PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and staff assessments using data from the cluster-randomized controlled IMPROVEjob study
AUTHORS	Schmidt, Manuela; Seifried-Dübon, Tanja; Göbel, Julian; Degen, Lukas; Werners, Brigitte; Grot, Matthias; Rind, Esther; Pieper, Claudia; Jöckel, Karl-Heinz; Linden, Karen; Rieger, Monika A.; Weltermann, Birgitta

VERSION 1 – REVIEW

REVIEWER	Khalique U Zaman Bahra University, Management Studies
REVIEW RETURNED	19-Sep-2022

GENERAL COMMENTS	Research gap to be defined clearly and in detail. Background needs to improved minorly. Define LMX theory in detail (mention the in group and out group) Limitation to be defined in better way. Future recommendation to be included.

REVIEWER	Ruwayda Petrus University of KwaZulu-Natal College of Humanities, School of Applied Human Sciences
REVIEW RETURNED	18-Oct-2022

GENERAL COMMENTS	Please see comments made in the PDF version of the article. Please also clarify that while the trial was registered in October; recruitment commenced in September. I note that you received ethics in Feb however for a trial to officially start it would need to be registered and only then do you approach The reviewer provided a marked copy with additional comments. Please contact the publisher for full details.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

- 1.1: Research gap to be defined clearly and in detail.

  Response: Thank you for your comment. The research gap is now more clearly described in the fully revised background.

  Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al. showed an association between stronger leadership and less burnout among different medical professionals in the United States [4]. In contrast, poor

leadership skills have a negative impact on job satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-day-mortality [11]. In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12].

Scientifically, leadership is conceptualized in several theories. One of the most studied leadership frameworks is the Full Range of Leadership Model (FRLM), which integrates transactional, transformational, and negative leadership [13, 14]. Transactional leadership describes leaders' structuring of work situations, the exchange of contingent rewards (e.g., work against salary), and the management by exception [13–15]. In contrast, transformational leadership moves beyond leaders' and staff's self-interests. It focuses on the staff's attitudes and values regarding overarching goals such as self-actualization, organizational achievements, and the well-being of others and society as a whole [13, 14]. Building on the FRLM, a recent further development, the so-called Implementation Leadership Scale, focusses on the role of leadership for implementation of organizational changes [12]. Another important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the relationship between leaders and staff. It concentrates on the perceived quality of the dyadic relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a dyadic social exchange process ranging from low LMX, described by limited social transactions with more transactional leadership to high LMX, which represents a transformational approach with a high degree of social exchange and a mature leader-member partnership [17]. High-quality relationships positively influence employees' work-related well-being and are associated with higher job satisfaction of health care workers [18].

Based on these theories, various questionnaires were developed, e.g., the Leader-Member Exchange questionnaire 'LMX-7' [16] and the German questionnaire 'Fragebogen für integrative Führung' (FIF; in English: Questionnaire for Integrative Leadership) [15, 19]. These instruments allow for a multi-rater perspective: the leader's and the staff's views on the leader's behavior are measured and compared providing 180° feedback. This method is valuable because assessments from different perspectives create a more comprehensive picture of the leaders' actual skills and performances [20]. Two recent reviews of 60 studies from various medical settings showed that such approaches are increasingly applied in medical education and graduate training [21, 22], but have not been used to evaluate GP leaders and their teams. Effective interventions to improve leadership were developed and evaluated in the hospital [23] and healthcare management setting [24]. For example, Saravo et al. showed an improvement in transformational and transactional leadership performance of 57 medical residents in hospital rotations after a 4-week intervention [23]. In addition, a 2018 study from Hill et al. highlighted positive effects of a leadership training for surgical residents on teamwork and team involvement in decision-making [25]. However, such interventions have not been implemented in German primary care, although high chronic stress and burnout rates are reported for this workforce [26, 27]. The need is even larger as about half of the German GPs who mainly work in GP-owned private practices [28], will reach retirement age in the next ten years [29]. Based on the leadership frameworks mentioned above, the publicly funded IMPROVEjob study aimed to improve the job satisfaction of physician leaders and practice personnel of German GP practices focusing on leadership, communication, and work processes [30, 31]. At baseline, GPs´ leadership skills were evaluated comparing GP leaders' self and staff ratings on practice level.

