicable

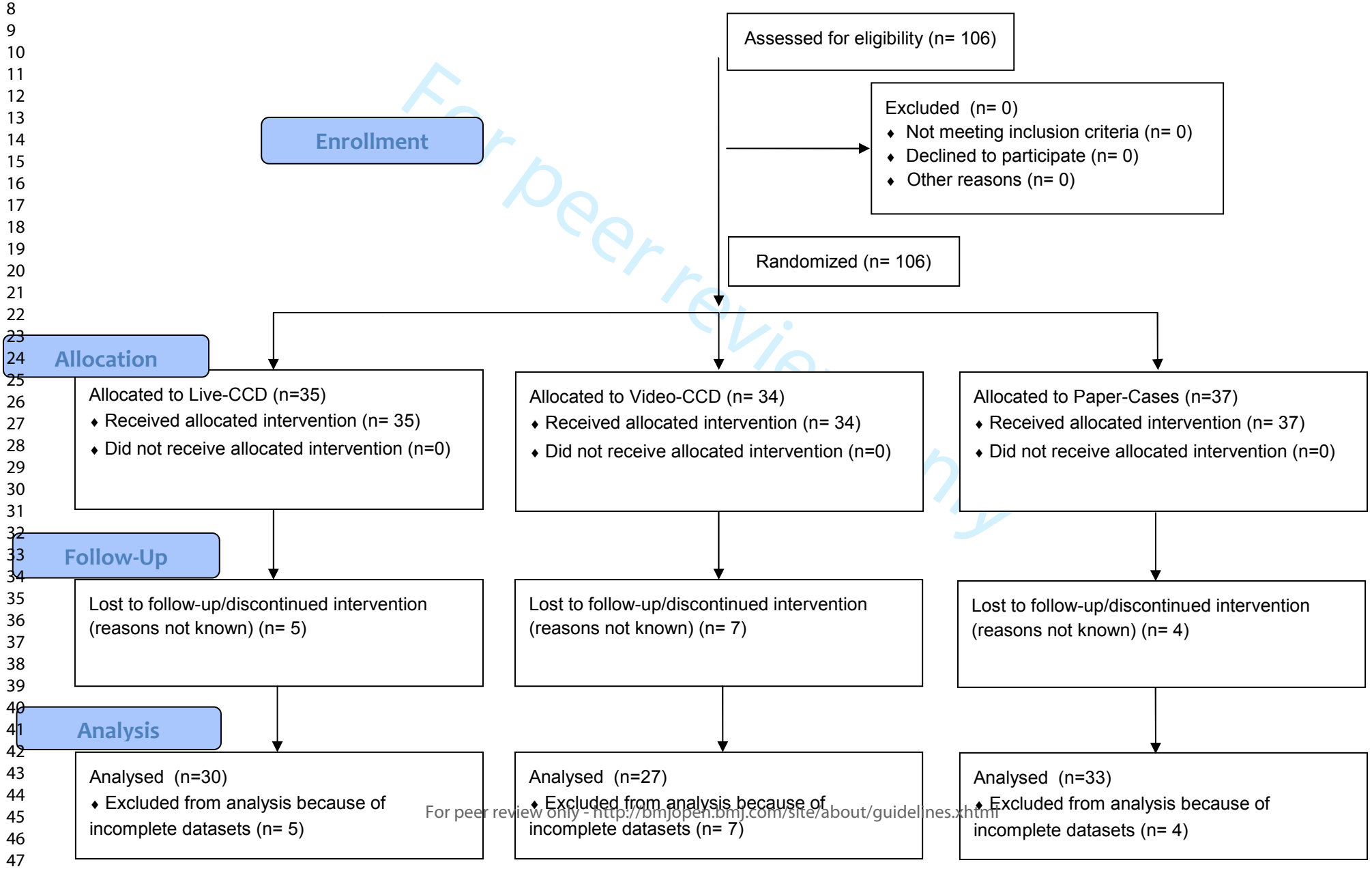
		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	10
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	10
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	attached as supplemental
	13b	For each group, losses and exclusions after randomisation, together with reasons	7
Recruitment	14a	Dates defining the periods of recruitment and follow-up	not applicable
	14b	Why the trial ended or was stopped	not applicable
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	attached as supplemental
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	11-12
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	10-13
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	not applicable
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	not applicable
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	not applicable
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	15
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	14-15
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	13-15
Other information			
Registration	23	Registration number and name of trial registry	not applicable
Protocol	24	Where the full trial protocol can be accessed, if available	17
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	16

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

CONSORT

TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



BMJ Open

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025973.R1
Article Type:	Research
Date Submitted by the Author:	19-Mar-2019
Complete List of Authors:	Weidenbusch, Marc; University Hospital of LMU Munich, Institute for Medical Education; University Hospital of LMU Munich, Department of Internal Medicine IV Lenzer, Benedikt; University Hospital of LMU Munich, Institute for Medical Education Sailer, Maximilian; LMU Munich, Department of Psychology Strobel, Christian; University Hospital of LMU Munich, Institute for Medical Education Kunisch, Raphael; University Hospital of LMU Munich, Institute for Medical Education Kiesewetter, Jan; University Hospital of LMU Munich, Institute for Medical Education Fischer, Martin; University Hospital of LMU Munich, Institute for Medical Education Zottmann, Jan; University Hospital of LMU Munich, Institute for Medical Education
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	Undergraduate medical education, Case-based learning, Clinical reasoning, Social interaction, Medical decision making

RESEARCH PAPER

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Dr. med. Marc Weidenbusch[1,2], Benedikt Lenzer[2*], Prof. Dr. phil. Maximilian Sailer[3], Dr. phil. Christian Strobel[2], Raphael Kunisch[2], Dr. phil. Jan Kiesewetter[2], Prof. Dr. med. Martin R. Fischer[2] & Dr. phil. Jan M. Zottmann[2]*

[1]Nephrologisches Zentrum, Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, LMU München

[2]Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München, LMU München

[3]Department Psychologie, LMU München

[*]equal contribution

Corresponding author:

Dr. phil. Jan M. Zottmann

Institut für Didaktik und Ausbildungsforschung in der Medizin

Klinikum der Universität München, LMU München

Pettenkoferstr. 8a, D-80336 Munich, Germany

Tel: +49 89 4400 57203

Fax: +49 89 4400 57202

E-Mail: jan.zottmann@med.uni-muenchen.de

Word count: 3600 words

Keywords: undergraduate medical education, case-based learning, clinical reasoning, social interaction, medical decision making

Abstract

Objective: Fostering clinical reasoning is a mainstay of medical education. Based on the Clinicopathological Conferences, we propose a case-based peer teaching approach called Clinical Case Discussions (CCDs) to promote the respective skills in medical students. This study compares the effectiveness of different CCD formats with varying degrees of social interaction in fostering clinical reasoning.

Design, Setting, Participants: A single-center randomised controlled trial with a parallel design was conducted at a German university. The 106 study participants were stratified (age, gender, year of study, prior CCD participation, performance in a pre-test) and tested regarding their clinical reasoning skills right after CCD participation and two weeks later.

Intervention: Participants worked either within a live discussion group (Live-CCD), a group watching recordings of the live discussions (Video-CCD), or a group working with printed cases (Paper-Cases). The presentation of case information followed an admission-, discussion-, summary-sequence.

Primary and secondary outcome measures: Clinical reasoning skills were measured with a knowledge application test addressing the students' conceptual, strategic, and conditional knowledge. Additionally, subjective learning outcomes were assessed.

Results: With respect to learning outcomes, the Live-CCD group displayed the best results, followed by Video-CCD and Paper-Cases. No difference was found between Live-CCD and Video-CCD groups in the delayed post-test; both outperformed the Paper-Cases group. Regarding subjective learning outcomes, the Live-CCD received significantly better ratings than the other formats.

Conclusions: This study demonstrates that the CCD approach is an effective and sustainable clinical reasoning teaching resource for medical students. Subjective learning outcomes underline the importance of learner (inter-)activity in the acquisition of clinical reasoning skills in the context of case-based learning. Higher efficacy of more interactive formats can be attributed to positive effects of collaborative learning. Future research should investigate how the Live-CCD format can further be improved and how video-based CCDs can be enhanced through instructional support.

Article summary

Strengths and limitations of this study:

- First empirical study on the implementation of Clinical Case Discussions in undergraduate medical education.
- Comparison of Clinical Case Discussions with differing grades of social interaction to determine their effectiveness on medical students' acquisition of clinical reasoning skills by between-group analyses.
- Implementation of multidimensional and multilayered test instruments in a pre-, post- and delayed post-test design to measure clinical reasoning skills by a knowledge application test and self-assessment.
- The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge).

Introduction

Curriculum developers face the challenge of implementing competence-oriented frameworks such as CanMEDS (Canada), the NKLM (Germany) or PROFILES (Switzerland), including the need to train clinical reasoning skills as a medical doctor's key competence.[1-3] As such, clinical reasoning skills are crucial not only for appropriate medical decision making, but also to avoid diagnostic errors and the associated harm for both patients and healthcare systems.[4]

Case-based learning has been proposed to foster clinical reasoning skills[5] and is well accepted amongst students.[6] Case-based learning found an early representation in Clinicopathological Conferences (CPC, first introduced by Cannon in 1900[7]) which are practiced until today. The Clinicopathological Conferences conducted at the Massachusetts General Hospital are published on a regular basis known as the *Case Records* series of the New England Journal of Medicine. In those CPCs the "medical mystery"[8] presented by the case under discussion calls readers to think about the possible diagnosis themselves, before it is finally disclosed at the last part of the CPC. Despite the absence of definitive evidence for efficacy as a teaching method, CPCs have widely been used in medical education since the early 20th century to foster clinical reasoning.[9-11] While these CPC-Case Records reaches lots of medical readers around the world, it has been criticised as being anachronistic with a diagnosing "star (i.e. the discussant), performing, acutely aware of being the center of attention".[12]

Case-based learning formats are embedded in a context, which is known to promote learning better than providing facts in an abstract, non-contextual form.[13] A definition found in the review by Merseth suggests three essential elements of a case: A case is real (i.e. based on a real-life situation or event); it relies on careful research and study; it

is “created explicitly for discussion and seeks to include sufficient detail and information to elicit active analysis and interpretation by users”.[14] Cases may be represented by means of text, pictures, videos, and the like. Realism and authenticity are varying features of cases,[15] but particularly elaborated and authentic cases provide increased diagnostic challenge, comprising an additional value for medical training.[16] However, due to their setup, CPCs are often a passive learning situation for participants, as they listen to the discussant laying out his or her clinical reasoning on the case under discussion. According to the ICAP framework by Chi et al.,[17] teaching formats increase their efficacy from passive < active < constructive < interactive learning environments. Based on the ICAP model, any intervention that would lead to more effective cognitive (i.e. constructive or interactive) learner activities should improve the learning outcomes of that format. Especially when students interactively engage in discussions among each other, learning is enhanced. Accordingly, case-based learning has been found to be particularly beneficial in collaborative settings.[15] However, another important aspect to consider in collaborative learning environments is that some students may participate passively while others contribute disproportionately much. To foster optimal learning effects, students should thus be encouraged to be interactively engaged. One prerequisite to achieve self-guided learning in groups is a low threshold for students to come forward with their questions and participate in ensuing discussions.[18] To this end, peer teaching has been established as an effective tool to stimulate discussions.[19] To make sure peer tutors are not overwhelmed in moderating these discussions, the presence of an experienced clinician appears to be warranted[20] in addition to a specific training of the tutors.

Taken together, while traditional CPCs encompass some important dimensions of effective case-based learning environments, they are not systematically aiming at constructive or interactive learner activities that are known features of effective teaching formats.[17,21] Therefore, we introduced Clinical Case Discussions (CCD) in undergraduate medical education to account for these features. We still use the Case Records of the Massachusetts General Hospital,[9] as these cases exemplify realistic patient encounters and fulfill the criteria for an interactive collaborative learning process as explained above. In the CCD approach, cases are presented only until the hospital admission of the patient, followed by an interactive discussion about possible diagnoses and diagnostic strategies. After all test results have been discussed, the actual diagnosis is disclosed and the pitfalls and take-home messages of the case are summarised.

To investigate the effectiveness of the CCD approach in undergraduate medical education, we designed an intervention trial and assessed clinical reasoning skills in medical students before and after *participating in* live CCDs or *being exposed* to video recordings of live CCDs. We compared these formats and its effects on clinical reasoning with the more traditional approach of working through written cases. When carrying out this randomised trial, we hypothesised that participation in live CCD sessions would lead to a higher increase of clinical reasoning skills than simply reading the cases. To better understand possible effects of the CCD learning environment with its social components on learning outcomes, participation in live CCDs as outlined above was additionally compared to the effects of watching videos of CCDs online. This comparison also seemed relevant from an economic point of view as video-streaming of lectures and seminars are prevalent at many instituitions in higher education allowing for flexible and scalable access to learning materials.[22] To

investigate the potential of different CCD formats for regular curricular use, we also measured subjective learning outcomes after the intervention and correlated student self-assessments with objective changes in their clinical reasoning skills.

Methods

Participants / Ethics

Initially, we recruited 106 volunteer medical students at the Medical Faculty of LMU Munich. Randomisation was performed in a two-step procedure: First, we selected a sample of roughly 100 enrolled students. Next, we stratified participants by creating triplets on the basis of the variables age, gender, year of study, prior CCD participation and performance in a knowledge application pre-test. This was done in an effort to limit the risk of random misdistribution of the selected sample. From each triplet we randomly assigned participants to the experimental groups. A total of 90 participants eventually completed the study, 31 of them were male and 59 female. They were 20 to 41 years old (M = 23; SD = 2.97) and in their first to eighth clinical semester (M = 3.5; SD = 1.78).

The study was approved by the ethics committee of the Medical Faculty of LMU Munich (approval reference no. 222-15). Written informed consent was obtained from all study participants and they received a financial reimbursement of 50 Euros upon completion of the trial.

Patient and public involvement

No patients or public were involved in this research.

Study design

We conducted a single-center randomised controlled trial consisting of a total of five course sessions with a parallel design (see Figure 1). One week prior to the first CCD session, participants were introduced to the principles of the CCD approach and the sequence of this trial in an introductory session where they also took a knowledge application pre-test (T_0). In the experimental phase, participants attended three weekly interventional course sessions of 90 minutes each in one of the three aforementioned groups with the respective CCD formats. Participants took a knowledge application post-test at the end of the last experimental course session (T_1), four weeks after pre-testing. A delayed knowledge application post-test was conducted two weeks after completion of the interventional courses (T_2); we deliberately chose that time interval to investigate the sustainability of possibly effects while balancing the risk of post-intervention confounding.[23]

****

Insert Figure 1 about here

****

Materials

In all experimental groups the intervention was based on the same three, independent internal medicine cases. Chief complaints in these cases were paresthesia (first session), fever and respiratory failure (second session), and rapidly progressive respiratory failure (third session).[24-26] Cases were worked through in an iterative approach in different formats: (a) peer-moderated live case discussions in an interactive setting (Live-CCD, n = 30), (b) a single-learner format utilizing an interactive multimedia platform displaying video recordings of the live case discussions (Video-CCD, n = 27), and (c) a single-learner format in which the students worked with the original paper cases of the NEJM

(Paper-Cases, $n = 33$). The cases were prepared in a way that participants in each format were exposed to the same case information.

Procedure

In all three groups cases were presented in a specified structured manner similar to the original Clinicopathological Conferences (see Figure 2). In each format the students ("discussants") had to fill out a form after the admission in which the case had to be summarised and a list of clinical problems and working diagnoses had to be provided. Subsequently, between discussion and summary a second case-summary had to be completed in which the final diagnostic test and the most likely diagnosis had to be proposed.

****

Insert Figure 2 about here

****

In the Live-CCD group, the case presentation was prepared beforehand by a voluntary discussant ("presenter"), who presented the facts in the admission (according to the structure shown in Figure 2). Electronic slides and flipcharts were used to transport case information. Original test results were revealed by the presenter during the discussion only when requested by the group of students. Furthermore, the presenter summarised the differential diagnosis, important pathophysiological features of the case at the end of the session and provided a short take home message. The moderating medical students ("moderator") were recruited among previous CCD participants. They had experience in CCD moderation and had had an introductory training (two days) in higher education methods and group facilitation prior to the study. The moderator facilitated the discussion process and ensured a reasonable approach to the patient encounter (e.g. with

respect to timing and hierarchy of ordered tests) in close communication with the discussants. Moreover, the moderator helped students develop their diagnostic strategy by co-evaluating their requested findings and the reasoning employed. Supervision of the correctness of medical facts and the correct diagnostic approach were ultimately granted by a clinician who could stop the discussion at any point when faulty reasoning was evident or discussants explicitly requested the facilitation of an experienced physician. The clinicians' level of involvement into the discussion was left at their own discretion. We varied the staff between each Live-CCD to minimise effects of personal teacher characteristics. Live sessions typically lasted 90 minutes and were recorded with multiple cameras.

