

# BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email [info.bmjopen@bmj.com](mailto:info.bmjopen@bmj.com)

# BMJ Open

## A 180° view on general practitioners' leadership: comparison of leader and practice team assessments using data from the IMPROVEjob study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2022-066298
Article Type:	Original research
Date Submitted by the Author:	07-Jul-2022
Complete List of Authors:	Schmidt, Manuela; University Hospital Bonn, Institute of General Practice and Family Medicine Seifried-Dübon, Tanja; University Hospital Tuebingen Osianderstraße 5, Department of Psychosomatic Medicine and Psychotherapy Göbel, Julian; University Hospital Bonn, Institute of General Practice and Family Medicine Degen, Lukas; University Hospital Bonn, Institute of General Practice and Family Medicine Werners, Brigitte; Ruhr University Bochum Grot, Matthias; Ruhr University Bochum Rind, Esther; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Pieper, Claudia; University of Duisburg-Essen Jöckel, Karl-Heinz; Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg- Essen Linden, Karen; University Hospital Bonn, Institute of General Practice and Family Medicine Rieger, Monika A.; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Weltermann, Birgitta; University Hospital Bonn, Institute of General Practice and Family Medicine
Keywords:	PRIMARY CARE, Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 1 **A 180° view on general practitioners' leadership: comparison of leader and practice team**  
4 **assessments using data from the IMPROVEjob study**  
5  
6  
7 3

8 4 **Authors:** Manuela Schmidt<sup>1\*</sup>, Tanja Seifried-Dübon<sup>2</sup>, Julian Göbel<sup>1</sup>, Lukas Degen<sup>1</sup>, Brigitte Werners<sup>3</sup>,  
9  
10 5 Matthias Grot<sup>3</sup>, Esther Rind<sup>4</sup>, Claudia Pieper<sup>5</sup>, Karl-Heinz Jöckel<sup>5 6</sup>, Karen Linden<sup>1</sup>, Monika A. Rieger<sup>4</sup>,  
11 6 Birgitta M. Weltermann<sup>1</sup> on behalf of the IMPROVEjob consortium  
12  
13  
14 7

15  
16 8 <sup>1</sup> Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn,  
17 9 Venusberg-Campus 1, 53127 Bonn, Germany

20 10 <sup>2</sup> Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tuebingen,  
21 11 Osianderstraße 5, 72076 Tuebingen, Germany

23  
24 12 <sup>3</sup> Institute of Management, Operations Research, Ruhr University Bochum, Universitätsstr. 150,  
25 13 44801 Bochum, Germany

27  
28 14 <sup>4</sup> Institute of Occupational and Social Medicine and Health Services Research, University Hospital  
29 15 Tuebingen, Wilhelmstr. 27, 72074 Tuebingen, Germany

31  
32 16 <sup>5</sup> Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University  
33 17 of Duisburg-Essen, Hufelandstr. 55, 45147 Essen, Germany

35  
36 18 <sup>6</sup> Center for Clinical Trials, University Hospital Essen, University of Duisburg-Essen, Hufelandstr. 55,  
37 19 45147 Essen, Germany.  
40  
41 20

42 21 \*Correspondence to: manuelaschmidt@uni-bonn.de  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 22 Abstract

23 **Objectives:** This study aims to identify general practitioners' leadership skills using self- and staff  
24 ratings based on the Full Range of Leadership Model (FRLM) and the Leader-Member Exchange (LMX).

25 **Setting:** Cross-sectional analysis with questionnaires among German general practice staff  
26 participating in the IMPROVE*job* trial.

27 **Participants:** The study population comprised 60 practices with 366 participants, of which 84 GP  
28 leaders and 282 staff members (254 practice assistants and 28 employed physicians).

29 **Primary and secondary outcome measures:** Leadership was measured using the Integrative  
30 Leadership Questionnaire (German *FIF*) and the Leader-Member Exchange (LMX-7) questionnaire, with  
31 leaders rating themselves and practice staff rating their leaders. The thus provided 180° feedback was  
32 analyzed by paired mean comparisons at the participant level and Kruskal-Wallis tests at the practice  
33 level.