- 1.2: Background needs to improved minorly.

  Response: Thank you. We have now thoroughly revised the background.

  Revised text: Please see the previous comment (1.1)

- 1.3: Define LMX theory in detail (mention the in group and out group)

  Response: Thank you for your advice. We detailed LMX theory including the importance of relationship quality and how it fits into a transactional/transformational construct. Based on the publication by Graen & Biel (1995, *Relationship-based approach to leadership: Development*

of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective) the in-group/out-group perspective has been replaced by a more prescriptive and practically useful approach regarding social transaction.

Revised text: Another important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the relationship between leaders and staff. It concentrates on the perceived quality of the dyadic relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a dyadic social exchange process ranging from low LMX, described by limited social transactions with more and associated with transactional leadership to high LMX, which represents a transformational approach with a high degree of social exchange and a mature leader-member partnership [17]. High-quality relationships positively influence employees' work-related well-being and are associated with higher job satisfaction of health care workers [18].

- 1.4: Limitation to be defined in better way.

  Response: Thank you for the helpful comment. The limitations section was revised.

  Revised text: Novel for the German GP setting, we investigated GP leadership in a large sample with analysis on practice level. Our data provide leadership ratings for each solo practice leaders, but not for each group practice leader, as we had asked staff to rate their leadership team to reflect current small team leadership situations. LMX data were missing for one of seven questions for the small number of employed physicians. However, the analysis of the available data yielded a high relationship quality with leaders like the results for practice assistants. A selection bias cannot be excluded as participating practices might have had a greater interest in the topic.

- 1.5: Future recommendation to be included.
  Response: Thank you, we added future recommendations.

  Revised text: Overall, our data from the IMPROVEjob study show trustful relationships between GP leaders and their staff. Future GPs´ trainings should enable GP leaders to implement goal-setting, innovation, and individuality focus more effectively. Our results support recent calls for leadership workshops on every level of the medical training for strengthening the GP and other health services workforce.

Reviewer 2:

- 2.1 :It is not clear from the title that what will be reported on Leaders Skills - perhaps rephrase to align more with the objective of the study (title)
  Response: Thank you for pointing this out. The title was changed.
  Revised text: **A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and staff assessments using data from the IMPROVE*job* study**

- 2.2: I recommend reading Söling, S., Pfaff, H., Karbach, U., Ansmann, L., & Köberlein-Neu, J. (2022). How is leadership behavior associated with organization-related variables? Translation and psychometric evaluation of the implementation leadership scale in German primary healthcare. BMC Health Services Research, 22(1), 1-13. (title)
  Response: Thank you for the literature recommendation. We included the aspect of implementation leadership in our introduction.

Revised text: In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12].
Building on the FRLM, a recent further development, the so-called Implementation Leadership Scale, focusses on the role of leadership for implementation of organizational changes [12].

- 2.3: This objective was not clear from the title of the study (Abstract, Objectives, line 23)
  Response: The title was revised.
  Revised text: **A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and staff assessments using data from the IMPROVE*job* study**

- 2.4: Also why are you doing this? What is the problem? (Abstract, Objectives, line 23)
  Response: Thank you, we detailed this more specifically.
  Revised text: Strong primary care leaders are needed to assure high quality services for patient populations. This study analyzed general practitioners' (GP) leadership skills comparing practice-level self- and staff assessments based on the Full Range of Leadership Model (FRLM) and the Leader-Member Exchange (LMX).

- 2.5: This refers to the context and not the analysis - describe the sampling setting (example: The study was conducted among Clinical Practices based in Germany (Abstract, Setting, line 25)
  Response: Thank you for pointing this out. We revised the section.
  Revised text: This questionnaire survey was conducted among German general practice leaders and their staff participating in the IMPROVE*job* trial.

- 2.6: Be specific 60 Medical Practices (Abstract, Participants, line 27)
  Response: Thank you, this was specified.
  Revised text: The study population comprised 60 German general practices with 366 participants: 84 GP practice leaders and 282 employees (28 physicians and 254 practice assistants).