Students in the Video-CCD format worked on a single-learner multimedia workstation on which a video recording of the Live-CCD were displayed. These recordings also contained the electronic slide presentation from the Live-CCD and enabled simultaneous observation of the discussion from multiple camera angles. Participants could pause and partially skip the videos.

In the Paper-Cases group participants received the case information of each CCD section sequentially (i.e. admission, discussion, summary) in a print format. In both single-learner formats students could choose their personal working speed. There was neither a prespecified minimum nor a maximum time they were required to work on the cases. In each of the three formats full access to the internet was permitted for additional information.

Instruments

Learning outcomes with respect to clinical reasoning were measured with a knowledge application test that consisted of 29 items (i.e. a maximum of 29 points could be achieved) and was to be filled out within 45 minutes. The knowledge application test was based on instruments previously developed at the Institute for Medical Education at LMU Munich.[27-29] It comprised multiple choice items, key feature problems and problem-solving tasks, addressing the conceptual, strategic, and conditional knowledge of the participants (see Figure 3). Meta-analyses on retest effects suggest that score increase is higher for identical forms than for parallel test forms.[30] In order to limit such effects, we applied parallel forms of the knowledge application test for pre- and post-measurements (i.e. topics covered by the individual items were the same, but the items were reformulated and their order was permutated). Overall test difficulty was chosen to be high in order to avoid ceiling effects, as students from all clinical years were allowed to participate in the study. Overall test reliability was satisfactory (Cronbach's $\alpha = .71$).

****

Insert Figure 3 about here

****

Subjective learning outcomes were measured at T_1 with a short questionnaire consisting of 9 items (e.g. "I learned a lot during the CCD course", "The CCD course increased my learning motivation" or "I recommend the implementation of the CCD teaching format into the curriculum"; the full questionnaire is available as a supplementary file). Participants were asked to rate these items on a Likert scale ranging from 1 (I don't agree) to 5 (I fully agree). Reliability of the corresponding scale was good (Cronbach's $\alpha = .95$). Additionally, study participants were asked to share their

views on positive and negative aspects of the respective training format through open items at the end of the questionnaire.

Statistical Analysis

The required sample size (N = 128) was estimated to detect medium effect sizes with a power of 80% and a significance level of $\alpha = .05$. For between-group analyses, one-way ANOVAs were conducted with *post-hoc* Bonferroni tests for multiple comparisons.

Results

Effects of the CCD format on learning outcomes related to Clinical Reasoning

Experimental groups differed significantly with respect to the knowledge application post-test (see Table 1), $F(2,87) = 27.07$, $p = .000$, partial $\eta^2 = .384$ (this corresponds to a Cohen's d of 1.580). The Live-CCD group ($M = 14.10$; $SD = 3.32$) outperformed both the Video-CCD ($M = 11.69$; $SD = 3.34$) and the Paper-Cases group ($M = 8.5$; $SD = 2.44$). Post hoc Bonferroni tests revealed significant differences between Live-CCD and Video-CCD ($p = .011$) as well as the Paper-Cases group ($p = .000$). The difference in the knowledge application post-test between Video-CCD and the Paper-Cases group was also significant ($p = .000$).

Two weeks after course completion, the effect of the teaching format was still found in a delayed knowledge application post-test, $F(2,87) = 30.91$, $p = .000$, partial $\eta^2 = .415$ (this corresponds to a Cohen's d of 1.685). Both Live-CCD ($M = 13.36$; $SD = 3.23$) and the Video-CCD ($M = 11.84$; $SD = 2.92$) outperformed the Paper-Cases group ($M = 7.89$; $SD = 2.41$). Post hoc Bonferroni tests revealed significant differences between the Live-CCD and Paper-Cases group ($p = .000$) as well as between the Video-CCD and Paper-

Cases group (p = .000). However, the difference between Live-CCD and Video-CCD was not significant in the delayed knowledge application post-test (p = .146).

Table 1. Overview of the findings of the study.

	Teaching format					
	Live-CCD		Video-CCD		Paper-Cases	
	M	(SD)	M	(SD)	M	(SD)
Knowledge application pre-test	5.34	(1.92)	4.76	(1.90)	5.76	(2.24)
	n = 30		n = 27		n = 33	
Knowledge application post-test	14.10	(3.32)	11.69	(3.34)	8.50	(2.44)
	n = 30		n = 27		n = 33	
Delayed knowledge application post-test	13.36	(3.23)	11.84	(2.92)	7.89	(2.41)
	n = 30		n = 27		n = 33	
Subjective learning outcomes	4.20	(.63)	3.18	(1.24)	3.00	(.99)
	n = 30		n = 27		n = 31	

Effects of the CCD format on subjective learning outcomes

Experimental groups differed significantly with respect to subjective learning outcomes (see Table 1), $F(2,85) = 13.16$, p = .000, partial η^2 = .236 (this corresponds to a Cohen's d of 1.112). Participants of the Live-CCD group (M = 4.20; SD = .63) assigned better ratings to their course format than participants in the Video-CCD group (M = 3.18; SD = 1.24) and the Paper-Cases group (M = 3.0; SD = .99). Post hoc Bonferroni tests showed that the Live-CCD differed from the Video-CCD (p = .001) and the Paper-Cases group (p = .000) in this regard. An additional Duncan post-hoc test confirmed that the Video-CCD and the Paper-Cases group did not differ from each other in this regard (p = .48).

To investigate the relations between the subjective assessment and the knowledge application tests applied at the end and two weeks after the course, we calculated correlations between the different outcome measures. Subjective learning outcomes correlated on a medium level with both the knowledge application post-test (r = .343, n = 88, p = .001) and the delayed knowledge application post-test (r = .339, n = 88, p = .001).

In the Live-CCD group, 83% of the students were in favour of implementing routine Live-CCD into the medical curriculum. Only 45% and 31% of students from the Video-CCD and Paper-Cases groups voted for an implementation of their respective course in the curriculum. With respect to the open items from the subjective learning outcomes questionnaire, participants from all groups praised the quality of the cases. Participants from the Live-CCD group particularly valued their course format for providing an opportunity to practice "diagnostic thinking" and the "focus on practice elements". They also mentioned that "you can look up theoretical knowledge, but you can't look up applied knowledge". Students in the Video-CCD group, on the other hand, praised features of the digital learning environment as they could "pause, reflect, or quickly do a Google search" when watching the case discussions. However, they also criticised it was not possible for them to "participate in a more active way".

Discussion

This randomised controlled study shows that even relatively short CCD interventions can lead to improved and sustainable learning outcomes with respect to clinical reasoning. This provides evidence that the CCD approach, which is based on Clinicopathological Conferences, is an effective teaching resource to foster clinical

reasoning skills in medical students. We had hypothesised that a more interactive course format would result in an improvement of clinical reasoning skills when compared with less interactive formats. Results show that the Live-CCD indeed leads to the highest learning outcomes in medical students compared to less interactive formats. Consistent with our hypothesis, clinical reasoning skills, as measured with our knowledge application test, had the highest gain in the Live-CCD group. These positive effects of the CCD teaching format on clinical reasoning skills proved sustainable as shown by the results in the delayed knowledge application post-test. Overall, these results are in line with a recently published study on diagnostic reasoning[31] where students who worked in pairs were more accurate in their diagnosis than individual students despite having comparable knowledge. Collaborative clinical reasoning has thus far been underrepresented in the literature, yet seems to solve many of the educational problems regarding diagnostic errors.[32]

The significant difference between the Live-CCD and the Video-CCD group can be explained by the findings of a meta-analysis that showed technology-assisted single-person learning to be inferior to group learning because of the decreased social interaction.[33] However, it is important to note that two weeks after the course, participants of the Live-CCD and Video-CCD groups did not differ significantly anymore while both groups still clearly outperformed the Paper-Case group. In other words, watching a video of the live case discussion was found to be more beneficial for learners regarding their clinical reasoning skills than just reading the printed cases. We cannot rule out that Live-CCD and Video-CCD groups did not differ in the delayed knowledge application post-test due to underpowering of the study. As our trial was not designed to detect smaller effect sizes, this finding has to be treated with caution.

Subjective learning outcomes suggest that students prefer the live discussion over the other formats and were linked to their performance in the knowledge application test. Additional qualitative data from the open item answers suggests that the Live-CCD format supported students in performing clinical reasoning and that the active discussion of cases was particularly valued by the students.

Generalisability

The conclusions of this study are applicable to a broader audience of medical students. The CCD approach and its respective formats can easily be implemented in routine medical education. Peer teaching courses hold the promise of being more easy to install and more easy to staff then courses led by faculty. Of course, live CCDs still come with certain personnel requirements, as faculty as well as a moderator need to be present. Special preparation is not necessary for the clinician though, so total time requirements might still be lower compared to other teaching formats. Likewise, the implementation of a singular two-day training for moderators should not require extensive ressources. The study population consisting of students with heterogeneous levels of clinical experience implies that the CCD is an effective teaching format not only for students at the beginning of their clinical career but also for intermediate students. On the other hand, generalisability is potentially limited as only students from one medical school participated in our study.

Limitations of the study

There are certain limitations of this study that have to be addressed: One important limitation is the single-centre nature of this study and the relatively small sample size.

Before the CCD approach can be implemented on a larger scale, a validation of our findings is therefore required. Caution is clearly warranted with the effect sizes shown in this trial, as it has been shown that effect sizes of learning intervention trials tend to be inflated compared to the effectiveness of the intervention when used in routine education.[34] Since we did not limit the time students had to work on the cases, we cannot entirely rule out that less time was spent on task in the single-learner formats and particularly the Paper-Cases group. Against this backdrop, we suggest replication to further validate the results found in this study and strengthen the outlined implications. Finally, the knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge). Larger item numbers could facilitate a reliable assessment of changes on the level of corresponding subscales.

Implications for policy makers / Future research questions

Based on our findings, the CCD approach is a useful asset for medical educators to widen the range of clinical reasoning teaching tools. Live-CCD can thus be seen as a prime candidate for routine implementation in clinical reasoning curricula. Future research should aim to identify which Live-CCD elements (the roles, case contents, or the course structure) contribute in which way to the improvement of clinical reasoning skills in medical students. The question if and to what extent such skills are applicable across domains is currently being discussed.[35] Future studies may also address the issue of transfer (i.e. to what extent can clinical reasoning skills obtained in case-based training later be applied to different cases?).[36] Regarding the Video-CCD, means of

instructional support to increase the effectiveness and interactivity of the video-based format should be investigated in an attempt to exploit its full potential.

Acknowledgements

The authors thank Johanna Huber and her team for technical support with the evaluation, Thomas Brendel and Thomas Bischoff for help with the video production and Mark S. Pecker for critical reading of our manuscript and valuable suggestions. The authors also thank the CCD student discussants and moderators for their contributions. We wish to sincerely address our gratitude to the CCD team for organisational support with the study: Nora Koenemann, Simone Reichert, Sandra Petrenz, Fabian Haak, Bjoern Stolte, Simon Berhe, Bastian Brandt, and Thomas Lautz. Marc Weidenbusch wishes to express special thanks to Bernd Gansbacher for introduction to CCDs.

Funding statement

This work was supported by the German Federal Ministry of Education and Research (grant no. 01PL12016) and an intramural grant of the Medical Faculty of the University of Munich (Lehre@LMU).

Competing Interests

Marc Weidenbusch declares to have no conflict of interest.

Benedikt Lenzer declares to have no conflict of interest.

Maximilian Sailer declares to have no conflict of interest.

Christian Strobel declares to have no conflict of interest.

Raphael Kunisch declares to have no conflict of interest.

Jan Kiesewetter declares to have no conflict of interest.

Martin R. Fischer declares to have no conflict of interest.

Jan M. Zottmann declares to have no conflict of interest.

Author contributions

MW, BL, MF and JZ planned the study.

MW, BL and CS were responsible for data acquisition.

MW, BL, RK, JK, JZ and MS analysed and interpreted the data.

MW, BL and JZ drafted and revised the manuscript, all authors contributed significant intellectual content and all authors gave final approval of the version to be published.

Data sharing statement

Dataset and detailed information about the CCD formats is available upon request.

References

1. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: Theory to practice. Medical Teacher. 2010;32(8):638-45.

2. Fischer MR, Bauer D, Mohn K, et al. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. GMS Journal for Medical Education. 2015;32(3):Doc35.

3. Harasym PH, Tsai TC, Hemmati P. Current trends in developing medical students' critical thinking abilities. The Kaohsiung Journal of Medical Sciences. 2008;24(7):341-55.

4. Donaldson MS, Corrigan JM, Kohn LT. To err is human: Building a safer health system. Washington: National Academies Press; 2000.

5. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing: let's preach what we practice. New England Journal of Medicine. 1983;309(15):921-3.

6. Hege I, Ropp V, Adler M, Radon K, Mäsch G, Lyon H, et al. Experiences with different integration strategies of case-based e-learning. Medical Teacher. 2007;29(8):791-7.

7. Cannon WB. The Case Method of Teaching Systematic Medicine. The Boston Medical and Surgical Journal. 1900;142(2):31-6.

8. Eva KW. What every teacher needs to know about clinical reasoning. Medical Education. 2005;39(1):98-106.

9. Harris NL. Case records of the Massachusetts General Hospital – Continuing to learn from the patient. New England Journal of Medicine. 2003;348(22):2252-4.

10. Cabot RC. Case teaching in medicine: A series of graduated exercises in the differential diagnosis, prognosis and treatment of actual cases of disease. Boston: Heath; 1906.

11. Sturdy S. Knowing Cases: Biomedicine in Edinburgh, 1887-1920. Social Studies of Science. 2007;37(5):659-89.

12. Relman AS. Two Views. New England Journal of Medicine. 1979;301(20):1112-3.

13. Ertmer PA, Newby TJ. Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance Improvement Quarterly. 1993;6(4):50-72.

14. Merseth KK. Cases and case methods in teacher education. In: Sikula J, Buttery TJ, Guyton R, editors. Handbook of research on teacher education. 2nd ed. New York, NY: Macmillan; 1996. p. 722-744.

15. Zottmann JM, Stegmann K, Strijbos J-W, Vogel F, Wecker C, Fischer F. Computer-supported collaborative learning with digital video cases in teacher education: The impact of teaching experience on knowledge convergence. Computers in Human Behavior. 2013;29(5):2100-8.

16. Powers BW, Navathe AS, Jain SH. Medical education's authenticity problem. BMJ. 2014;348:g2651.

17. Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist. 2014;49(4):219-43.

18. Duncan RG, Rivet AE. Science learning progressions. Science. 2013;339(6118):396-7.

19. De Menezes S, Premnath D. Near-peer education: A novel teaching program. International journal of medical education. 2016;7:160-167.

20. Ince-Cushman D, Rudkin T, Rosenberg E. Supervised near-peer clinical teaching in the ambulatory clinic: an exploratory study of family medicine residents' perspectives. Perspectives on medical education. 2015;4(1):8-13.

21. Chi MTH. Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science. 2009;1(1):73-105.

22. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. Academic Medicine. 2006;81(3):207-12.

23. Miller DC, Salkind NJ, editors. Handbook of Research Design & Social Measurement. Thousand Oaks, CA: SAGE; 2002.

24. Kotton DN, Muse VV, Nishino M. Case 2-2012. New England Journal of Medicine. 2012;366(3):259-69.

25. Marks PW, Zukerberg LR. Case 30-2004. New England Journal of Medicine. 2004;351(13):1333-41.

26. Uyeki TM, Sharma A, Branda JA. Case 40-2009. New England Journal of Medicine. 2009;361(26):2558-69.

27. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, et al. Representation scaffolds improve diagnostic efficiency in medical students. Medical Education. 2017; 51(11):1118-1126.

28. Kopp V, Stark R, Kühne-Eversmann L, Fischer MR. Do worked examples foster medical students' diagnostic knowledge of hyperthyroidism? Medical Education. 2009;43(12):1210-7

29. Schmidmaier R, Eiber S, Ebersbach R, Schiller M, Hege I, Holzer M, et al. Learning the facts in medical school is not enough: Which factors predict successful application of procedural knowledge in a laboratory setting? BMC Medical Education. 2013;13:28.

30. Hausknecht JP, Halbert JA, Di Paolo NT, Moriarty Gerrard MO. Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. Journal of Applied Psychology. 2007;92(2):373–385.