34 **Results:** Practice owners' self-ratings were higher than practice assistants' ratings for all leadership  
35 dimensions. Interestingly, employed physicians' ratings were higher for the dimensions 'performance  
36 development' and 'providing a vision', as well as for transactional leadership. Statistically significant  
37 differences were found for transformational leadership ( $p < .001$ ,  $d = .41$ ), especially for the dimensions  
38 'innovation' ( $p < .001$ ,  $d = .69$ ) and 'individuality focus' ( $p < .001$ ,  $d = .50$ ). For transactional leadership, only  
39 the dimension 'goal setting' showed significant differences ( $p < .01$ ,  $d = .30$ ). There were no significant  
40 differences between single and group practices. The LMX-7 scale (ranging from 7 to 35) showed a high  
41 relationship quality for both leaders and staff (26 for practice assistants and 28 for practice owners).

42 **Conclusions:** This analysis of GP leaders' leadership with self- and staff ratings showed consistent  
43 ratings of good relationships, but also highlighted the potential for leadership interventions to improve  
44 goal-setting, innovation, and individuality focus.

45 **Trial registration:** German Clinical Trials Register, DRKS00012677. Registered 16 October 2019.

46

47 **Keywords:** leadership, leadership quality, general physician, practice, 180-degree feedback, working  
48 conditions

49

### 50 Strengths & limitations:

- 51 • Presentation of an innovative 180° feedback approach in the GP setting which allows for  
52 analysis of the different occupational fields in the practice
- 53 • The data reflects a typical spectrum of German general practices including teaching and non-  
54 teaching practices as well as solo and group practices
- 55 • Interpretation of the results is limited by the data's cross-sectional nature therefore cause-  
56 effect analyses are not possible

- 1  
2  
3 57 • The results for group practices are limited as staff were asked to rate practice leaders in  
4 general, not on individual level. However, this approach reflects leadership in group practices  
5 58  
6 59 typically executed by a leadership team.  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
- The results for group practices are limited as staff were asked to rate practice leaders in general, not on individual level. However, this approach reflects leadership in group practices typically executed by a leadership team.

For peer review only

## 60 **Background**

61 Leadership has become an important topic in the medical field, as associations between poor  
62 leadership and various outcomes at leader, staff, and patient levels are described. Leadership issues  
63 comprise lack of organization, recognition for work, and opportunities to use one's abilities. Among  
64 the related adverse outcomes are low job satisfaction [1, 2], lack of staff well-being [3], and poor  
65 quality of patient care [1, 4, 5].

66  
67 Successful efforts to improve leadership have been undertaken in the hospital [6] and healthcare  
68 management settings [7], while leadership in primary care practices remains poorly addressed.  
69 However, improving leadership in this setting is important because general practitioners (GPs) and  
70 practice assistants were twice as likely to report high chronic stress as the general population [8]. Also,  
71 higher rates of burnout were documented, especially for young, female, and part-time employed  
72 physicians in GP group practices [9]. So far, there are no studies addressing leadership issues in primary  
73 care grounded on current leadership theories and frameworks for action, although theory-based  
74 leadership intervention(s) have improved respective skills in other medical [6] and non-medical fields  
75 [10, 11].

76  
77 One of the most significant leadership frameworks is the Full Range of Leadership Model (FRLM), which  
78 integrates transactional, transformational, and negative leadership [12, 13]. Transactional leadership  
79 describes leaders' structuring of work situations, the exchange of contingent rewards (e.g., work  
80 against salary), and management by exception [12–14]. In contrast, transformational leadership moves  
81 beyond leaders' and staff's self-interests. It focuses on the staff's attitudes and values regarding  
82 overarching goals such as self-actualization, organizational achievements, and the well-being of others  
83 and society as a whole [12, 13]. Associated leadership behaviors include inspirational motivation,  
84 intellectual stimulation, and individualized consideration. Additional aspects are charismatic  
85 relationships based on trust and confidence (so-called attributed idealized behavior) and a collective  
86 sense of action (so-called idealized influence behavior) [14]. Although they are described as different  
87 leadership behaviors, studies indicate that transactional and transformational leadership are highly  
88 interrelated [15]. Transactional leadership is often valued as the leadership foundation, while  
89 additional transformational leadership creates an 'augmentation effect' that may stimulate staff's  
90 extra efforts and high performance, as shown for health care workers [16]. In various settings,  
91 transformational and transactional leadership influenced organizational outcomes positively, e.g.,  
92 performance indicators and employees' job satisfaction [14, 15]. Transformational leadership is  
93 associated with lower job stress and strain, less anxiety, higher well-being, and better outcomes for  
94 occupational safety [17]. While transformational and transactional leadership include desired