- 2.7: This is confusing - suggest you rephrase.
  Example: The participants for this study were recruited from 60 Medical Practices based in Germany. In total n = 366 persons took part in the study. Of the total sample; n = 84 were General Practitioners who served in a leadership capacity; with the remaining n = 282 medical staff (n = 28 GP and n = 254 practice assistants) that work in those respective practices. (Abstract, Participants, line 27+28)
  Response: Thank you, we revised this section.
  Revised text**:** The study population comprised 60 German general practices with 366 participants: 84 GP practice leaders and 282 employees (28 physicians and 254 practice assistants).

- 2.8: Leadership Skills? Leadership practice? Leadership style? be specific (Abstract, Primary and secondary outcome measures, line 29)
  Response: Thank you, we clarified this.
  Revised text: Leadership skills of the practice leaders were measured using the Integrative Leadership Questionnaire (German FIF) and the Leader-Member Exchange (LMX-7) questionnaire. Leaders rated themselves and practice staff rated their leaders. The data was analyzed by paired mean comparisons on the practice level.

- 2.9: Also, whose leadership was measured? The practice owners? Or of everyone? (Abstract, Primary and secondary outcome measures, line 29)
  Response: This was clarified.

Revised text: Leadership skills of the practice leaders were measured using the Integrative Leadership Questionnaire (German FIF) and the Leader-Member Exchange (LMX-7) questionnaire. Leaders rated themselves and practice staff rated their leaders. The data was analyzed by paired mean comparisons on the practice level.

- 2.10: Confusing presentation of this information.
  Rephrase: Data was analyzed by..... to provide a holistic view of the perceived leadership skills of the ..... (Abstract, Primary and secondary outcome measures, lines 31-33)
  Response: Thank you for your suggestion. The information was rephrased.
  Revised text: Leadership skills of the practice leaders were measured using the Integrative Leadership Questionnaire (German FIF) and the Leader-Member Exchange (LMX-7) questionnaire. Leaders rated themselves and practice staff rated their leaders. The data was analyzed by paired mean comparisons on the practice level.

- 2.11: This section needs to be rephrased so that it is clear what the study found. Currently it does not read clearly whether only self rated scores and those of employees were compared or whether within group analysis was also conducted (Abstract, Results, lines 34-41)
  Response: Thank you for the comment. We clarified this.
  Revised text: For most leadership dimensions, practice leaders rated themselves higher than their employees rated them. Differences were found for transformational leadership ($p<.001$, $d=.41$), especially for the dimensions 'innovation' ($p<.001$, $d=.69$) and 'individuality focus' ($p<.001$, $d=.50$). For transactional leadership, the dimension 'goal setting' differed significantly ($p<.01$, $d=.30$) but not the other dimensions. Scores for negative leadership were low and showed no differences between leaders and employees. Interestingly, employed physicians' rated their practice leaders higher on the two transformational ('performance development', 'providing a vision') and all transactional dimensions. The LMX-7 scale showed high quality relationships between leaders and employees.

- 2.12: Of the practice leaders (Abstract, results, line 35)
  Response: Thank you. The text was revised.
  Revised text: Interestingly, employed physicians' rated their practice leaders higher on the two transformational ('performance development', 'providing a vision') and all transactional dimensions.

- 2.13: Did you analyze the results from the owners; the GP and the Practice assistances separately? Did you group the employees together and compare their rating to that of the leadership self rating? (Abstract, results, line 41)
  Response: Thank you for pointing this out, this was corrected.
  Revised text: For most leadership dimensions, practice leaders rated themselves higher than their employees rated them. Differences were found for transformational leadership ($p<.001$, $d=.41$), especially for the dimensions 'innovation' ($p<.001$, $d=.69$) and 'individuality focus' ($p<.001$, $d=.50$). For transactional leadership, the dimension 'goal setting' differed significantly ($p<.01$, $d=.30$) but not the other dimensions. Scores for negative leadership were low and showed no differences between leaders and employees. Interestingly, employed physicians' rated their practice leaders higher on the two transformational ('performance development', 'providing a vision') and all transactional dimensions. The LMX-7 scale showed high quality relationships between leaders and employees.

- 2.14: Given that the FIF assesses transformational, transactional, instrumental and abusive leadership as well as leaders- communication - group your results accordingly (Abstract, Conclusions, line 44)
  Response: Thank you for this point. We have revised the results and conclusion section.