31. Hautz WE, Kämmer JE, Schauber SK, Spies CD, Gaissmaier W. Diagnostic performance by medical students working individually or in teams. JAMA. 2015;313(3):303-4.

32. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Medical Education. 2015;49(10):961-73.

33. Lou Y, Abrami PC, d'Apollonia S. Small group and individual learning with technology: A meta-analysis. Review of Educational Research. 2001;71(3):449-521.

34. Springer L, Stanne ME, Donovan SS. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of Educational Research. 1999;69(1):21-51.

35. Fischer F, Chinn C, Engelmann K, Osborne J, editors. Scientific reasoning and argumentation. The roles of general and specific knowledge. New York, NY: Routledge; 2018.

36. Keemink Y, Custers E, van Dijk S, ten Cate O. Illness script development in pre-clinical education through case-based clinical reasoning training International Journal of Medical Education. 2018;9:35-41.

Figure legends

Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the *admission* part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive *discussion* part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

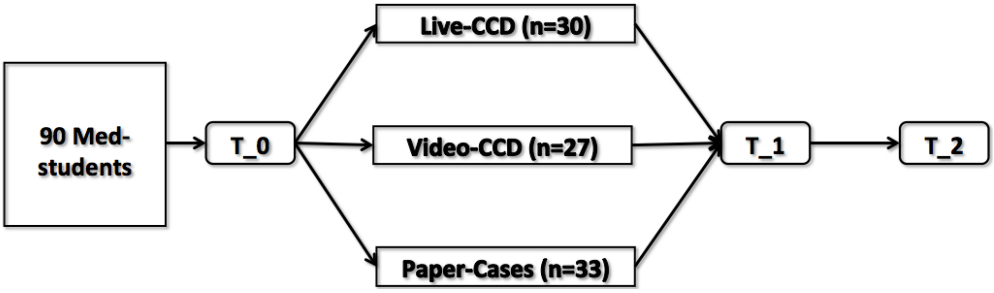


Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

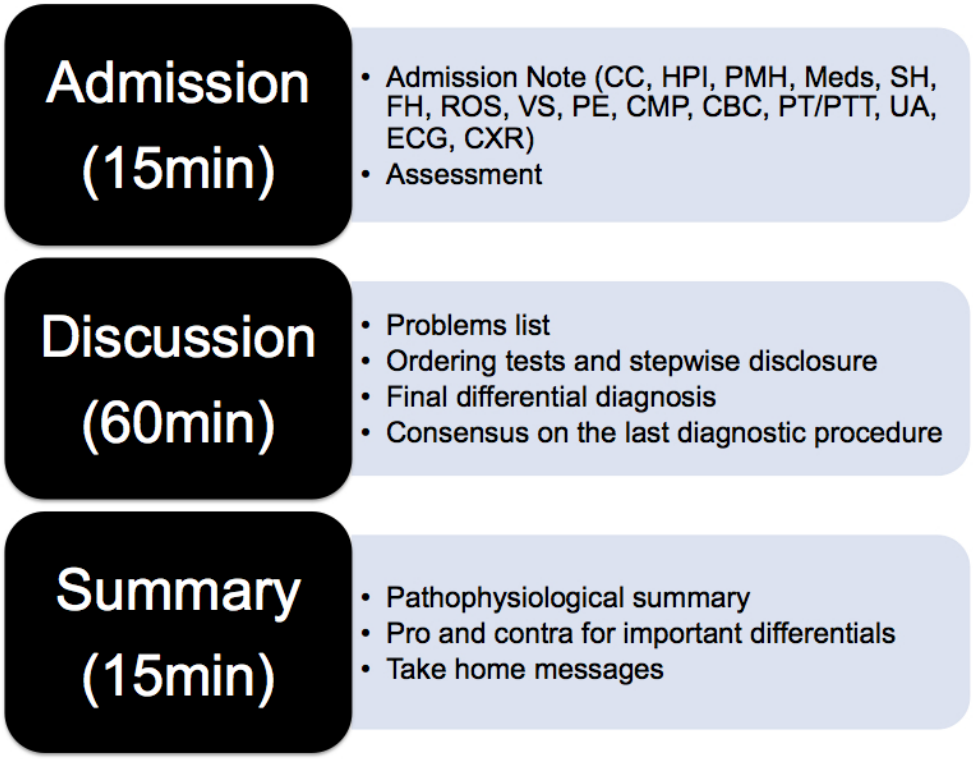


Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the admission part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive discussion part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

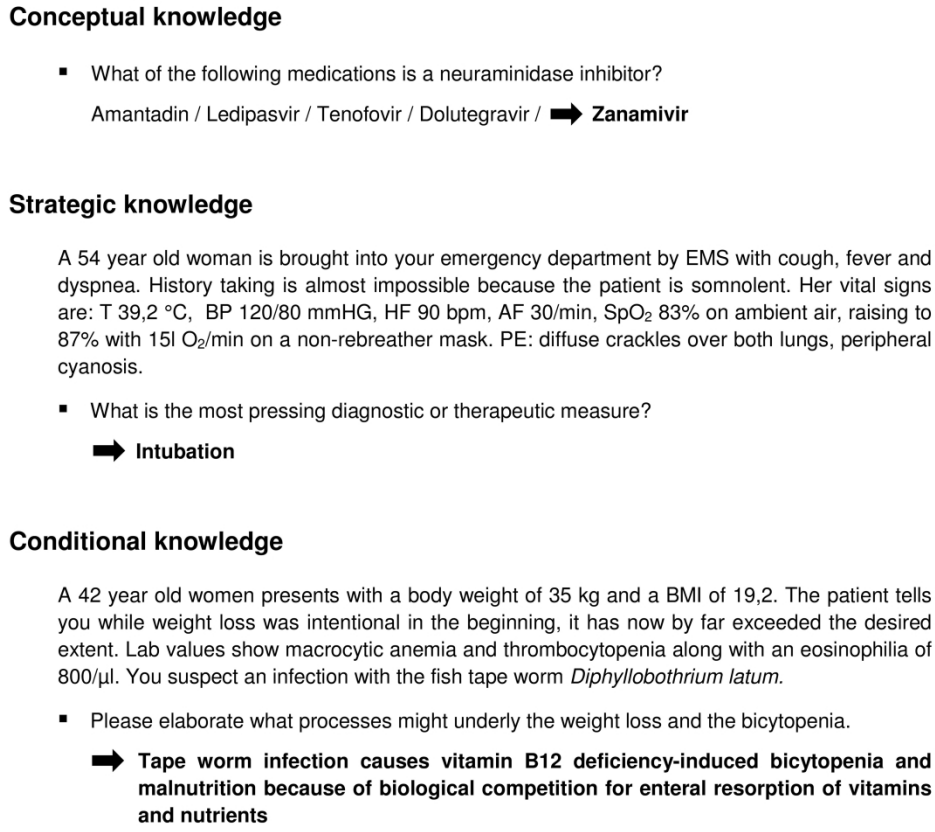


Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

154x138mm (600 x 600 DPI)

Supplemental

Baseline data: Demographic characteristics of the study participants.

	CCD format							
	Live-CCD *n* = 30		Video-CCD *n* = 27		Paper-Cases *n* = 33		All formats *N* = 90	
Gender distribution *N* m/f (% f)	14/16 (53.3% f)		5/22 (81.5% f)		12/21 (63.6% f)		31/59 (65.6% f)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Age	23.77	(4.09)	22.26	(1.77)	22.91	(2.40)	23.0	(2.97)
Clinical semester	3.23	(1.96)	3,41	(1.47)	3.82	(1.84)	3.50	(1.78)
High school grade	1.53	(0.36)	1,35	(0.42)	1.48	(0.68)	1.46	(0.52)
First National Board Exam Score	245	(30)	226	(78)	246	(30)	240	(49)
Participants with prior CCD experience n (%)	6	(20.0%)	6	(22.2%)	5	(15.1%)	17	(18.9%)

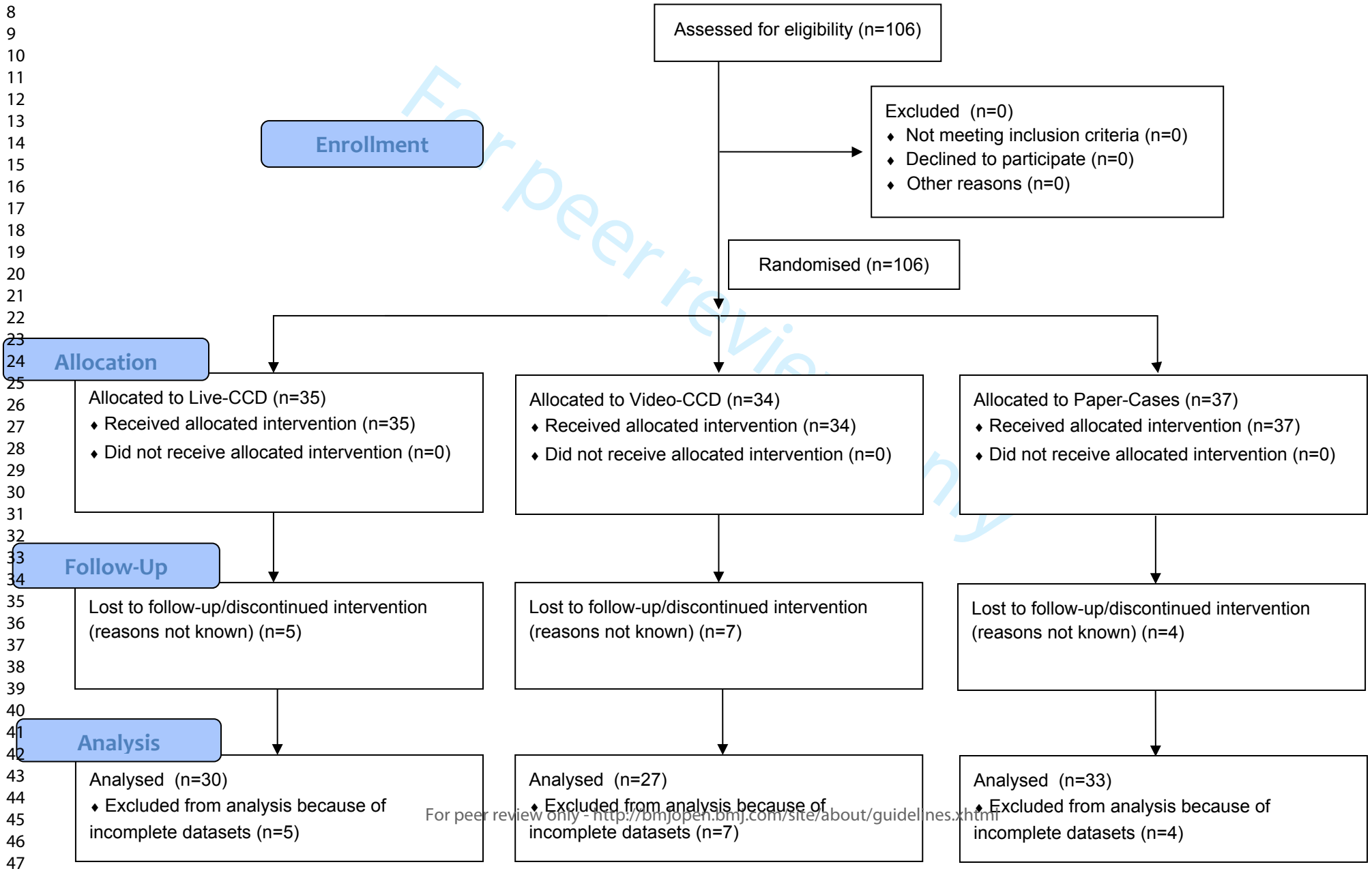
Questionnaire items (5-point Likert scale) for the assessment of subjective learning outcomes.

1. I perceived this CCD format as meaningful.
2. I learned a lot during the CCD course.
3. The CCD course increased my learning motivation.
4. I would like to participate in this CCD format again in the future.
5. I enjoyed the CCD course.
6. This CCD format should be offered as part of the curriculum.
7. I was able to follow the case discussions.
8. Learning in the CCD format is easier for me than learning in traditional lectures or seminars.
9. How would you rate the course overall?

CONSORT

TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-7
	2b	Specific objectives or hypotheses	6-7
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	8
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	not applicable
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	8-10
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8, 11-12
	6b	Any changes to trial outcomes after the trial commenced, with reasons	not applicable
Sample size	7a	How sample size was determined	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines	not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	7
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	7
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	not applicable
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	not applicable
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	not applicable

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	attached as supplemental
	13b	For each group, losses and exclusions after randomisation, together with reasons	7
Recruitment	14a	Dates defining the periods of recruitment and follow-up	not applicable
	14b	Why the trial ended or was stopped	not applicable
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	attached as supplemental
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	13
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	12-14
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	not applicable
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	not applicable
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	not applicable
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	16-17
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	16
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	14-16
Other information			
Registration	23	Registration number and name of trial registry	not applicable
Protocol	24	Where the full trial protocol can be accessed, if available	19
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	18

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

BMJ Open

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025973.R2
Article Type:	Research
Date Submitted by the Author:	20-Jun-2019
Complete List of Authors:	Weidenbusch, Marc; University Hospital of LMU Munich, Institute for Medical Education; University Hospital of LMU Munich, Department of Internal Medicine IV Lenzer, Benedikt; University Hospital of LMU Munich, Institute for Medical Education Sailer, Maximilian; LMU Munich, Department of Psychology Strobel, Christian; University Hospital of LMU Munich, Institute for Medical Education Kunisch, Raphael; University Hospital of LMU Munich, Institute for Medical Education Kiesewetter, Jan; University Hospital of LMU Munich, Institute for Medical Education Fischer, Martin; University Hospital of LMU Munich, Institute for Medical Education Zottmann, Jan; University Hospital of LMU Munich, Institute for Medical Education
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	Undergraduate medical education, Case-based learning, Clinical reasoning, Social interaction, Medical decision making

RESEARCH PAPER

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Dr. med. Marc Weidenbusch[1,2], Benedikt Lenzer[2*], Prof. Dr. phil. Maximilian Sailer[3], Dr. phil. Christian Strobel[2], Raphael Kunisch[2], Dr. phil. Jan Kiesewetter[2], Prof. Dr. med. Martin R. Fischer[2] & Dr. phil. Jan M. Zottmann[2]*

[1]Nephrologisches Zentrum, Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, LMU München

[2]Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München, LMU München

[3]Department Psychologie, LMU München

[*]equal contribution

Corresponding author:

Dr. phil. Jan M. Zottmann

Institut für Didaktik und Ausbildungsforschung in der Medizin

Klinikum der Universität München, LMU München

Pettenkoferstr. 8a, D-80336 Munich, Germany

Tel: +49 89 4400 57203

Fax: +49 89 4400 57202

E-Mail: jan.zottmann@med.uni-muenchen.de

Word count: 3645 words

Keywords: undergraduate medical education, case-based learning, clinical reasoning, social interaction, medical decision making

Abstract

Objective: Fostering clinical reasoning is a mainstay of medical education. Based on the Clinicopathological Conferences, we propose a case-based peer teaching approach called Clinical Case Discussions (CCDs) to promote the respective skills in medical students. This study compares the effectiveness of different CCD formats with varying degrees of social interaction in fostering clinical reasoning.

Design, Setting, Participants: A single-center randomised controlled trial with a parallel design was conducted at a German university. The 106 study participants were stratified (age, gender, year of study, prior CCD participation, performance in a pre-test) and tested regarding their clinical reasoning skills right after CCD participation and two weeks later.

Intervention: Participants worked either within a live discussion group (Live-CCD), a group watching recordings of the live discussions (Video-CCD), or a group working with printed cases (Paper-Cases). The presentation of case information followed an admission-, discussion-, summary-sequence.

Primary and secondary outcome measures: Clinical reasoning skills were measured with a knowledge application test addressing the students' conceptual, strategic, and conditional knowledge. Additionally, subjective learning outcomes were assessed.

Results: With respect to learning outcomes, the Live-CCD group displayed the best results, followed by Video-CCD and Paper-Cases. No difference was found between Live-CCD and Video-CCD groups in the delayed post-test; both outperformed the Paper-Cases group. Regarding subjective learning outcomes, the Live-CCD received significantly better ratings than the other formats.

Conclusions: This study demonstrates that the CCD approach is an effective and sustainable clinical reasoning teaching resource for medical students. Subjective learning outcomes underline the importance of learner (inter-)activity in the acquisition of clinical reasoning skills in the context of case-based learning. Higher efficacy of more interactive formats can be attributed to positive effects of collaborative learning. Future research should investigate how the Live-CCD format can further be improved and how video-based CCDs can be enhanced through instructional support.