1  
2  
3 95 behaviors, a leader can also behave in a way that is detrimental to employees and the organization as  
4  
5 96 a whole [18]. This negative leadership style includes laissez-faire leadership, which represents a highly  
6  
7 97 passive leadership style where the leader's activities are at a minimum [12, 14], and destructive  
8  
9 98 leadership, where harmful interpersonal behavior takes place which is not related to the leadership  
10  
11 99 task [18].

12 100  
13 101 Another essential leadership theory, the Leader-Member Exchange (LMX), reflects the relationship  
14  
15 102 between leaders and staff. It concentrates on the perceived quality of the dyadic relationship between  
16  
17 103 a staff member and their immediate leader [12, 19]. A high-quality relationship positively influences  
18  
19 104 employees' work-related well-being and is associated with higher job satisfaction for health care  
20  
21 105 workers [16].

22 106  
23 107 The described focus on the relationship between leaders and staff is reflected in current  
24  
25 108 methodological strategies. For example, the Leader-Member Exchange questionnaire 'LMX-7' [19] and  
26  
27 109 the newly developed German questionnaire 'Fragebogen für integrative Führung' (FIF; in English:  
28  
29 110 Questionnaire for Integrative Leadership) (Rowold & Poethke, 2017; Rowold & Schlotz, 2009) allow for  
30  
31 111 a multi-rater perspective: the leaders' and the staff's views on the leaders' behavior are measured and  
32  
33 112 compared providing 180° feedback. Including assessments from different perspectives creates a more  
34  
35 113 comprehensive picture of the leaders' actual skills and performances [20, 21]. While such approaches  
36  
37 114 are increasingly applied in medical education and graduate training [22, 23], they have not been used  
38  
39 115 to evaluate GP leaders.

40 116  
41 117 The publicly funded *IMPROVEjob* study aims to improve the job satisfaction of physician leaders and  
42  
43 118 practice personnel of German GP practices focusing on leadership, communication, and work  
44  
45 119 processes [24, 25]. Based on the described leadership concepts, a 180° feedback approach compared  
46  
47 120 GP leaders' self-ratings and their staff's ratings of their leadership skills.

## 48 122 **Methods**

49 123 This analysis draws on the baseline data of the *IMPROVEjob* study [25], which is designed as a cluster-  
50  
51 124 randomized controlled trial (cRCT) to improve job satisfaction among practice personnel. The details  
52  
53 125 are described in the study protocol [24].

## 54 126 55 127 **Participants**

56  
57 128 A total of 56 GP practices in the North Rhine region in Germany were recruited by the Institute of  
58  
59 129 General Practice and Family Medicine of the University of Bonn, aiming for approximately equal strata



1  
2  
3 130 of single and group practices as well as of teaching and non-teaching practices. The study aimed to  
4  
5 131 recruit all practice team staff, including physician leaders, employed physicians, and practice  
6  
7 132 assistants. Practice owners answered a short questionnaire on practice characteristics and the  
8  
9 133 questionnaire for practice leaders. Employed physicians and practice assistants filled out different  
10  
11 134 versions of the same employee questionnaire. For details, see [24].  
12

135

### 136 ***Patient and public involvement***

137 No patient involved

138

### 139 ***Ethics***

140 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn  
141 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the  
142 Medical Association North-Rhine (Lfd-Nr.: 2019107) and of the Medical Faculty, University Hospital of  
143 Tuebingen (project no.: 446/2019BO2) approved the study protocol. The study was performed in  
144 accordance with the Declaration of Helsinki. All participating practice team members received written  
145 information and signed informed consent forms.