Revised text:
Results: For most leadership dimensions, practice leaders rated themselves higher than their employees rated them. Differences were found for transformational leadership ($p<.001$, $d=.41$), especially for the dimensions 'innovation' ($p<.001$, $d=.69$) and 'individuality focus' ($p<.001$, $d=.50$). For transactional leadership, the dimension 'goal setting' differed significantly ($p<.01$, $d=.30$) but not the other dimensions. Scores for negative leadership were low and showed no differences between leaders and employees. Interestingly, employed physicians' rated their practice leaders higher on the two transformational ('performance development', 'providing a vision') and all transactional dimensions. The LMX-7 scale showed high quality relationships between leaders and employees.
Conclusions: This 180° analysis of GPs´ leadership skills with self- and employee ratings indicated good relationships. There is a potential to improve leadership regarding goal-setting, innovation and focusing on individual team members. These results allow for the development of targeted interventions.

- 2.15: Also what does this mean in the context of the problem? (Abstract, Conclusions, line 44)
  Response: Thank you, we revised the conclusion.
  Revised text: This 180° analysis of GPs´ leadership skills with self- and employee ratings indicated good relationships. There is a potential to improve leadership regarding goal-setting, innovation and focusing on individual team members. These results allow for the development of targeted interventions.

- 2.16: This is an example of a background

  Primary health care (PHC) in Ethiopia serves as the main entry point for preventive, promotive and curative health services. The district health office is responsible for the planning, implementation and evaluation of all district health activities. In addition, district health offices manage service delivery facilities working on provision of PHC – primary hospitals, health centers and health posts. As the leader of the health care system tier, district health management must ensure direction, alignment and commitment within teams and organizations and make sure that achievements are consistent with the vision, values and strategy of the organization. USAID Transform: Primary Health Care provides diverse support to improve district health manager competencies including in-service trainings followed by planning and implementation of performance improvement projects and on-the-job mentoring and support.

  Your paper is missing this crucial element (Background, line 60)
  Response: Thank you for the example. We revised the background substantially.

  Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al. showed an association between stronger leadership and less burnout among different medical professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-day-mortality [11]. In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12].

  Scientifically, leadership is conceptualized in several theories. One of the most studied leadership frameworks is the Full Range of Leadership Model (FRLM), which integrates transactional, transformational, and negative leadership [13, 14]. Transactional leadership describes leaders' structuring of work situations, the exchange of contingent rewards (e.g., work against salary), and the management by exception [13–15]. In contrast, transformational leadership moves beyond leaders' and staff's self-interests. It focuses on the staff's attitudes

and values regarding overarching goals such as self-actualization, organizational achievements, and the well-being of others and society as a whole [13, 14]. Building on the FRLM, a recent further development, the so-called Implementation Leadership Scale, focusses on the role of leadership for implementation of organizational changes [12]. Another important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the relationship between leaders and staff. It concentrates on the perceived quality of the dyadic relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a dyadic social exchange process ranging from low LMX, described by limited social transactions with more transactional leadership to high LMX, which represents a transformational approach with a high degree of social exchange and a mature leader-member partnership [17]. High-quality relationships positively influence employees' work-related well-being and are associated with higher job satisfaction of health care workers [18].

Based on these theories, various questionnaires were developed, e.g., the Leader-Member Exchange questionnaire 'LMX-7' [16] and the German questionnaire 'Fragebogen für integrative Führung' (FIF; in English: Questionnaire for Integrative Leadership) [15, 19]. These instruments allow for a multi-rater perspective: the leader's and the staff's views on the leader's behavior are measured and compared providing 180° feedback. This method is valuable because assessments from different perspectives create a more comprehensive picture of the leaders' actual skills and performances [20]. Two recent reviews of 60 studies from various medical settings showed that such approaches are increasingly applied in medical education and graduate training [21, 22], but have not been used to evaluate GP leaders and their teams. Effective interventions to improve leadership were developed and evaluated in the hospital [23] and healthcare management setting [24]. For example, Saravo et al. showed an improvement in transformational and transactional leadership performance of 57 medical residents in hospital rotations after a 4-week intervention [23]. In addition, a 2018 study from Hill et al. highlighted positive effects of a leadership training for surgical residents on teamwork and team involvement in decision-making [25]. However, such interventions have not been implemented in German primary care, although high chronic stress and burnout rates are reported for this workforce [26, 27]. The need is even larger as about half of the German GPs who mainly work in GP-owned private practices [28], will reach retirement age in the next ten years [29]. Based on the leadership frameworks mentioned above, the publicly funded IMPROVEjob study aimed to improve the job satisfaction of physician leaders and practice personnel of German GP practices focusing on leadership, communication, and work processes [30, 31]. At baseline, GPs´ leadership skills were evaluated comparing GP leaders' self and staff ratings on practice level.