Article summary

Strengths and limitations of this study:

- First empirical study on the implementation of Clinical Case Discussions in undergraduate medical education.
- Comparison of Clinical Case Discussions with differing grades of social interaction to determine their effectiveness on medical students' acquisition of clinical reasoning skills by between-group analyses.
- Implementation of multidimensional and multilayered test instruments in a pre-, post- and delayed post-test design to measure clinical reasoning skills by a knowledge application test and self-assessment.
- The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge).

Introduction

Curriculum developers face the challenge of implementing competence-oriented frameworks such as CanMEDS (Canada), the NKLM (Germany) or PROFILES (Switzerland), including the need to train clinical reasoning skills as a medical doctor's key competence.[1-3] As such, clinical reasoning skills are crucial not only for appropriate medical decision making, but also to avoid diagnostic errors and the associated harm for both patients and healthcare systems.[4]

Case-based learning has been proposed to foster clinical reasoning skills[5] and is well accepted among students.[6] Case-based learning found an early representation in Clinicopathological Conferences (CPC, first introduced by Cannon in 1900[7]) which are practiced until today. The Clinicopathological Conferences conducted at the Massachusetts General Hospital are published on a regular basis known as the *Case Records* series of the New England Journal of Medicine. In those CPCs the "medical mystery"[8] presented by the case under discussion calls readers to think about the possible diagnosis themselves, before it is finally disclosed at the last part of the CPC. Despite the absence of definitive evidence for efficacy as a teaching method, CPCs have widely been used in medical education since the early 20th century to foster clinical reasoning.[9-11] While these CPC-Case Records reaches lots of medical readers around the world, it has been criticised as being anachronistic with a diagnosing "star (i.e. the discussant), performing, acutely aware of being the center of attention".[12]

Case-based learning formats are embedded in a context, which is known to promote learning better than providing facts in an abstract, non-contextual form.[13] A definition found in the review by Merseth suggests three essential elements of a case: A case is real (i.e. based on a real-life situation or event); it relies on careful research and study; it

is “created explicitly for discussion and seeks to include sufficient detail and information to elicit active analysis and interpretation by users”.[14] Cases may be represented by means of text, pictures, videos, and the like. Realism and authenticity are varying features of cases,[15] but particularly elaborated and authentic cases provide increased diagnostic challenge, comprising an additional value for medical training.[16] However, due to their setup, CPCs are often a passive learning situation for participants, as they listen to the discussant laying out his or her clinical reasoning on the case under discussion. According to the ICAP framework by Chi et al.,[17] teaching formats increase their efficacy from passive < active < constructive < interactive learning environments. Learning is enhanced when students interactively engage in discussions among each other. Accordingly, case-based learning has been found to be particularly beneficial in collaborative settings.[15] However, another important aspect to consider in collaborative learning environments is that some students may participate passively while others contribute disproportionately much. To foster optimal learning effects, students should thus be encouraged to be interactively engaged. One prerequisite to achieve self-guided learning in groups is a low threshold for students to come forward with their questions and participate in ensuing discussions.[18] To this end, peer teaching has been established as an effective tool to stimulate discussions.[19] To make sure peer tutors are not overwhelmed in moderating these discussions, the presence of an experienced clinician appears to be warranted[20] in addition to a specific training of the tutors.

Taken together, while traditional CPCs encompass some important dimensions of effective case-based learning environments, they are not systematically aiming at constructive or interactive learner activities that are known features of effective teaching

formats.[17,21] Therefore, we introduced Clinical Case Discussions (CCD) in undergraduate medical education to account for these features. We still use the Case Records of the Massachusetts General Hospital,[9] as these cases exemplify realistic patient encounters and fulfill the criteria for an interactive collaborative learning process as explained above. In the CCD approach, cases are typically presented with information until the admission of the patient to the hospital. This event is usually the starting point of an interactive discussion phase of the group about possible diagnoses and diagnostic strategies. After all test results have been discussed, the actual diagnosis is disclosed and the pitfalls and take-home messages of the case are summarised.

To investigate the effectiveness of the CCD approach in undergraduate medical education, we designed an intervention trial and assessed clinical reasoning skills in medical students before and after *participating in* live CCDs or *being exposed* to video recordings of live CCDs. We compared these formats and its effects on clinical reasoning with the more traditional approach of working through written cases. When carrying out this randomised trial, we hypothesised that participation in live CCD sessions would lead to a higher increase of clinical reasoning skills than simply reading the cases. To better understand possible effects of the CCD learning environment with its social components on learning outcomes, participation in live CCDs as outlined above was additionally compared to the effects of watching videos of CCDs online. This comparison also seemed relevant from an economic point of view as video-streaming of lectures and seminars are prevalent at many instituitions in higher education allowing for flexible and scalable access to learning materials.[22] To investigate the potential of different CCD formats for regular curricular use, we also

measured subjective learning outcomes after the intervention and correlated student self-assessments with objective changes in their clinical reasoning skills.

Methods

Participants / Ethics

Initially, we recruited 106 volunteer medical students at the Medical Faculty of LMU Munich. Randomisation was performed in a two-step procedure: First, we selected a sample of roughly 100 enrolled students. Next, we stratified participants by creating triplets on the basis of the variables age, gender, year of study, prior CCD participation and performance in a knowledge application pre-test. This was done in an effort to limit the risk of random misdistribution of the selected sample. From each triplet we randomly assigned participants to the experimental groups. A total of 90 participants eventually completed the study, 31 of them were male and 59 female. They were 20 to 41 years old (M = 23; SD = 2.97) and in their first to eighth clinical semester (M = 3.5; SD = 1.78).

The study was approved by the ethics committee of the Medical Faculty of LMU Munich (approval reference no. 222-15). Written informed consent was obtained from all study participants and they received a financial reimbursement of 50 Euros upon completion of the trial.

Patient and public involvement

No patients or public were involved in this research.

Study design

We conducted a single-center randomised controlled trial consisting of a total of five course sessions with a parallel design (see Figure 1). One week prior to the first CCD session, participants were introduced to the principles of the CCD approach and the sequence of this trial in an introductory session where they also took a knowledge application pre-test (T_0). In the experimental phase, participants attended three weekly interventional course sessions of 90 minutes each in one of the three aforementioned groups with the respective CCD formats. Participants took a knowledge application post-test at the end of the last experimental course session (T_1), four weeks after pre-testing. A delayed knowledge application post-test was conducted two weeks after completion of the interventional courses (T_2); we deliberately chose that time interval to investigate the sustainability of possible effects while balancing the risk of post-intervention confounding.[23]

****

Insert Figure 1 about here

****

Materials

In all experimental groups the intervention was based on the same three, independent internal medicine cases. Chief complaints in these cases were paresthesia (first session), fever and respiratory failure (second session), and rapidly progressive respiratory failure (third session).[24-26] Cases were worked through in an iterative approach in different formats: (a) peer-moderated live case discussions in an interactive setting (Live-CCD, *n* = 30), (b) a single-learner format utilizing an interactive multimedia platform displaying video recordings of the live case discussions (Video-CCD, *n* = 27), and (c) a single-learner format in which the students worked with the original paper cases of the NEJM

(Paper-Cases, $n = 33$). The cases were prepared in a way that participants in each format were exposed to the same case information.

Procedure

In all three groups cases were presented in a specified structured manner similar to the original Clinicopathological Conferences (see Figure 2). In each format the students ("discussants") had to fill out a form after the admission in which the case had to be summarised and a list of clinical problems and working diagnoses had to be provided. Subsequently, between discussion and summary a second case-summary had to be completed in which the final diagnostic test and the most likely diagnosis had to be proposed.

****

Insert Figure 2 about here

****

In the Live-CCD group, the case presentation was prepared beforehand by a voluntary discussant ("presenter"), who presented the facts in the admission (according to the structure shown in Figure 2). Electronic slides and flipcharts were used to transport case information. Original test results were revealed by the presenter during the discussion only when requested by the group of students. Furthermore, the presenter summarised the differential diagnosis, important pathophysiological features of the case at the end of the session and provided a short take home message. The moderating medical students ("moderator") were recruited among previous CCD participants. They had experience in CCD moderation and had had an introductory training (two days) in higher education methods and group facilitation prior to the study. The moderator facilitated the discussion process and ensured a reasonable approach to the patient encounter (e.g. with

respect to timing and hierarchy of ordered tests) in close communication with the discussants. Moreover, the moderator helped students develop their diagnostic strategy by co-evaluating their requested findings and the reasoning employed. Supervision of the correctness of medical facts and the correct diagnostic approach were ultimately granted by a clinician who could stop the discussion at any point when faulty reasoning was evident or discussants explicitly requested the facilitation of an experienced physician. The clinicians' level of involvement into the discussion was left at their own discretion. We varied the staff between each Live-CCD to minimise effects of personal teacher characteristics. Live sessions typically lasted 90 minutes and were recorded with multiple cameras.

Students in the Video-CCD format worked on a single-learner multimedia workstation on which a video recording of the Live-CCD were displayed. These recordings also contained the electronic slide presentation from the Live-CCD and enabled simultaneous observation of the discussion from multiple camera angles. Participants could pause and partially skip the videos.

In the Paper-Cases group participants received the case information of each CCD section sequentially (i.e. admission, discussion, summary) in a print format. In both single-learner formats students could choose their personal working speed. There was neither a prespecified minimum nor a maximum time they were required to work on the cases. In each of the three formats full access to the internet was permitted for additional information.

Instruments

Learning outcomes with respect to clinical reasoning were measured with a knowledge application test that consisted of 29 items (i.e. a maximum of 29 points could be achieved) and was to be filled out within 45 minutes. The knowledge application test was based on instruments previously developed at the Institute for Medical Education at LMU Munich.[27-29] It comprised multiple choice items, key feature problems and problem-solving tasks, addressing the conceptual, strategic, and conditional knowledge of the participants (see Figure 3). Meta-analyses on retest effects suggest that score increase is higher for identical forms than for parallel test forms.[30] In order to limit such effects, we applied parallel forms of the knowledge application test for pre- and post-measurements (i.e. topics covered by the individual items were the same, but the items were reformulated and their order was permutated). Overall test difficulty was chosen to be high in order to avoid ceiling effects, as students from all clinical years were allowed to participate in the study. Overall test reliability was satisfactory (Cronbach's $\alpha = .71$).

****

Insert Figure 3 about here

****

Subjective learning outcomes were measured at T_1 with a short questionnaire consisting of 9 items (e.g. "I learned a lot during the CCD course", "The CCD course increased my learning motivation" or "I recommend the implementation of the CCD teaching format into the curriculum"; the full questionnaire is available as a supplementary file). Participants were asked to rate these items on a Likert scale ranging from 1 (I don't agree) to 5 (I fully agree). Reliability of the corresponding scale was good (Cronbach's $\alpha = .95$). Additionally, study participants were asked to share their

views on positive and negative aspects of the respective training format through open items at the end of the questionnaire.

Statistical Analysis

The required sample size (N = 128) was estimated to detect medium effect sizes with a power of 80% and a significance level of $\alpha = .05$. For between-group analyses, one-way ANOVAs were conducted with *post-hoc* Bonferroni tests for multiple comparisons.

Results

Effects of the CCD format on learning outcomes related to Clinical Reasoning

Experimental groups differed significantly with respect to the knowledge application post-test (see Table 1), $F(2,87) = 27.07$, $p = .000$, partial $\eta^2 = .384$. The Live-CCD group ($M = 14.10$; $SD = 3.32$) outperformed both the Video-CCD ($M = 11.69$; $SD = 3.34$) and the Paper-Cases group ($M = 8.5$; $SD = 2.44$). Post hoc Bonferroni tests revealed significant differences between Live-CCD and Video-CCD ($p = .011$) as well as the Paper-Cases group ($p = .000$). The difference in the knowledge application post-test between Video-CCD and the Paper-Cases group was also significant ($p = .000$).

Two weeks after course completion, the effect of the teaching format was still found in a delayed knowledge application post-test, $F(2,87) = 30.91$, $p = .000$, partial $\eta^2 = .415$. Both Live-CCD ($M = 13.36$; $SD = 3.23$) and the Video-CCD ($M = 11.84$; $SD = 2.92$) outperformed the Paper-Cases group ($M = 7.89$; $SD = 2.41$). Post hoc Bonferroni tests revealed significant differences between the Live-CCD and Paper-Cases group ($p = .000$) as well as between the Video-CCD and Paper-Cases group ($p = .000$). However,

the difference between Live-CCD and Video-CCD was not significant in the delayed knowledge application post-test (p = .146).

Table 1. Overview of the findings of the study.

	Teaching format					
	Live-CCD		Video-CCD		Paper-Cases	
	M	(SD)	M	(SD)	M	(SD)
Knowledge application pre-test	5.34	(1.92)	4.76	(1.90)	5.76	(2.24)
	n = 30		n = 27		n = 33	
Knowledge application post-test	14.10	(3.32)	11.69	(3.34)	8.50	(2.44)
	n = 30		n = 27		n = 33	
Delayed knowledge application post-test	13.36	(3.23)	11.84	(2.92)	7.89	(2.41)
	n = 30		n = 27		n = 33	
Subjective learning outcomes	4.20	(.63)	3.18	(1.24)	3.00	(.99)
	n = 30		n = 27		n = 31	

Effects of the CCD format on subjective learning outcomes

Experimental groups differed significantly with respect to subjective learning outcomes (see Table 1), $F(2,85) = 13.16$, p = .000, partial η^2 = .236. Participants of the Live-CCD group (M = 4.20; SD = .63) assigned better ratings to their course format than participants in the Video-CCD group (M = 3.18; SD = 1.24) and the Paper-Cases group (M = 3.0; SD = .99). Post hoc Bonferroni tests showed that the Live-CCD differed from the Video-CCD (p = .001) and the Paper-Cases group (p = .000) in this regard. An additional Duncan post-hoc test confirmed that the Video-CCD and the Paper-Cases group did not differ from each other in this regard (p = .48).

To investigate the relations between the subjective assessment and the knowledge application tests applied at the end and two weeks after the course, we calculated correlations between the different outcome measures. Subjective learning outcomes correlated on a medium level with both the knowledge application post-test (r = .343, n = 88, p = .001) and the delayed knowledge application post-test (r = .339, n = 88, p = .001).

In the Live-CCD group, 83% of the students were in favour of implementing routine Live-CCD into the medical curriculum. Only 45% and 31% of students from the Video-CCD and Paper-Cases groups voted for an implementation of their respective course in the curriculum. With respect to the open items from the subjective learning outcomes questionnaire, participants from all groups praised the quality of the cases. Participants from the Live-CCD group particularly valued their course format for providing an opportunity to practice "diagnostic thinking" and the "focus on practice elements". They also mentioned that "you can look up theoretical knowledge, but you can't look up applied knowledge". Students in the Video-CCD group, on the other hand, praised features of the digital learning environment as they could "pause, reflect, or quickly do a Google search" when watching the case discussions. However, they also criticised it was not possible for them to "participate in a more active way".

Discussion

This randomised controlled study shows that even relatively short CCD interventions can lead to improved and sustainable learning outcomes with respect to clinical reasoning. This provides evidence that the CCD approach, which is based on Clinicopathological Conferences, is an effective teaching resource to foster clinical

reasoning skills in medical students. We had hypothesised that a more interactive course format would result in an improvement of clinical reasoning skills when compared with less interactive formats. Results show that the Live-CCD indeed leads to the highest learning outcomes in medical students compared to less interactive formats. Consistent with our hypothesis, clinical reasoning skills, as measured with our knowledge application test, had the highest gain in the Live-CCD group. These positive effects of the CCD teaching format on clinical reasoning skills proved sustainable as shown by the results in the delayed knowledge application post-test. Overall, these results are in line with a recently published study on diagnostic reasoning[31] where students who worked in pairs were more accurate in their diagnosis than individual students despite having comparable knowledge. Collaborative clinical reasoning has thus far been underrepresented in the literature, yet seems to solve many of the educational problems regarding diagnostic errors.[32]

The significant difference between the Live-CCD and the Video-CCD group can be explained by the findings of a meta-analysis that showed technology-assisted single-person learning to be inferior to group learning because of the decreased social interaction.[33] However, it is important to note that two weeks after the course, participants of the Live-CCD and Video-CCD groups did not differ significantly anymore while both groups still clearly outperformed the Paper-Case group. In other words, watching a video of the live case discussion was found to be more beneficial for learners regarding their clinical reasoning skills than just reading the printed cases. We cannot rule out that Live-CCD and Video-CCD groups did not differ in the delayed knowledge application post-test due to underpowering of the study. As our trial was not designed to detect smaller effect sizes, this finding has to be treated with caution.

Subjective learning outcomes suggest that students prefer the live discussion over the other formats. The subjective assessment correlated with the students' performance in both knowledge application post-tests. Additional qualitative data from the open item answers suggests that the Live-CCD format supported students in performing clinical reasoning and that the active discussion of cases was particularly valued by the students.