146

### 147 ***Measures***

148 This analysis uses the IMPROVE<sub>job</sub> participants' baseline data on sociodemographic, professional, and  
149 work-related characteristics [25], as well as the following two leadership scales:

150

#### 151 1. Integrative Leadership Questionnaire (FIF)

152 Transformational, transactional, and negative leadership were measured using the FIF  
153 questionnaire [26]. Its scales' validity and internal consistency are confirmed for different  
154 populations [26, 27]. The FIF has been used in non-medical [28] and hospital settings [29], but not  
155 in primary care.

156 All 40 items of the FIF are answered on a 5-point Likert scale and are worded to reflect either the  
157 leaders' or the staff's position [26].

158 The measures comprise:

- 159 - the transformational leadership scale, consisting of six dimensions: innovation, team spirit,  
160 performance development, individuality focus, providing a vision, and being a role model;
- 161 - the transactional leadership scale with two dimensions: goal setting, and management by  
162 exception; and
- 163 - the negative leadership scale with two dimensions: laissez-faire, and destructive leadership.

164

1  
2  
3 165 2. Leader-Member Exchange (LMX-7)

4  
5 166 The relationship quality between leaders and staff is measured using the Leader-Member  
6 167 Exchange questionnaire (LMX-7) with seven items (5-point Likert scales), which are also worded to  
7  
8 168 reflect the leader or staff position [19, 30, 31].  
9

10 169

11 170 **180° feedback**

12  
13 171 The multi-rater, 180° approach is applied to the leadership scales, measuring leadership skills from  
14 172 different perspectives within GP practices to create a more comprehensive picture of the actual  
15 173 performance of the leader in question [20, 21] and improve the accuracy of the assessment [22, 23].  
16 174 While results of such assessments are usually shared with the ratee [26], previous studies showed  
17 175 mixed reactions in the medical setting [32, 33]. Therefore, the results of this 180° feedback are not  
18 176 shared with the participating practices but are used for research purposes only on an aggregated level.  
19  
20  
21  
22

23 177

24  
25 178 **Statistical analysis**

26 179 Statistical analyses were conducted with SAS version 9.4 for the sociodemographic characteristics and  
27 180 SPSS on Windows version 26 for additional analyses. All analyses were carried out at the participant  
28 181 and the practice level.  
29  
30  
31

32 182

33 183 The FIF data were analyzed according to the official manual [26]. Mean scores for transformational,  
34 184 transactional, and negative leadership were summarized both for the respective main scale and the  
35 185 dimensions: for transactional and transformational leadership, they ranged from 1 (worst rating) to 5  
36 186 (best rating); for negative leadership, they ranged from 1 (best rating) to 5 (worst rating). To allow for  
37 187 a comparison with other settings, scores were standardized using T-scaling tables as defined by Rowold  
38 188 & Poethke [26]. These T-values are based on a normal distribution around 50 (SD 10). Thus, values  
39 189 above 70 only reflect about 2% of the reference population from German-speaking countries [26].  
40  
41  
42  
43  
44

45 190

46 191 The LMX-7 was analyzed per standard protocol by creating a sum score of all seven items without  
47 192 transformation [31]. Thus, the LMX-7 score may range from 7 to 35, with five standard categories for  
48 193 interpretation: score 7 to 14 = very low; 15 to 19 = low; 20 to 24 = moderate; 25 to 29 = high, 30 to  
49 194 35 = very high [34]. Inadvertently, question seven was missing on all employed physicians'  
50 195 questionnaires, which reduced the answered questions to six. As the LMX-7 manual does not suggest  
51 196 a standard approach for missing values, we excluded employed physicians from further analyses.  
52  
53  
54  
55  
56

57 197

58 198 Following the 180° feedback approach, the combined mean scores of employed physicians and  
59 199 practice assistants per practice were compared to the self-assessment of their respective leaders using  
60

200 paired t-tests, as the data satisfied the condition of a normal distribution with the Kolmogorov-Smirnov  
 201 test. Cohen's d was applied to determine the effect size of mean comparisons with the following  
 202 standard interpretations: small effects from  $d=.2$ , medium from  $d=.5$ , and high from  $d=.8$  [35].