- 2.17: Incomplete sentence/ confusing sentence
  rephrase:
  Leadership has become an important topic in the medical field as poor leadership practices has been shown to have a negative impact on job satisfaction of staff; patient centered care and staff wellbeing. (Background, line 61+62)
  Response: Thank you for the comment. We have adjusted the paragraph.

  Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al. showed an association between stronger leadership and less burnout among different medical professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-day-mortality [11]. In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12]

- 2.18: Why is leadership important? You have not stated this upfront so it appears that your study is not anchored by a research problem. (Background, line 61+62)

Response: We have defined the research gap in more detail and highlighted why leadership is relevant in the background.

Revised text: Please see revised text in comment 2.17.

- 2.19: GENERAL COMMENT:
  Please have a language editor work with you for flow and understanding (Background, line 66)
  Response: The manuscript was checked by Sarah Chalmers who is a certified medical translator (www.medi-translate.com)

- 2.20: Leadership has become an important focus in the medical field
  Issues/challenges in leadership practices have been proven to result in poor job satisfaction, poor staff wellbeing and poor quality of patient care. (Background, line 66)
  Response: Thank you for your suggestion. We adjusted the paragraph.
  Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al. showed an association between stronger leadership and less burnout among different medical professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-day-mortality [11]. In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12].

- 2.21: Please foreground what does healthcare look like in your setting? Primary care means something different based on HMIC and LMIC. (Background, line 67 + 68)
  Response: Thanks for the important comment. We specified this.
  Revised text: In Germany, primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician, practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years. Similar to other regions worldwide, the size of group practices is increasing.

- 2.22: The effort is the doing so how can this be success? (Background, line 67)
  Response: Thank you, this was revised.
  Revised text: Effective interventions to improve leadership were developed and evaluated in the hospital [23] and healthcare management setting [24].

- 2.23: How is stress and leadership related? You need to set the scene for your reader/
  You need to do a PPS+Q
  Problem Statement - what is the problem your study is addressing
  Purpose Statement; why must this problem be addressed
  Research Question - the actual question that you will be answering or researching (Background, line 69+70)
  Response: Thank you for the comment. We clarified the problem our study is addressing and the resulting research question in the revised background.
  Revised text: Please see revised text in comment 2.16.

- 2.24: This is a poorly placed sentence that does not flow or provide further support for what you have said above (Background, lines 70-72)

Response: Thank you, the background section was fully revised.
Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al. showed an association between stronger leadership and less burnout among different medical professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-day-mortality [11]. In addition to individual outcomes, leadership is important to promote organizational changes (e.g., the implementation of IT-supported care) [12].

- 2.25: Are these non medical personnel? Or are they nurses? Please clarify so that the reader may understand (Background, line 70)
  Response: Thank you. Practice assistants in German GP practices are professionally trained, non-physician personnel, performing medical and administrative tasks to assist the GPs.
  Revised text: In Germany, primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician, practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years. Similar to other regions worldwide, the size of group practices is increasing.