Generalisability

The conclusions of this study are applicable to a broader audience of medical students. The CCD approach and its respective formats can easily be implemented in routine medical education. Peer teaching courses hold the promise of being more easy to install and more easy to staff then courses led by faculty. Of course, live CCDs still come with certain personnel requirements, as faculty as well as a moderator need to be present. Extensive preparation was not necessary for the clinicians involved though as they served as facilitators and provided guidance only in situations when they were explicitly asked for their clinical judgement or when they felt that the discussion went astray. Total time requirements might still be lower compared to other teaching formats. Likewise, the implementation of a singular two-day training for moderators should not require extensive ressources. The study population consisting of students with heterogeneous levels of clinical experience implies that the CCD is an effective teaching format not only for students at the beginning of their clinical career but also for intermediate students. On the other hand, generalisability is potentially limited as only students from one medical school participated in our study.

Limitations of the study

There are certain limitations of this study that have to be addressed: One important limitation is the single-centre nature of this study and the relatively small sample size. Before the CCD approach can be implemented on a larger scale, a validation of our findings is therefore required. Caution is clearly warranted with the effect sizes shown in this trial, as it has been shown that effect sizes of learning intervention trials tend to be inflated compared to the effectiveness of the intervention when used in routine education.[34] Since we did not limit the time students had to work on the cases, we cannot entirely rule out that less time was spent on task in the single-learner formats and particularly the Paper-Cases group. Against this backdrop, we suggest replication to further validate the results found in this study and strengthen the outlined implications. The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge). Larger item numbers could facilitate a reliable assessment of changes on the level of corresponding subscales. Finally, we cannot relate the underlying reasoning process with the measured knowledge gains. Further studies on clinical reasoning processes of individuals and groups are methodologically challenging but urgently needed for the advancement of a model of clinical reasoning and for improving teaching clinical reasoning.[35]

Implications for policy makers / Future research questions

Based on our findings, the CCD approach is a useful asset for medical educators to widen the range of clinical reasoning teaching tools. Live-CCD can thus be seen as a prime candidate for routine implementation in clinical reasoning curricula. Future

research should aim to identify which Live-CCD elements (the roles, case contents, or the course structure) contribute in which way to the improvement of clinical reasoning skills in medical students. The question if and to what extent such skills are applicable across domains is currently being discussed.[36] Future studies may also address the issue of transfer (i.e. to what extent can clinical reasoning skills obtained in case-based training later be applied to different cases?).[37] Regarding the Video-CCD, means of instructional support to increase the effectiveness and interactivity of the video-based format should be investigated in an attempt to exploit its full potential.

Acknowledgements

The authors thank Johanna Huber and her team for technical support with the evaluation, Thomas Brendel and Thomas Bischoff for help with the video production and Mark S. Pecker for critical reading of our manuscript and valuable suggestions. The authors also thank the CCD student discussants and moderators for their contributions. We wish to sincerely address our gratitude to the CCD team for organisational support with the study: Nora Koenemann, Simone Reichert, Sandra Petrenz, Fabian Haak, Bjoern Stolte, Simon Berhe, Bastian Brandt, and Thomas Lautz. Marc Weidenbusch wishes to express special thanks to Bernd Gansbacher for introduction to CCDs.

Funding statement

This work was supported by the German Federal Ministry of Education and Research (grant no. 01PL12016) and an intramural grant of the Medical Faculty of the University of Munich (Lehre@LMU).

Competing Interests

Marc Weidenbusch declares to have no conflict of interest.

Benedikt Lenzer declares to have no conflict of interest.

Maximilian Sailer declares to have no conflict of interest.

Christian Strobel declares to have no conflict of interest.

Raphael Kunisch declares to have no conflict of interest.

Jan Kiesewetter declares to have no conflict of interest.

Martin R. Fischer declares to have no conflict of interest.

Jan M. Zottmann declares to have no conflict of interest.

Author contributions

MW, BL, MF and JZ planned the study.

MW, BL and CS were responsible for data acquisition.

MW, BL, RK, JK, JZ, MF and MS analysed and interpreted the data.

MW, BL and JZ drafted and revised the manuscript, all authors contributed significant intellectual content and all authors gave final approval of the version to be published.

Data sharing statement

Dataset and detailed information about the CCD formats is available upon request.

References

1. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: Theory to practice. Medical Teacher. 2010;32(8):638-45.

2. Fischer MR, Bauer D, Mohn K, et al. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. GMS Journal for Medical Education. 2015;32(3):Doc35.

3. Harasym PH, Tsai TC, Hemmati P. Current trends in developing medical students' critical thinking abilities. The Kaohsiung Journal of Medical Sciences. 2008;24(7):341-55.

4. Donaldson MS, Corrigan JM, Kohn LT. To err is human: Building a safer health system. Washington: National Academies Press; 2000.

5. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing: let's preach what we practice. New England Journal of Medicine. 1983;309(15):921-3.

6. Hege I, Ropp V, Adler M, Radon K, Mäsch G, Lyon H, et al. Experiences with different integration strategies of case-based e-learning. Medical Teacher. 2007;29(8):791-7.

7. Cannon WB. The Case Method of Teaching Systematic Medicine. The Boston Medical and Surgical Journal. 1900;142(2):31-6.

8. Eva KW. What every teacher needs to know about clinical reasoning. Medical Education. 2005;39(1):98-106.

9. Harris NL. Case records of the Massachusetts General Hospital – Continuing to learn from the patient. New England Journal of Medicine. 2003;348(22):2252-4.

10. Cabot RC. Case teaching in medicine: A series of graduated exercises in the differential diagnosis, prognosis and treatment of actual cases of disease. Boston: Heath; 1906.

11. Sturdy S. Knowing Cases: Biomedicine in Edinburgh, 1887-1920. Social Studies of Science. 2007;37(5):659-89.

12. Relman AS. Two Views. New England Journal of Medicine. 1979;301(20):1112-3.

13. Ertmer PA, Newby TJ. Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance Improvement Quarterly. 1993;6(4):50-72.

14. Merseth KK. Cases and case methods in teacher education. In: Sikula J, Buttery TJ, Guyton R, editors. Handbook of research on teacher education. 2nd ed. New York, NY: Macmillan; 1996. p. 722-744.

15. Zottmann JM, Stegmann K, Strijbos J-W, Vogel F, Wecker C, Fischer F. Computer-supported collaborative learning with digital video cases in teacher education: The impact of teaching experience on knowledge convergence. Computers in Human Behavior. 2013;29(5):2100-8.

16. Powers BW, Navathe AS, Jain SH. Medical education's authenticity problem. BMJ. 2014;348:g2651.

17. Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist. 2014;49(4):219-43.

18. Duncan RG, Rivet AE. Science learning progressions. Science. 2013;339(6118):396-7.

19. De Menezes S, Premnath D. Near-peer education: A novel teaching program. International journal of medical education. 2016;7:160-167.

20. Ince-Cushman D, Rudkin T, Rosenberg E. Supervised near-peer clinical teaching in the ambulatory clinic: an exploratory study of family medicine residents' perspectives. Perspectives on medical education. 2015;4(1):8-13.

21. Chi MTH. Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science. 2009;1(1):73-105.

22. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. Academic Medicine. 2006;81(3):207-12.

23. Miller DC, Salkind NJ, editors. Handbook of Research Design & Social Measurement. Thousand Oaks, CA: SAGE; 2002.

24. Kotton DN, Muse VV, Nishino M. Case 2-2012. New England Journal of Medicine. 2012;366(3):259-69.

25. Marks PW, Zukerberg LR. Case 30-2004. New England Journal of Medicine. 2004;351(13):1333-41.

26. Uyeki TM, Sharma A, Branda JA. Case 40-2009. New England Journal of Medicine. 2009;361(26):2558-69.

27. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, et al. Representation scaffolds improve diagnostic efficiency in medical students. Medical Education. 2017; 51(11):1118-1126.

28. Kopp V, Stark R, Kühne-Eversmann L, Fischer MR. Do worked examples foster medical students' diagnostic knowledge of hyperthyroidism? Medical Education. 2009;43(12):1210-7.

29. Schmidmaier R, Eiber S, Ebersbach R, Schiller M, Hege I, Holzer M, et al. Learning the facts in medical school is not enough: Which factors predict successful application of procedural knowledge in a laboratory setting? BMC Medical Education. 2013;13:28.

30. Hausknecht JP, Halbert JA, Di Paolo NT, Moriarty Gerrard MO. Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. Journal of Applied Psychology. 2007;92(2):373–385.

31. Hautz WE, Kämmer JE, Schauber SK, Spies CD, Gaissmaier W. Diagnostic performance by medical students working individually or in teams. JAMA. 2015;313(3):303-4.

32. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Medical Education. 2015;49(10):961-73.

33. Lou Y, Abrami PC, d'Apollonia S. Small group and individual learning with technology: A meta-analysis. Review of Educational Research. 2001;71(3):449-521.

34. Springer L, Stanne ME, Donovan SS. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of Educational Research. 1999;69(1):21-51.

35. Heitzmann N, Fischer MR, Fischer F. Towards more systematic and better theorised research on simulations. Medical Education. 2017;51(2):129-131.

36. Fischer F, Chinn C, Engelmann K, Osborne J, editors. Scientific reasoning and argumentation. The roles of general and specific knowledge. New York, NY: Routledge; 2018.

37. Keemink Y, Custers E, van Dijk S, ten Cate O. Illness script development in pre-clinical education through case-based clinical reasoning training. International Journal of Medical Education. 2018;9:35-41.

Figure legends

Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the *admission* part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive *discussion* part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. Abbreviations: CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

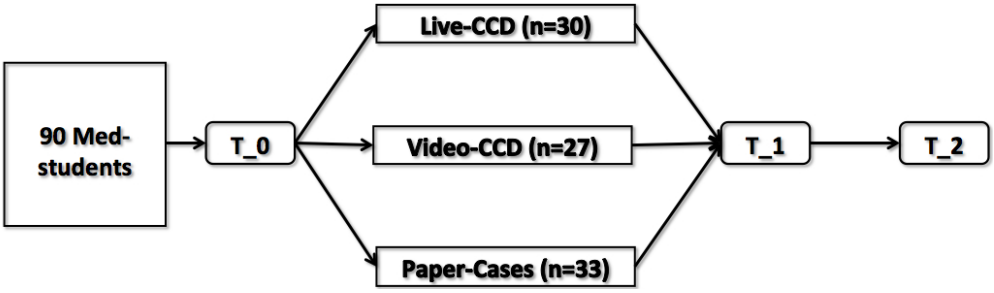


Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

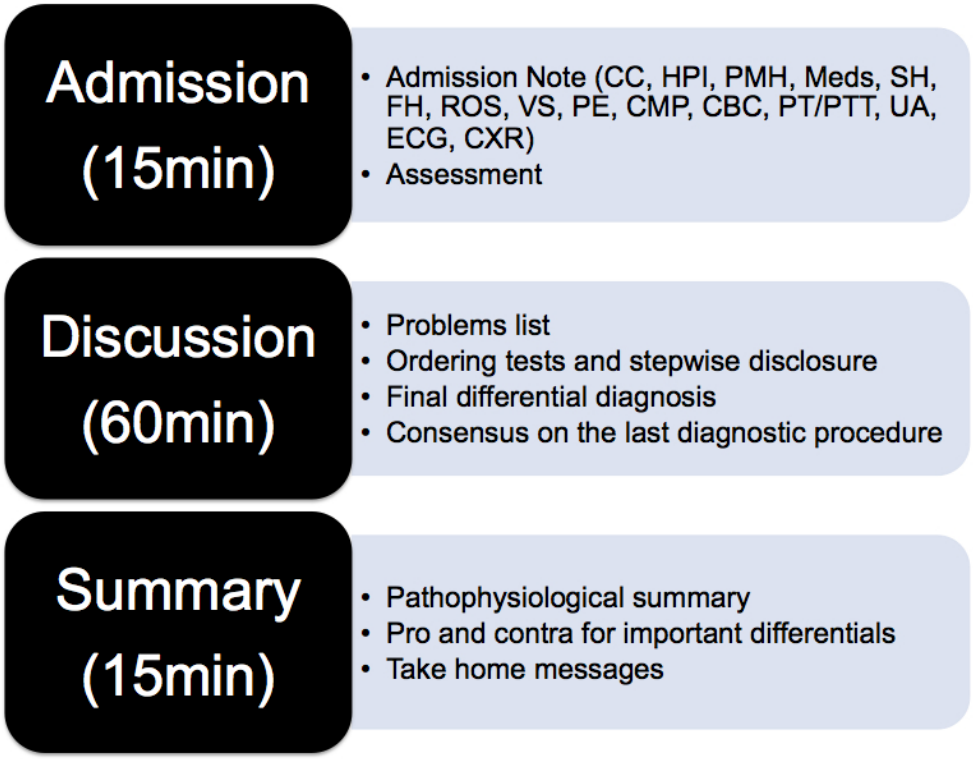


Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the admission part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive discussion part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. Abbreviations: CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

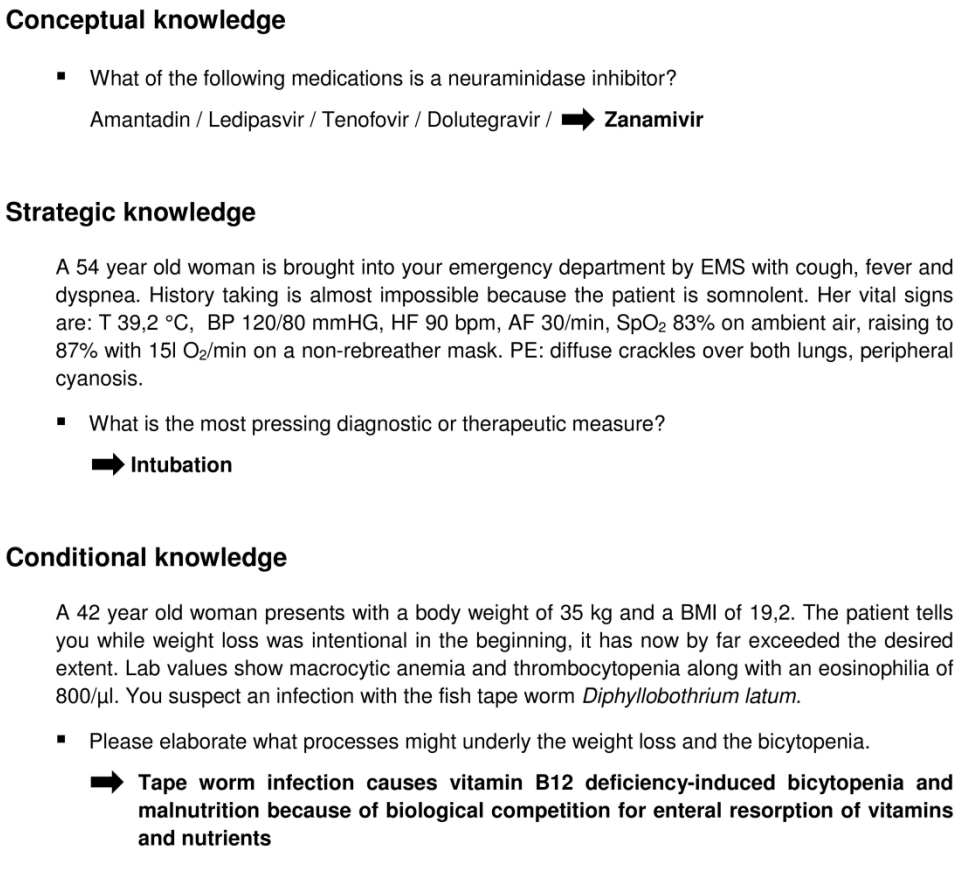


Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

150x135mm (600 x 600 DPI)

Supplemental

Baseline data: Demographic characteristics of the study participants.

	CCD format							
	Live-CCD *n* = 30		Video-CCD *n* = 27		Paper-Cases *n* = 33		All formats *N* = 90	
Gender distribution *N* m/f (% f)	14/16 (53.3% f)		5/22 (81.5% f)		12/21 (63.6% f)		31/59 (65.6% f)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Age	23.77	(4.09)	22.26	(1.77)	22.91	(2.40)	23.0	(2.97)
Clinical semester	3.23	(1.96)	3,41	(1.47)	3.82	(1.84)	3.50	(1.78)
High school grade	1.53	(0.36)	1,35	(0.42)	1.48	(0.68)	1.46	(0.52)
First National Board Exam Score	245	(30)	226	(78)	246	(30)	240	(49)
Participants with prior CCD experience n (%)	6	(20.0%)	6	(22.2%)	5	(15.1%)	17	(18.9%)

Questionnaire items (5-point Likert scale) for the assessment of subjective learning outcomes.