203  
 204 Self-ratings were available for each leader, as they were part of the individual questionnaire, allowing  
 205 for 84 ratings in 60 practices. In single practices, the staff ratings were compared to the leader's  
 206 assessment. In practices with more than one owner (group practices), each leader's self-rating was  
 207 compared with the respective rating of the practice personnel, who were asked to rate leadership in  
 208 the entire practice, not stratified by individuals.

209  
 210 In addition, the ratings of the transformational and transactional leadership scales were compared by  
 211 practice type (group vs. single and teaching vs. non-teaching practices) using the Kruskal-Wallis tests  
 212 because the data for practice comparisons did not satisfy the conditions for parametric tests.

213

## 214 **Results**

### 215 ***Population***

216 The baseline data of the IMPROVE*job* study [25] included 366 participants from 60 practices, consisting  
 217 of 84 practice owners, 28 employed physicians, and 254 practice assistants. The mean age of all  
 218 participants was 44.4 years, with a mean of 54.3 years for practice owners, 44.8 for employed  
 219 physicians, and 41.0 for practice assistants. Among the practice assistants, 99.6% were female, as were  
 220 76.6% of the employed physicians and half of the practice owners (52.4%). Most practice owners  
 221 worked full-time (90.5%), as did about a quarter of the employed physicians (28.6%) and 41.5% of the  
 222 practice assistants (see Table 1). For details on the sociodemographic descriptions, see [25].

223

224 **Table 1: Sociodemographic description of participants at baseline [25]**

Variable	Total sample N=366	Practice owners N=84	Employed physicians N=28	Practice assistants N=254
Female, %	87.1	52.4	78.6	99.6
Age in years, mean (SD)	44.4 (12.8)	54.3 (6.2)	44.8 (9.8)	41.0 (13.0)
Years in current practice, mean (SD)	10.0 (9.1)	15.3 (8.4)	3.9 (5.4)	8.8 (8.9)
Working full-time, %	52.0	90.5	28.6	41.5
Persons in household over 18 years, mean (SD)	2.2 (1.0)	2.1 (1.0)	2.0 (0.5)	2.2 (1.1)
Persons in household under 18 years, mean (SD)	1.2 (1.0)	1.3 (1.3)	1.4 (1.0)	1.0 (0.9)
Care for next-of-kin, %	20.8	21.7	0.0	22.9

225

226 On average, practice owners had been accredited for 26.6 years and licensed for the statutory health  
 227 insurance for 16.4 years. Seven (25%) of the employed physicians were in GP training. Practice  
 228 assistants had graduated on average 19.9 years ago, while 7.5% were still in training.

229  
 230 Of the 60 practices, 21 (35%) were single and 39 (65%) were group practices; of these, 34 were  
 231 teaching (57%) and 26 (43%) were non-teaching practices. On average, practices had been located in  
 232 their area for 20.4 years (SD 14 years).

233

### 234 **Leadership**

235 The transactional and transformational leadership scales showed a high internal consistency with  
 236 Cronbach's  $\alpha = .74$  to  $.93$  for the staff members' assessment and Cronbach's  $\alpha = .72$  to  $.87$  for the  
 237 leaders' assessment. For negative leadership, the scales showed a sufficient internal consistency for  
 238 staff members' (Cronbach's  $\alpha = .73$  to  $.80$ ) but not for leaders' assessments (Cronbach's  $\alpha = .47$  to  $.68$ ).

239

240 The mean results of the FIF were within the reference population's average (one standard deviation)  
 241 range. Based on raw values, employed physicians rated their leaders consistently better than practice  
 242 assistants and—in part—also better than the leaders themselves. While practice assistants rated their  
 243 leaders more poorly than the practice owners in raw values, reference T-values showed only minor  
 244 differences; see Table 2 for details.

245

246 The LMX-7 scale showed an internal consistency of Cronbach's  $\alpha = .88$  for staff members (practice  
 247 assistants) and  $\alpha = .71$  for leaders. All groups showed a high relationship quality, scoring 28 for practice  
 248 owners and 26 for practice assistants. As the seventh question was missing for employed physicians,  
 249 they were excluded from the analysis. However, the sum score of the remaining six questions also  
 250 showed a high score of 24.9 out of 30; see Table 2 for details.