- 2.26: Be careful of stating there are no studies

  See:
  -Swanwick, T., & Varnam, R. (2019). Leadership development and primary care. BMJ Leader, 3(2).
  -Gilson, L., Elloker, S., Olckers, P., & Lehmann, U. (2014). Advancing the application of systems thinking in health: South African examples of a leadership of sensemaking for primary health care. Health Research Policy and Systems, 12(1), 1-13.
  -Cleary, S., Toit, A. D., Scott, V., & Gilson, L. (2018). Enabling relational leadership in primary healthcare settings: lessons from the DIALHS collaboration. Health Policy and Planning, 33(suppl_2), ii65-ii74.
  -i Solà, G. J., i Badia, J. G., Hito, P. D., Osaba, M. A. C., & García, J. L. D. V. (2016). Self-perception of leadership styles and behaviour in primary health care. BMC health services research, 16.
  -Sfantou, D. F., Laliotis, A., Patelarou, A. E., Sifaki-Pistolla, D., Matalliotakis, M., & Patelarou, E. (2017, October). Importance of leadership style towards quality of care measures in healthcare settings: a systematic review. In Healthcare (Vol. 5, No. 4, p. 73). MDPI.
  -Alrubaysh, M. A., Alshehri, M. H., Alsuhaibani, E. A., Allowaihiq, L. H.,
  -Alnasser, A. A., & Alwazzan, L. (2022). The leadership styles of primary healthcare center managers and center performance outcomes in Riyadh, Saudi Arabia: A correlational study. Journal of Family & Community Medicine, 29(1), 56.
  (Background, lines 72-74)
  Response: Thank you very much for the references, which were very helpful. We included these in the background and/or discussion of the manuscript.

- 2.27: You jump from one idea to the next - no linking sentences, no connection between paragraphs. Signpost for your reader (Background, line 75+76)
  Response: We fundamentally revised the background.
  Revised text: Please see the revised text of comment 2.16.

- 2.28: This is not suited here. You are not providing background or integrated presentation of what your study is focusing on. You are defining tools and frameworks. (Background, lines 77-94)
  Response: Thank you for pointing this out. The respective sections were focused and rephrased.
  Revised text: Scientifically, leadership is conceptualized in several theories. One of the most studied leadership frameworks is the Full Range of Leadership Model (FRLM), which integrates transactional, transformational, and negative leadership [13, 14]. Transactional leadership describes leaders' structuring of work situations, the exchange of contingent rewards (e.g., work against salary), and the management by exception [13–15]. In contrast, transformational leadership moves beyond leaders' and staff's self-interests. It focuses on the staff's attitudes and values regarding overarching goals such as self-actualization, organizational achievements, and the well-being of others and society as a whole [13, 14]. Building on the FRLM, a recent further development, the so-called Implementation Leadership Scale, focusses on the role of leadership for implementation of organizational changes [12]. Another important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the relationship between leaders and staff. It concentrates on the perceived quality of the dyadic relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a dyadic social exchange process ranging from low LMX, described by limited social transactions with more transactional leadership to high LMX, which represents a transformational approach with a high degree of social exchange and a mature leader-member partnership [17]. High-quality relationships positively influence employees' work-related well-being and are associated with higher job satisfaction of health care workers [18].

- 2.29: You have not orientated the reader to the healthcare setting or the type of practice you included - do they only deal with chronic patients? Is this a child clinic? (Background, line 77-94)
  Response: We included this information in the background and methods section.

  Revised text: The need is even larger as about half of the German GPs who mainly work in GP-owned private practices [28], will reach retirement age in the next ten years [29].

  In Germany, primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician, practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years. Similar to other regions worldwide, the size of group practices is increasing.

- 2.30: Because you have not presented a problem your study is addressing - all of this information does not make sense (Background, lines 95-99)
  Response: Thank you, the background was revised.
  Revised text: Please see revised text of comment 2.16.

- 2.31: Is it 60 or 56? (Methods, line 128)
  Response: The sample size calculation resulted in 56 practices for the effectiveness study. On purpose, we included 4 more practices in case of drop-outs. The recruitment process is detailed in our baseline paper (Degen et al., 2021; Int. Journal of Environmental Research and Public Health)
  Revised text: In short, a total of 60 GP practices in the North Rhine region in Germany were recruited by the Institute of General Practice and Family Medicine of the University of Bonn […]

- 2.32: What is the difference between these two? what do you mean? (Methods, line 130)

Response: Historically, German GP practices are one-physician practices. During the last 2 decades, the number of practices with more than one provider has increased markedly. We detailed this in the revised text.
Revised text: In Germany, primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician, practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years. Similar to other regions worldwide, the size of group practices is increasing.

- 2.33: Explain this to non German residents who might no know how your healthcare system is structured (Methods, line 130)
  Response: Thank you, we described this in more detail.
  Revised text: Strong primary care leaders and a strong primary care workforce are important to assure the health of populations and primary care teams [1–3].