1. I perceived this CCD format as meaningful.
2. I learned a lot during the CCD course.
3. The CCD course increased my learning motivation.
4. I would like to participate in this CCD format again in the future.
5. I enjoyed the CCD course.
6. This CCD format should be offered as part of the curriculum.
7. I was able to follow the case discussions.
8. Learning in the CCD format is easier for me than learning in traditional lectures or seminars.
9. How would you rate the course overall?

CONSORT 2010 checklist of information to include when reporting a randomised trial*

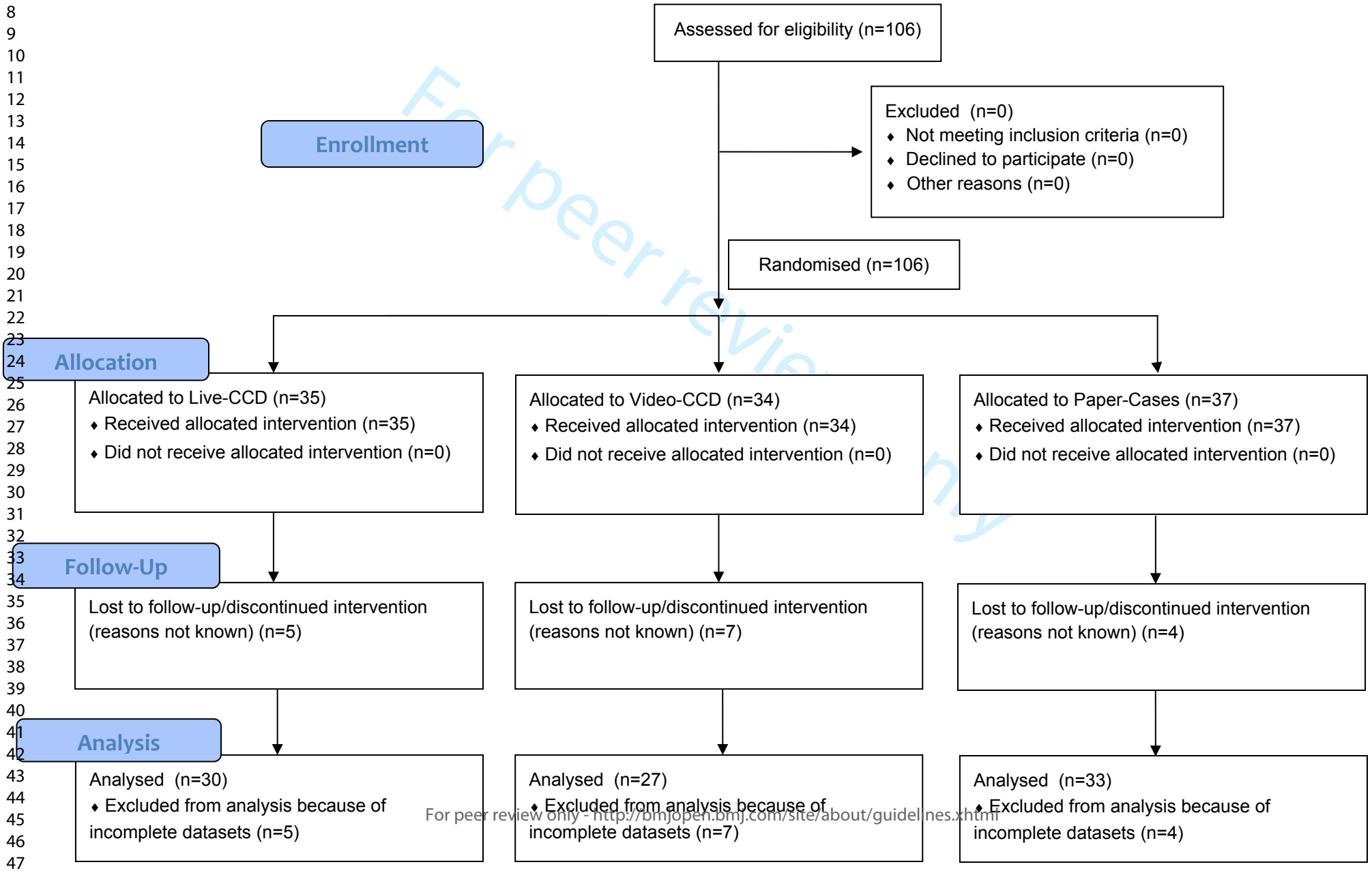
Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-7
	2b	Specific objectives or hypotheses	6-7
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	8
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	not applicable
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	8-10
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8, 11-12
	6b	Any changes to trial outcomes after the trial commenced, with reasons	not applicable
Sample size	7a	How sample size was determined	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines	not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	7
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	7
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	not applicable
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	not applicable
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	not applicable

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	attached as supplemental
	13b	For each group, losses and exclusions after randomisation, together with reasons	7
Recruitment	14a	Dates defining the periods of recruitment and follow-up	not applicable
	14b	Why the trial ended or was stopped	not applicable
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	attached as supplemental
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	13
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	12-14
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	not applicable
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	not applicable
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	not applicable
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	17
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	16
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	14-16
Other information			
Registration	23	Registration number and name of trial registry	not applicable
Protocol	24	Where the full trial protocol can be accessed, if available	19
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	18

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



BMJ Open

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025973.R3
Article Type:	Research
Date Submitted by the Author:	11-Jul-2019
Complete List of Authors:	Weidenbusch, Marc; University Hospital of LMU Munich, Institute for Medical Education; University Hospital of LMU Munich, Department of Internal Medicine IV Lenzer, Benedikt; University Hospital of LMU Munich, Institute for Medical Education Sailer, Maximilian; University of Passau, Department of Education Strobel, Christian; University Hospital of LMU Munich, Institute for Medical Education Kunisch, Raphael; University Hospital of LMU Munich, Institute for Medical Education Kiesewetter, Jan; University Hospital of LMU Munich, Institute for Medical Education Fischer, Martin; University Hospital of LMU Munich, Institute for Medical Education Zottmann, Jan; University Hospital of LMU Munich, Institute for Medical Education
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Medical education and training
Keywords:	Undergraduate medical education, Case-based learning, Clinical reasoning, Social interaction, Medical decision making

RESEARCH PAPER

Can Clinical Case Discussions foster clinical reasoning skills in undergraduate medical education? A randomised controlled trial

Dr. med. Marc Weidenbusch[1,2], Benedikt Lenzer[2*], Prof. Dr. phil. Maximilian Sailer[3], Dr. phil. Christian Strobel[2], Raphael Kunisch[2], Dr. phil. Jan Kiesewetter[2], Prof. Dr. med. Martin R. Fischer[2] & Dr. phil. Jan M. Zottmann[2]*

[1]Nephrologisches Zentrum, Medizinische Klinik und Poliklinik IV, Klinikum der Universität München, LMU München

[2]Institut für Didaktik und Ausbildungsforschung in der Medizin, Klinikum der Universität München, LMU München

[3]Department of Education, University of Passau

[*]equal contribution

Corresponding author:

Dr. phil. Jan M. Zottmann

Institut für Didaktik und Ausbildungsforschung in der Medizin

Klinikum der Universität München, LMU München

Pettenkoferstr. 8a, D-80336 Munich, Germany

Tel: +49 89 4400 57203

Fax: +49 89 4400 57202

E-Mail: jan.zottmann@med.uni-muenchen.de

Word count: 3540 words

Keywords: undergraduate medical education, case-based learning, clinical reasoning, social interaction, medical decision making

Abstract

Objective: Fostering clinical reasoning is a mainstay of medical education. Based on the Clinicopathological Conferences, we propose a case-based peer teaching approach called Clinical Case Discussions (CCDs) to promote the respective skills in medical students. This study compares the effectiveness of different CCD formats with varying degrees of social interaction in fostering clinical reasoning.

Design, setting, participants: A single-center randomised controlled trial with a parallel design was conducted at a German university. Study participants (N=106) were stratified and tested regarding their clinical reasoning skills right after CCD participation and two weeks later.

Intervention: Participants worked either within a live discussion group (Live-CCD), a group watching recordings of the live discussions (Video-CCD), or a group working with printed cases (Paper-Cases). The presentation of case information followed an admission-, discussion-, summary-sequence.

Primary and secondary outcome measures: Clinical reasoning skills were measured with a knowledge application test addressing the students' conceptual, strategic, and conditional knowledge. Additionally, subjective learning outcomes were assessed.

Results: With respect to learning outcomes, the Live-CCD group displayed the best results, followed by Video-CCD and Paper-Cases, $F(2,87)=27.07$, $p<0.001$, partial $\eta^2=0.384$. No difference was found between Live-CCD and Video-CCD groups in the delayed post-test; however, both outperformed the Paper-Cases group, $F(2,87)=30.91$, $p<0.001$, partial $\eta^2=0.415$. Regarding subjective learning outcomes, the Live-CCD received significantly better ratings than the other formats, $F(2,85)=13.16$, $p<0.001$, partial $\eta^2=0.236$.

Conclusions: This study demonstrates that the CCD approach is an effective and sustainable clinical reasoning teaching resource for medical students. Subjective learning outcomes underline the importance of learner (inter-)activity in the acquisition of clinical reasoning skills in the context of case-based learning. Higher efficacy of more interactive formats can be attributed to positive effects of collaborative learning. Future research should investigate how the Live-CCD format can further be improved and how video-based CCDs can be enhanced through instructional support.

Article summary

Strengths and limitations of this study:

- First empirical study on the implementation of Clinical Case Discussions in undergraduate medical education.
- Comparison of Clinical Case Discussions with differing grades of social interaction to determine their effectiveness on medical students' acquisition of clinical reasoning skills by between-group analyses.
- Implementation of multidimensional and multilayered test instruments in a pre-, post- and delayed post-test design to measure clinical reasoning skills by a knowledge application test and self-assessment.
- The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge).

Introduction

Curriculum developers face the challenge of implementing competence-oriented frameworks such as CanMEDS (Canada), the NKLM (Germany) or PROFILES (Switzerland), including the need to train clinical reasoning skills as a medical doctor's key competence.[1-3] As such, clinical reasoning skills are crucial not only for appropriate medical decision making, but also to avoid diagnostic errors and the associated harm for both patients and healthcare systems.[4]

Case-based learning has been proposed to foster clinical reasoning skills[5] and is well accepted among students.[6] Case-based learning found an early representation in Clinicopathological Conferences (CPC, first introduced by Cannon in 1900[7]) which are practiced until today. The Clinicopathological Conferences conducted at the Massachusetts General Hospital are published on a regular basis known as the *Case Records* series of the New England Journal of Medicine. In those CPCs the "medical mystery"[8] presented by the case under discussion calls readers to think about the possible diagnosis themselves, before it is finally disclosed at the last part of the CPC. Despite the absence of definitive evidence for efficacy as a teaching method, CPCs have widely been used in medical education since the early 20th century to foster clinical reasoning.[9-11] While these CPC-Case Records reaches lots of medical readers around the world, it has been criticised as being anachronistic with a diagnosing "star (i.e. the discussant), performing, acutely aware of being the center of attention".[12]

Case-based learning formats are embedded in a context, which is known to promote learning better than providing facts in an abstract, non-contextual form.[13] A definition found in the review by Merseth suggests three essential elements of a case: A case is real (i.e. based on a real-life situation or event); it relies on careful research and study; it

is “created explicitly for discussion and seeks to include sufficient detail and information to elicit active analysis and interpretation by users”.[14] Cases may be represented by means of text, pictures, videos, and the like. Realism and authenticity are varying features of cases,[15] but particularly elaborated and authentic cases provide increased diagnostic challenge, comprising added value for medical training.[16]

However, due to their setup, CPCs are often a passive learning situation for participants, as they listen to the discussant laying out his or her clinical reasoning on the case under discussion. According to the ICAP framework by Chi et al.,[17] teaching formats increase their efficacy from passive < active < constructive < interactive learning environments. Learning is enhanced when students interactively engage in discussions among each other. Accordingly, case-based learning has been found to be particularly beneficial in collaborative settings.[15] However, another important aspect to consider in collaborative learning environments is that some students may participate passively while others contribute disproportionately much. To foster optimal learning effects, students should thus be encouraged to be interactively engaged. One prerequisite to achieve self-guided learning in groups is a low threshold for students to come forward with their questions and participate in ensuing discussions.[18] To this end, peer teaching has been established as an effective tool to stimulate discussions.[19] To make sure peer tutors are not overwhelmed in moderating these discussions, the presence of an experienced clinician appears to be warranted[20] in addition to a specific training of the tutors.

Taken together, while traditional CPCs encompass some important dimensions of effective case-based learning environments, they are not systematically aiming at constructive or interactive learner activities that are known features of effective teaching

formats.[17,21] Therefore, we introduced Clinical Case Discussions (CCD) in undergraduate medical education to account for these features. We still use the Case Records of the Massachusetts General Hospital,[9] as these cases exemplify realistic patient encounters and fulfill the criteria for an interactive collaborative learning process as explained above. In the CCD approach, cases are typically presented with information until the admission of the patient to the hospital. This event is usually the starting point of an interactive discussion phase of the group about possible diagnoses and diagnostic strategies. After all test results have been discussed, the actual diagnosis is disclosed and the pitfalls and take-home messages of the case are summarised.

To investigate the effectiveness of the CCD approach in undergraduate medical education, we designed an intervention trial and assessed clinical reasoning skills in medical students before and after *participating in* live CCDs or *being exposed* to video recordings of live CCDs. We compared these formats and its effects on clinical reasoning with the more traditional approach of working through written cases. When carrying out this randomised trial, we hypothesised that participation in live CCD sessions would lead to a higher increase of clinical reasoning skills than simply reading the cases. To better understand possible effects of the CCD learning environment with its social components on learning outcomes, participation in live CCDs as outlined above was additionally compared to the effects of watching videos of CCDs online. This comparison also seemed relevant from an economic point of view as video-streaming of lectures and seminars are prevalent at many instituitions in higher education allowing for flexible and scalable access to learning materials.[22] To investigate the potential of different CCD formats for regular curricular use, we also

measured subjective learning outcomes after the intervention and correlated student self-assessments with objective changes in their clinical reasoning skills.

Methods

Participants / Ethics

Initially, we recruited 106 volunteer medical students at the Medical Faculty of LMU Munich. Randomisation was performed in a two-step procedure: First, we selected a sample of roughly 100 enrolled students. Next, we stratified participants by creating triplets on the basis of the variables age, gender, year of study, prior CCD participation and performance in a knowledge application pre-test. This was done in an effort to limit the risk of random misdistribution of the selected sample. From each triplet we randomly assigned participants to the experimental groups. A total of 90 participants eventually completed the study, 31 of them were male and 59 female. They were 20 to 41 years old (M=23; SD=2.97) and in their first to eighth clinical semester (M=3.50; SD=1.78).

The study was approved by the ethics committee of the Medical Faculty of LMU Munich (approval reference no. 222-15). Written informed consent was obtained from all study participants and they received a financial reimbursement of 50 Euros upon completion of the trial.

Patient and public involvement

No patients or public were involved in this research.

Study design

We conducted a single-center randomised controlled trial consisting of a total of five course sessions with a parallel design (see Figure 1). One week prior to the first CCD session, participants were introduced to the principles of the CCD approach and the sequence of this trial in an introductory session where they also took a knowledge application pre-test (T_0). In the experimental phase, participants attended three weekly interventional course sessions of 90 minutes each in one of the three aforementioned groups with the respective CCD formats. Participants took a knowledge application post-test at the end of the last experimental course session (T_1), four weeks after pre-testing. A delayed knowledge application post-test was conducted two weeks after completion of the interventional courses (T_2); we deliberately chose that time interval to investigate the sustainability of possible effects while balancing the risk of post-intervention confounding.[23]

****

Insert Figure 1 about here

****

Materials

In all experimental groups the intervention was based on the same three, independent internal medicine cases. Chief complaints in these cases were paresthesia (first session), fever and respiratory failure (second session), and rapidly progressive respiratory failure (third session).[24-26] Cases were worked through in an iterative approach in different formats: (a) peer-moderated live case discussions in an interactive setting (Live-CCD, *n*=30), (b) a single-learner format utilizing an interactive multimedia platform displaying video recordings of the live case discussions (Video-CCD, *n*=27), and (c) a single-learner format in which the students worked with the original paper cases of the

NEJM (Paper-Cases, *n*=33). The cases were prepared in a way that participants in each format were exposed to the same case information.

Procedure

In all three groups cases were presented in a specified structured manner similar to the original Clinicopathological Conferences (see Figure 2). In each format the students ("discussants") had to fill out a form after the admission in which the case had to be summarised and a list of clinical problems and working diagnoses had to be provided. Subsequently, between discussion and summary a second case-summary had to be completed in which the final diagnostic test and the most likely diagnosis had to be proposed.

****

Insert Figure 2 about here

****

In the Live-CCD group, the case presentation was prepared beforehand by a voluntary discussant ("presenter"), who presented the facts in the admission (according to the structure shown in Figure 2). Electronic slides and flipcharts were used to transport case information. Original test results were revealed by the presenter during the discussion only when requested by the group of students. Furthermore, the presenter summarised the differential diagnosis, important pathophysiological features of the case at the end of the session and provided a short take home message. The moderating medical students ("moderator") were recruited among previous CCD participants. They had experience in CCD moderation and had had an introductory training (two days) in higher education methods and group facilitation prior to the study. The moderator facilitated the discussion process and ensured a reasonable approach to the patient encounter (e.g. with

respect to timing and hierarchy of ordered tests) in close communication with the discussants. Moreover, the moderator helped students develop their diagnostic strategy by co-evaluating their requested findings and the reasoning employed. Supervision of the correctness of medical facts and the correct diagnostic approach were ultimately granted by a clinician who could stop the discussion at any point when faulty reasoning was evident or discussants explicitly requested the facilitation of an experienced physician. The clinicians' level of involvement into the discussion was left at their own discretion. We varied the staff between each Live-CCD to minimise effects of personal teacher characteristics. Live sessions typically lasted 90 minutes and were recorded with multiple cameras.

Students in the Video-CCD format worked on a single-learner multimedia workstation on which a video recording of the Live-CCD were displayed. These recordings also contained the electronic slide presentation from the Live-CCD and enabled simultaneous observation of the discussion from multiple camera angles. Participants could pause and partially skip the videos.

In the Paper-Cases group participants received the case information of each CCD section sequentially (i.e. admission, discussion, summary) in a print format. In both single-learner formats students could choose their personal working speed. There was neither a prespecified minimum nor a maximum time they were required to work on the cases. In each of the three formats full access to the internet was permitted for additional information.

Instruments

Learning outcomes with respect to clinical reasoning were measured with a knowledge application test that consisted of 29 items (i.e. a maximum of 29 points could be achieved) and was to be filled out within 45 minutes. The knowledge application test was based on instruments previously developed at the Institute for Medical Education at LMU Munich.[27-29] It comprised multiple choice items, key feature problems and problem-solving tasks, addressing the conceptual, strategic, and conditional knowledge of the participants (see Figure 3). Meta-analyses on retest effects suggest that score increase is higher for identical forms than for parallel test forms.[30] In order to limit such effects, we applied parallel forms of the knowledge application test for pre- and post-measurements (i.e. topics covered by the individual items were the same, but the items were reformulated and their order was permutated). Overall test difficulty was chosen to be high in order to avoid ceiling effects, as students from all clinical years were allowed to participate in the study. Overall test reliability was satisfactory (Cronbach's α=0.71).

****

Insert Figure 3 about here

****

Subjective learning outcomes were measured at T_1 with a short questionnaire consisting of 9 items (e.g. "I learned a lot during the CCD course", "The CCD course increased my learning motivation" or "I recommend the implementation of the CCD teaching format into the curriculum"; the full questionnaire is available as a supplementary file). Participants were asked to rate these items on a Likert scale ranging from 1 (I don't agree) to 5 (I fully agree). Reliability of the corresponding scale was good (Cronbach's α=0.95). Additionally, study participants were asked to share their

views on positive and negative aspects of the respective training format through open items at the end of the questionnaire.

Statistical Analysis

The required sample size (N=128) was estimated to detect medium effect sizes with a power of 80% and a significance level of α=0.05. For between-group analyses, one-way ANOVAs were conducted with *post-hoc* Bonferroni tests for multiple comparisons.

Results

Effects of the CCD format on learning outcomes related to Clinical Reasoning

Experimental groups differed significantly with respect to the knowledge application post-test (see Table 1), $F(2,87)=27.07$, $p<0.001$, partial $\eta^2=0.384$. The Live-CCD group (M=14.10; SD=3.32) outperformed both the Video-CCD (M=11.69; SD=3.34) and the Paper-Cases group (M=8.50; SD=2.44). Post hoc Bonferroni tests revealed significant differences between Live-CCD and Video-CCD (p=0.011) as well as the Paper-Cases group ($p<0.001$). The difference in the knowledge application post-test between Video-CCD and the Paper-Cases group was also significant ($p<0.001$).

Two weeks after course completion, the effect of the teaching format was still found in a delayed knowledge application post-test, $F(2,87)=30.91$, $p<0.001$, partial $\eta^2=0.415$. Both Live-CCD (M=13.36; SD=3.23) and the Video-CCD (M=11.84; SD=2.92) outperformed the Paper-Cases group (M=7.89; SD=2.41). Post hoc Bonferroni tests revealed significant differences between the Live-CCD and Paper-Cases group ($p<0.001$) as well as between the Video-CCD and Paper-Cases group ($p<0.001$).

However, the difference between Live-CCD and Video-CCD was not significant in the delayed knowledge application post-test (p=0.146).

Table 1. Overview of the findings of the study.

	Teaching format					
	Live-CCD		Video-CCD		Paper-Cases	
	M	(SD)	M	(SD)	M	(SD)
Knowledge application pre-test	5.34	(1.92)	4.76	(1.90)	5.76	(2.24)
	n=30		n=27		n=33	
Knowledge application post-test	14.10	(3.32)	11.69	(3.34)	8.50	(2.44)
	n=30		n=27		n=33	
Delayed knowledge application post-test	13.36	(3.23)	11.84	(2.92)	7.89	(2.41)
	n=30		n=27		n=33	
Subjective learning outcomes	4.20	(0.63)	3.18	(1.24)	3.00	(0.99)
	n=30		n=27		n=31	

Effects of the CCD format on subjective learning outcomes

Experimental groups differed significantly with respect to subjective learning outcomes (see Table 1), $F(2,85)=13.16$, $p<0.001$, partial $\eta^2=0.236$. Participants of the Live-CCD group (M=4.20; SD=0.63) assigned better ratings to their course format than participants in the Video-CCD group (M=3.18; SD=1.24) and the Paper-Cases group (M=3.00; SD=0.99). Post hoc Bonferroni tests showed that the Live-CCD differed from the Video-CCD (p=0.001) and the Paper-Cases group ($p<0.001$) in this regard. An additional Duncan post-hoc test confirmed that the Video-CCD and the Paper-Cases group did not differ from each other in this regard (p=0.48).

To investigate the relations between the subjective assessment and the knowledge application tests applied at the end and two weeks after the course, we calculated correlations between the different outcome measures. Subjective learning outcomes correlated on a medium level with both the knowledge application post-test (r=0.343, n=88, p=0.001) and the delayed knowledge application post-test (r=0.339, n=88, p=0.001).

In the Live-CCD group, 83% of the students were in favour of implementing routine Live-CCD into the medical curriculum. Only 45% and 31% of students from the Video-CCD and Paper-Cases groups voted for an implementation of their respective course in the curriculum. With respect to the open items from the subjective learning outcomes questionnaire, participants from all groups praised the quality of the cases. Participants from the Live-CCD group particularly valued their course format for providing an opportunity to practice "diagnostic thinking" and the "focus on practice elements". They also mentioned that "you can look up theoretical knowledge, but you can't look up applied knowledge". Students in the Video-CCD group, on the other hand, praised features of the digital learning environment as they could "pause, reflect, or quickly do a Google search" when watching the case discussions. However, they also criticised it was not possible for them to "participate in a more active way".

Discussion

This randomised controlled study shows that even relatively short CCD interventions can lead to improved and sustainable learning outcomes with respect to clinical reasoning. This provides evidence that the CCD approach, which is based on Clinicopathological Conferences, is an effective teaching resource to foster clinical

reasoning skills in medical students. We had hypothesised that a more interactive course format would result in an improvement of clinical reasoning skills when compared with less interactive formats. Results show that the Live-CCD indeed leads to the highest learning outcomes in medical students compared to less interactive formats. Consistent with our hypothesis, clinical reasoning skills, as measured with our knowledge application test, had the highest gain in the Live-CCD group. These positive effects of the CCD teaching format on clinical reasoning skills proved sustainable as shown by the results in the delayed knowledge application post-test. Overall, these results are in line with a recently published study on diagnostic reasoning[31] where students who worked in pairs were more accurate in their diagnosis than individual students despite having comparable knowledge. Collaborative clinical reasoning has thus far been underrepresented in the literature, yet seems to solve many of the educational problems regarding diagnostic errors.[32]

The significant difference between the Live-CCD and the Video-CCD group can be explained by the findings of a meta-analysis that showed technology-assisted single-person learning to be inferior to group learning because of the decreased social interaction.[33] However, it is important to note that two weeks after the course, participants of the Live-CCD and Video-CCD groups did not differ significantly anymore while both groups still clearly outperformed the Paper-Case group. In other words, watching a video of the live case discussion was found to be more beneficial for learners regarding their clinical reasoning skills than just reading the printed cases. We cannot rule out that Live-CCD and Video-CCD groups did not differ in the delayed knowledge application post-test due to underpowering of the study. As our trial was not designed to detect smaller effect sizes, this finding has to be treated with caution.

Subjective learning outcomes suggest that students prefer the live discussion over the other formats. The subjective assessment correlated with the students' performance in both knowledge application post-tests. Additional qualitative data from the open item answers suggests that the Live-CCD format supported students in performing clinical reasoning and that the active discussion of cases was particularly valued by the students.

Generalisability

The conclusions of this study are applicable to a broader audience of medical students. The CCD approach and its respective formats can easily be implemented in routine medical education. Peer teaching courses hold the promise of being more easy to install and more easy to staff then courses led by faculty. Of course, live CCDs still come with certain personnel requirements, as faculty as well as a moderator need to be present. Extensive preparation was not necessary for the clinicians involved though as they served as facilitators and provided guidance only in situations when they were explicitly asked for their clinical judgement or when they felt that the discussion went astray. Total time requirements might still be lower compared to other teaching formats. Likewise, the implementation of a singular two-day training for moderators should not require extensive ressources. The study population consisting of students with heterogeneous levels of clinical experience implies that the CCD is an effective teaching format not only for students at the beginning of their clinical career but also for intermediate students. On the other hand, generalisability is potentially limited as only students from one medical school participated in our study.

Limitations of the study

There are certain limitations of this study that have to be addressed: One important limitation is the single-centre nature of this study and the relatively small sample size. Before the CCD approach can be implemented on a larger scale, a validation of our findings is therefore required. Caution is clearly warranted with the effect sizes shown in this trial, as it has been shown that effect sizes of learning intervention trials tend to be inflated compared to the effectiveness of the intervention when used in routine education.[34] Since we did not limit the time students had to work on the cases, we cannot entirely rule out that less time was spent on task in the single-learner formats and particularly the Paper-Cases group. Against this backdrop, we suggest replication to further validate the results found in this study and strengthen the outlined implications. The knowledge application test utilized in this study did not allow for a more in-depth analysis of clinical reasoning skills (i.e. a distinction of conceptual, strategic, and conditional knowledge). Larger item numbers could facilitate a reliable assessment of changes on the level of corresponding subscales. Finally, we cannot relate the underlying reasoning process with the measured knowledge gains. Further studies on clinical reasoning processes of individuals and groups are methodologically challenging but urgently needed for the advancement of a model of clinical reasoning and for improving teaching clinical reasoning.[35]

Implications for policy makers / Future research questions

Based on our findings, the CCD approach is a useful asset for medical educators to widen the range of clinical reasoning teaching tools. Live-CCD can thus be seen as a prime candidate for routine implementation in clinical reasoning curricula. Future

research should aim to identify which Live-CCD elements (the roles, case contents, or the course structure) contribute in which way to the improvement of clinical reasoning skills in medical students. The question if and to what extent such skills are applicable across domains is currently being discussed.[36] Future studies may also address the issue of transfer (i.e. to what extent can clinical reasoning skills obtained in case-based training later be applied to different cases?).[37] Regarding the Video-CCD, means of instructional support to increase the effectiveness and interactivity of the video-based format should be investigated in an attempt to exploit its full potential.

Acknowledgements

The authors thank Johanna Huber and her team for technical support with the evaluation, Thomas Brendel and Thomas Bischoff for help with the video production and Mark S. Pecker for critical reading of our manuscript and valuable suggestions. The authors also thank the CCD student discussants and moderators for their contributions. We wish to sincerely address our gratitude to the CCD team for organisational support with the study: Nora Koenemann, Simone Reichert, Sandra Petrenz, Fabian Haak, Bjoern Stolte, Simon Berhe, Bastian Brandt, and Thomas Lautz. Marc Weidenbusch wishes to express special thanks to Bernd Gansbacher for introduction to CCDs.

Funding statement

This work was supported by the German Federal Ministry of Education and Research (grant no. 01PL12016) and an intramural grant of the Medical Faculty of the University of Munich (Lehre@LMU).

Competing Interests

Marc Weidenbusch declares to have no conflict of interest.

Benedikt Lenzer declares to have no conflict of interest.

Maximilian Sailer declares to have no conflict of interest.

Christian Strobel declares to have no conflict of interest.

Raphael Kunisch declares to have no conflict of interest.

Jan Kiesewetter declares to have no conflict of interest.

Martin R. Fischer declares to have no conflict of interest.

Jan M. Zottmann declares to have no conflict of interest.

Author contributions

MW, BL, MF and JZ planned the study.

MW, BL and CS were responsible for data acquisition.

MW, BL, RK, JK, JZ, MF and MS analysed and interpreted the data.

MW, BL and JZ drafted and revised the manuscript, all authors contributed significant intellectual content and all authors gave final approval of the version to be published.

Data sharing statement

Dataset and detailed information about the CCD formats is available upon request.

References

1. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: Theory to practice. Medical Teacher. 2010;32(8):638-45.

2. Fischer MR, Bauer D, Mohn K, et al. Finally finished! National Competence Based Catalogues of Learning Objectives for Undergraduate Medical Education (NKLM) and Dental Education (NKLZ) ready for trial. GMS Journal for Medical Education. 2015;32(3):Doc35.

3. Harasym PH, Tsai TC, Hemmati P. Current trends in developing medical students' critical thinking abilities. The Kaohsiung Journal of Medical Sciences. 2008;24(7):341-55.

4. Donaldson MS, Corrigan JM, Kohn LT. To err is human: Building a safer health system. Washington: National Academies Press; 2000.

5. Kassirer JP. Teaching clinical medicine by iterative hypothesis testing: let's preach what we practice. New England Journal of Medicine. 1983;309(15):921-3.

6. Hege I, Kopp V, Adler M, Radon K, Mäsch G, Lyon H, et al. Experiences with different integration strategies of case-based e-learning. Medical Teacher. 2007;29(8):791-7.

7. Cannon WB. The Case Method of Teaching Systematic Medicine. The Boston Medical and Surgical Journal. 1900;142(2):31-6.

8. Eva KW. What every teacher needs to know about clinical reasoning. Medical Education. 2005;39(1):98-106.

9. Harris NL. Case records of the Massachusetts General Hospital – Continuing to learn from the patient. New England Journal of Medicine. 2003;348(22):2252-4.

10. Cabot RC. Case teaching in medicine: A series of graduated exercises in the differential diagnosis, prognosis and treatment of actual cases of disease. Boston: Heath; 1906.

11. Sturdy S. Knowing Cases: Biomedicine in Edinburgh, 1887-1920. Social Studies of Science. 2007;37(5):659-89.

12. Relman AS. Two Views. New England Journal of Medicine. 1979;301(20):1112-3.

13. Ertmer PA, Newby TJ. Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance Improvement Quarterly. 1993;6(4):50-72.

14. Merseth KK. Cases and case methods in teacher education. In: Sikula J, Buttery TJ, Guyton R, editors. Handbook of research on teacher education. 2nd ed. New York, NY: Macmillan; 1996. p. 722-744.

15. Zottmann JM, Stegmann K, Strijbos J-W, Vogel F, Wecker C, Fischer F. Computer-supported collaborative learning with digital video cases in teacher education: The impact of teaching experience on knowledge convergence. Computers in Human Behavior. 2013;29(5):2100-8.

16. Powers BW, Navathe AS, Jain SH. Medical education's authenticity problem. BMJ. 2014;348:g2651.

17. Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist. 2014;49(4):219-43.

18. Duncan RG, Rivet AE. Science learning progressions. Science. 2013;339(6118):396-7.

19. De Menezes S, Premnath D. Near-peer education: A novel teaching program. International journal of medical education. 2016;7:160-167.