  In Germany, primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician, practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years. Similar to other regions worldwide, the size of group practices is increasing.

- 2.34: This does not go here. Under participants you describe your sample. (Methods, lines 132-134)
  Response: Thank you for pointing this out. This was revised.

  Revised text:

  Practice leaders answered a short questionnaire on practice characteristics and the questionnaire for practice leaders. Employed physicians and practice assistants filled out different versions of the same employee questionnaire.

- 2.35: Please be consistent in how you present your scales. (Methods, line 168)
  Response: Thank you. We adjusted this.

  Revised text: The relationship quality between leaders and staff is measured using the Leader-Member Exchange questionnaire (LMX-7) with seven items on a 5-point Likert scale, which are worded to reflect the leader or the staff position [16, 34, 35].

- 2.36: Why not just use SPSS for all analysis? (Methods, lines 179-181)
  Response: Thank you. One of the researchers preferred to work with SAS. We controlled the reported values using SPSS and changed the text.

  Revised text: Statistical analyses were conducted using SPSS Statistics 27 (IBM Cooperation, Armonk, Ny, USA, 2020).

- 2.37: It may be worthwhile to send reviewers the tool and how the manual says to interpret the scores (Methods, line 183)
  Response: Thank you for this suggestion. Since the manual is protected by copyright, sharing of the original manuals is not possible. Sorry for this.

- 2.38: This is problematic because if the personnel rated the entire group based on their like or dislike of one leader - would this not skew the results? (Methods, lines 204-208)

Response: To study the leadership teams of primary care practices, we asked staff to rate their leaders as a team. We were more interested in practice leadership in total, not the results for individual leaders.
Revised text: In practices with more than one owner (group practices), each leader's self-rating was compared with the respective ratings of the practice personnel, who were asked to rate the leadership team of the practice, not stratified by individual leaders. This approach was chosen because practice owners of German practices typically work as a leadership team.

Our data provide leadership ratings for each solo practice leaders, but not for each group practice leader, as we had asked staff to rate their leadership team to reflect current small team leadership situations.

- 2.39: What is the justification for not asking them to rate the leaders individually? (Methods, lines 204-208)
  Response: Please see comment 2.38.

- 2.40: Is it 60 or 56 practices? (Results, line 216)
- Response: The sample size calculation resulted in 56 practices for the effectiveness study. On purpose, we included 4 more practices in case of drop-outs. The recruitment process is detailed in our baseline paper (Degen et al., 2021; Int. Journal of Environmental Research and Public Health)
- Revised text: In short, a total of 60 GP practices in the North Rhine region in Germany were recruited by the Institute of General Practice and Family Medicine of the University of Bonn […]

- 2.41: What does this have to do with leadership? Only present what is relevant to the objective of the study (Results, Table 1)
  Response: Thank you for pointing this out. The variables were excluded.
  Revised text: Please see table 1.

- 2.42: In both the protocol paper and the beginning of this paper - you stated 56 practices. Again you reference 60 practices???? Please be consistent (Results, lines 230-232)
- Response: The sample size calculation resulted in 56 practices for the effectiveness study. On purpose, we included 4 more practices in case of drop-outs. The recruitment process is detailed in our baseline paper (Degen et al., 2021; Int. Journal of Environmental Research and Public Health)
- Revised text: In short, a total of 60 GP practices in the North Rhine region in Germany were recruited by the Institute of General Practice and Family Medicine of the University of Bonn […]

- 2.43: Based on the extensive corrections that need to be done prior to the discussion section, I cannot comment on the discussion as the problem was not clearly identified; the background did not contextualize the need for the study and situate it within prior research or literature. Therefore it would not make sense to review the discussion as I have nothing to relate it to. (Discussion, line 280)
  Response: We understand your criticism and revised the manuscript profoundly.

- 2.44: But this is not what you set out to do? Or this was not made explicit (Discussion, line 282)
  Response: We revised the background and discussion sections.

- 2.45: General comment: Please also clarify that while the trial was registered in October; recruitment commenced in September. I note that you received ethics in Feb however for a trial to officially start it would need to be registered and only then do you approach
  Response: The ethics vote was based on the full study protocol which was approved in February 2019. Registration was prepared prior to starting recruitment. Due to a misunderstanding of personnel involved, the final registration was performed after recruitment started.