20. Ince-Cushman D, Rudkin T, Rosenberg E. Supervised near-peer clinical teaching in the ambulatory clinic: an exploratory study of family medicine residents' perspectives. Perspectives on medical education. 2015;4(1):8-13.

21. Chi MTH. Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science. 2009;1(1):73-105.

22. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. Academic Medicine. 2006;81(3):207-12.

23. Miller DC, Salkind NJ, editors. Handbook of Research Design & Social Measurement. Thousand Oaks, CA: SAGE; 2002.

24. Kotton DN, Muse VV, Nishino M. Case 2-2012. New England Journal of Medicine. 2012;366(3):259-69.

25. Marks PW, Zukerberg LR. Case 30-2004. New England Journal of Medicine. 2004;351(13):1333-41.

26. Uyeki TM, Sharma A, Branda JA. Case 40-2009. New England Journal of Medicine. 2009;361(26):2558-69.

27. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, et al. Representation scaffolds improve diagnostic efficiency in medical students. Medical Education. 2017; 51(11):1118-1126.

28. Kopp V, Stark R, Kühne-Eversmann L, Fischer MR. Do worked examples foster medical students' diagnostic knowledge of hyperthyroidism? Medical Education. 2009;43(12):1210-7.

29. Schmidmaier R, Eiber S, Ebersbach R, Schiller M, Hege I, Holzer M, et al. Learning the facts in medical school is not enough: Which factors predict successful application of procedural knowledge in a laboratory setting? BMC Medical Education. 2013;13:28.

30. Hausknecht JP, Halbert JA, Di Paolo NT, Moriarty Gerrard MO. Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. Journal of Applied Psychology. 2007;92(2):373–385.

31. Hautz WE, Kämmer JE, Schauber SK, Spies CD, Gaissmaier W. Diagnostic performance by medical students working individually or in teams. JAMA. 2015;313(3):303-4.

32. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Medical Education. 2015;49(10):961-73.

33. Lou Y, Abrami PC, d'Apollonia S. Small group and individual learning with technology: A meta-analysis. Review of Educational Research. 2001;71(3):449-521.

34. Springer L, Stanne ME, Donovan SS. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of Educational Research. 1999;69(1):21-51.

35. Heitzmann N, Fischer MR, Fischer F. Towards more systematic and better theorised research on simulations. Medical Education. 2017;51(2):129-131.

36. Fischer F, Chinn C, Engelmann K, Osborne J, editors. Scientific reasoning and argumentation. The roles of general and specific knowledge. New York, NY: Routledge; 2018.

37. Keemink Y, Custers E, van Dijk S, ten Cate O. Illness script development in pre-clinical education through case-based clinical reasoning training. International Journal of Medical Education. 2018;9:35-41.

Figure legends

Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the *admission* part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive *discussion* part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. Abbreviations: CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

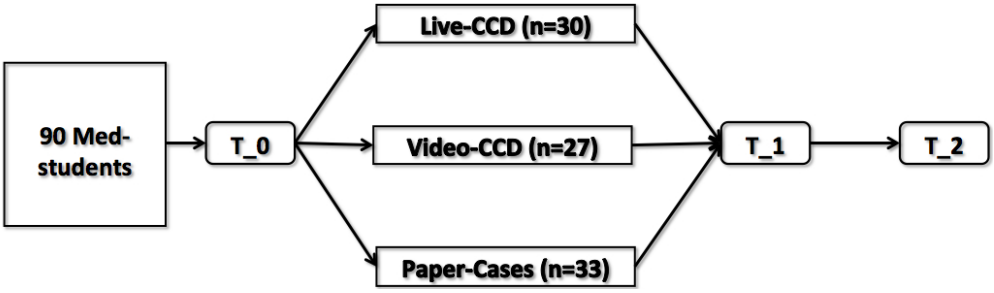


Figure 1: Study design. Full data sets of 90 medical students were analysed. T_0: knowledge application pre-test, T_1: knowledge application post-test, T_2: delayed knowledge application post-test.

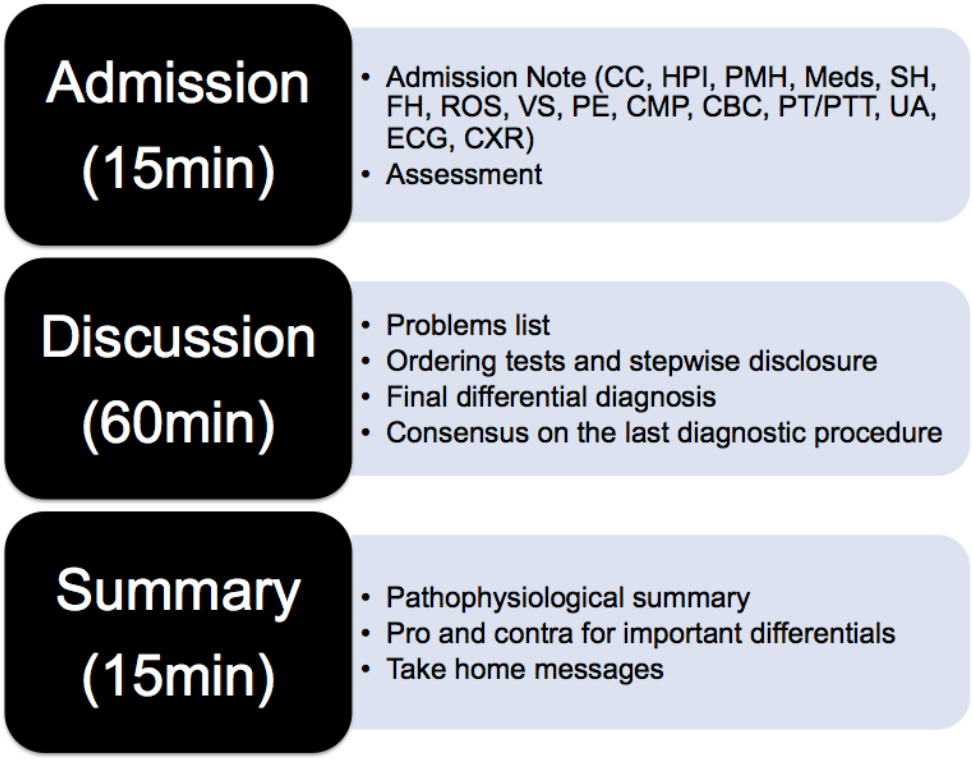


Figure 2: Live-CCD Structure. CCD sessions are divided into three parts: In the admission part the presenting student shows the discussants his prepared slides (based on the original NEJM-case record), after which the group has to agree on an assessment of the patient under discussion. In the interactive discussion part the students prioritise the medical problems, link them to possible etiologies and order tests to further corroborate or discard differential diagnoses. After all the tests that were performed in the case record, the discussants order the putative diagnostic test. The result is disclosed along with the pathological discussion and "take home messages" on important differentials in the third part of the session. Abbreviations: CC chief complaint, HPI history of present illness, PMH past medical history, Meds medications, SH social history, FH family history, ROS review of systems, VS vital signs, PE physical examination, CMP comprehensive metabolic panel, CBC complete blood count, PT prothrombin time, PTT partial thromboplastin time, UA urine analysis, ECG electrocardiogram, CXR chest radiograph.

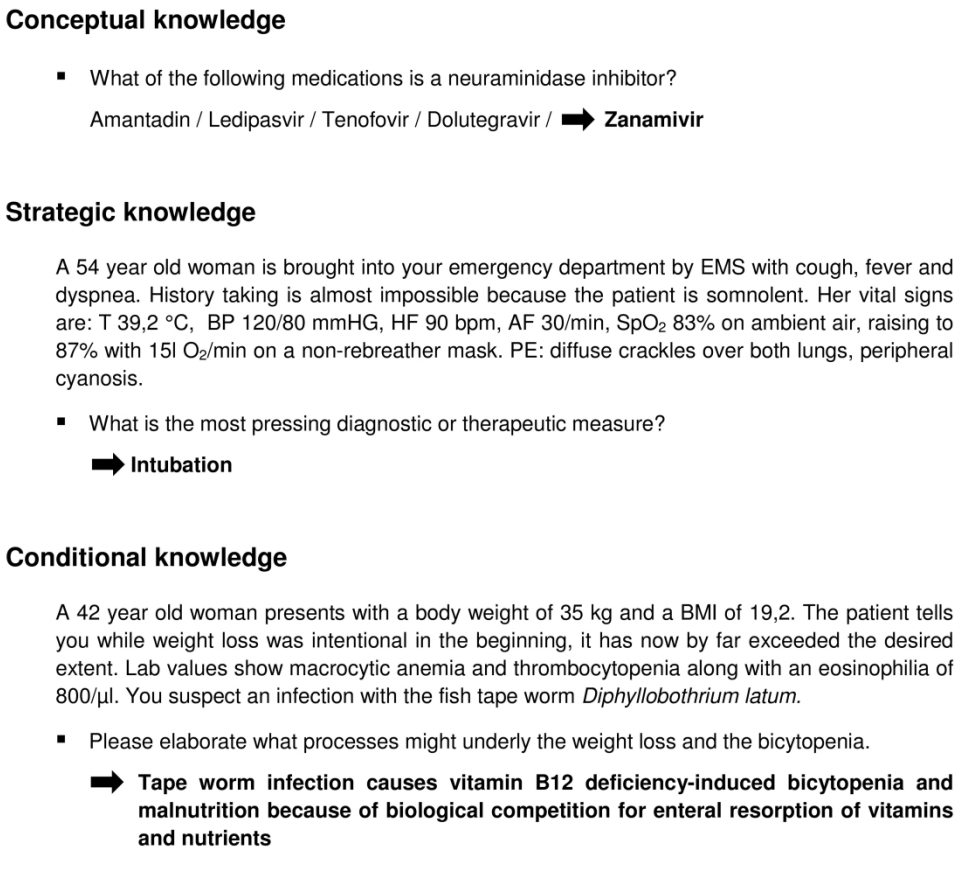


Figure 3: Knowledge application test. Exemplary items are shown for each of the knowledge types addressed (arrows point to the correct answers). The test included 11 items on conceptual knowledge, 9 items on strategic knowledge, and 9 items on conditional knowledge.

150x135mm (600 x 600 DPI)

Supplemental

Baseline data: Demographic characteristics of the study participants.

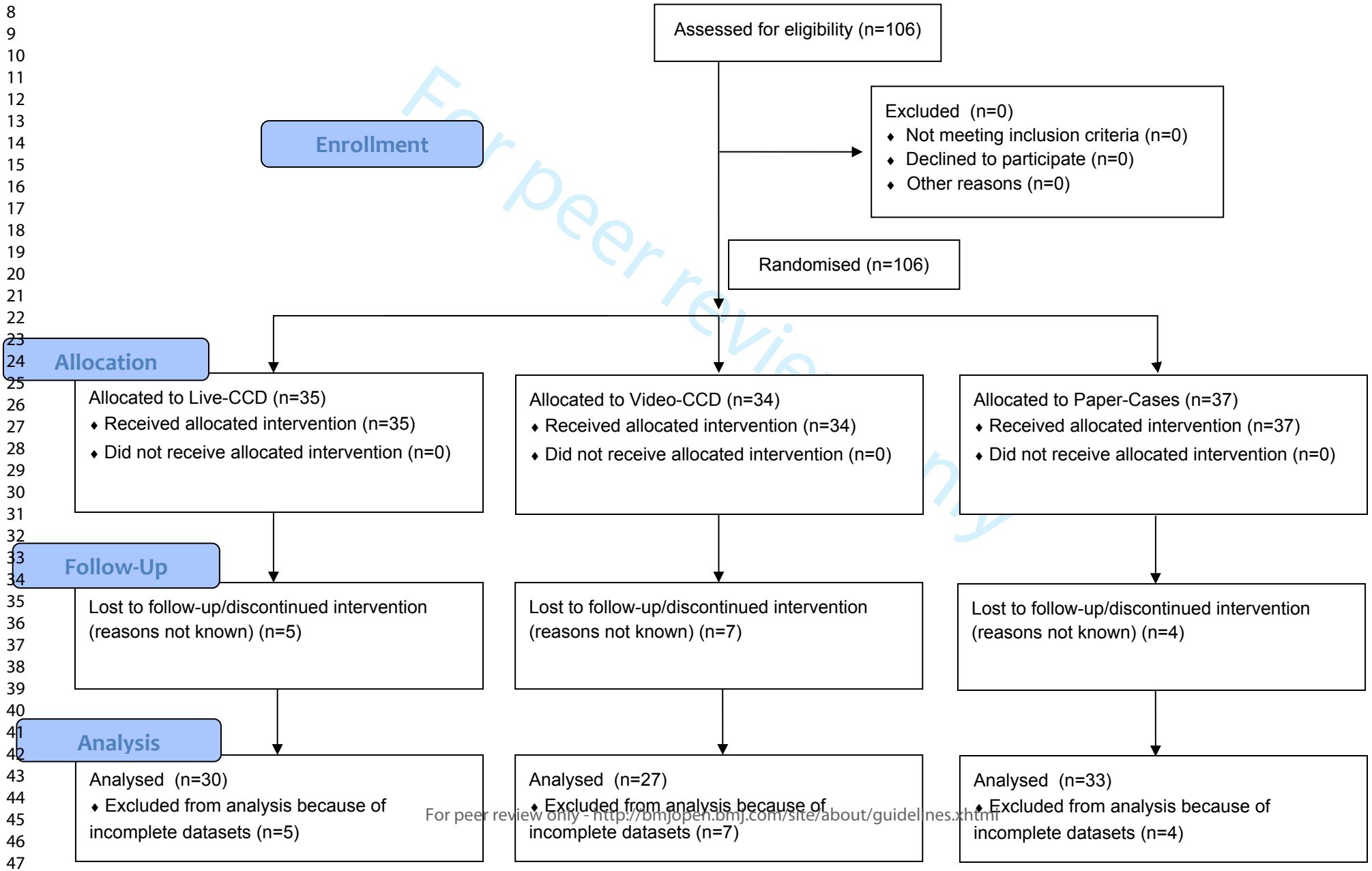
	CCD format							
	Live-CCD *n* = 30		Video-CCD *n* = 27		Paper-Cases *n* = 33		All formats *N* = 90	
Gender distribution *N* m/f (% f)	14/16 (53.3% f)		5/22 (81.5% f)		12/21 (63.6% f)		31/59 (65.6% f)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Age	23.77	(4.09)	22.26	(1.77)	22.91	(2.40)	23.0	(2.97)
Clinical semester	3.23	(1.96)	3.41	(1.47)	3.82	(1.84)	3.50	(1.78)
High school grade	1.53	(0.36)	1.35	(0.42)	1.48	(0.68)	1.46	(0.52)
First National Board Exam Score	245	(30)	226	(78)	246	(30)	240	(49)
Participants with prior CCD experience n (%)	6	(20.0%)	6	(22.2%)	5	(15.1%)	17	(18.9%)

Questionnaire items (5-point Likert scale) for the assessment of subjective learning outcomes.

1. I perceived this CCD format as meaningful.
2. I learned a lot during the CCD course.
3. The CCD course increased my learning motivation.
4. I would like to participate in this CCD format again in the future.
5. I enjoyed the CCD course.
6. This CCD format should be offered as part of the curriculum.
7. I was able to follow the case discussions.
8. Learning in the CCD format is easier for me than learning in traditional lectures or seminars.
9. How would you rate the course overall?

CONSORT
TRANSPARENT REPORTING of TRIALS

CONSORT 2010 Flow Diagram



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2-3
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	4-7
	2b	Specific objectives or hypotheses	6-7
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	8
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	not applicable
Participants	4a	Eligibility criteria for participants	7
	4b	Settings and locations where the data were collected	7
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	8-10
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	8, 11-12
	6b	Any changes to trial outcomes after the trial commenced, with reasons	not applicable
Sample size	7a	How sample size was determined	12
	7b	When applicable, explanation of any interim analyses and stopping guidelines	not applicable
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	7
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	7
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	not applicable
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	not applicable
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those	not applicable

		assessing outcomes) and how	
	11b	If relevant, description of the similarity of interventions	not applicable
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	12
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	12
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	attached as supplemental
	13b	For each group, losses and exclusions after randomisation, together with reasons	7
Recruitment	14a	Dates defining the periods of recruitment and follow-up	not applicable
	14b	Why the trial ended or was stopped	not applicable
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	attached as supplemental
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	13
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	12-14
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	not applicable
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	not applicable
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	not applicable
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	17
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	16
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	14-16
Other information			
Registration	23	Registration number and name of trial registry	not applicable
Protocol	24	Where the full trial protocol can be accessed, if available	19
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	18

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.