251

252 **Table 2: Leadership assessment by employment group:** main scales (in bold) and dimensions (LMX  
 253 values can range from 7 to 35, FIF scales from 1 to 5)

	Practice owners (N=84)				Employed physicians (N=28)				Practice assistants (N=254)			
	M	SD	T*	n	M	SD	T*	N	M	SD	T*	n
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.6</b>	<b>45</b>	<b>84</b>	<b>3.9</b>	<b>0.7</b>	<b>56</b>	<b>27</b>	<b>3.5</b>	<b>0.8</b>	<b>52</b>	<b>237</b>
Innovation	4.2	0.6	49	84	4.0	1.0	55	28	3.7	0.9	52	247
Team spirit	4.1	0.7	49	84	3.8	1.0	54	28	3.6	1.1	52	251
Performance development	3.6	0.8	44	84	4.1	0.7	57	27	3.5	1.0	51	247
Individuality focus	3.9	0.7	47	84	3.7	1.0	54	28	3.5	1.1	53	249

Providing a vision	3.5	0.9	45	84	3.6	0.9	55	28	3.2	1.1	51	245
Being a role model	4.1	0.6	45	84	4.0	0.8	55	27	3.7	1.0	52	246
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.7</b>	<b>47</b>	<b>83</b>	<b>3.5</b>	<b>0.7</b>	<b>54</b>	<b>27</b>	<b>3.2</b>	<b>0.8</b>	<b>50</b>	<b>244</b>
Goal setting	3.5	0.7	44	83	3.7	0.9	56	27	3.1	1.0	50	246
Management by exception	3.3	0.8	51	83	3.4	0.8	52	27	3.3	0.9	51	245
<b>Negative leadership</b>	<b>1.5</b>	<b>0.5</b>	<b>51</b>	<b>83</b>	<b>1.5</b>	<b>0.6</b>	<b>45</b>	<b>28</b>	<b>1.7</b>	<b>0.7</b>	<b>47</b>	<b>248</b>
Laissez-faire	1.6	0.6	52	83	1.6	0.8	45	28	1.7	0.8	46	249
Destructive	1.4	0.5	51	83	1.4	0.6	46	28	1.6	0.7	48	248
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	-	<b>81</b>	<b>n/a</b>	<b>n/a</b>	-	<b>n/a</b>	<b>26.7</b>	<b>4.8</b>	-	<b>222</b>

\*Reference T-values range from 0 to 100, as defined by Rowold & Poethke 2017

### 180° leadership feedback

Practice owners self-rated their leadership skills slightly better than their staff for all dimensions except for 'management by exception'. There were no statistically significant differences for negative leadership. For transactional leadership, goal setting showed a statistically significant difference with a low effect size ( $p=.009$ ,  $d=.30$ ). Leaders' scores on transformational leadership were significantly higher than the scores of the teams, with the dimension for innovation reaching the strongest effect size ( $p<.001$ ,  $d=.69$ ), followed by individuality focus with a medium effect size ( $p<.001$ ,  $d=.50$ ), and team spirit as well as being a role model with slightly smaller, but significant effect sizes. The main scale for transformational leadership also showed a statistically significant difference with a medium effect size ( $p<.001$ ,  $d=.41$ ). For details, see Table 3.

**Table 3: Comparison of leaders' self- and staff ratings (n=84 leader-team pairs): main scales (in bold) and dimensions**

	Practice owners		Practice staff		Paired t-test		
	M	SD	M	SD	t(df)	p	d
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.5</b>	<b>3.6</b>	<b>0.6</b>	<b>3.721(82)</b>	<b>&lt;.001</b>	<b>0.41</b>
Innovation	4.2	0.6	3.8	0.6	6.359(83)	<.001	0.69
Team spirit	4.1	0.7	3.8	0.7	3.462(82)	.001	0.38
Performance development	3.6	0.8	3.7	0.6	-0.208(83)	.836	-
Individuality focus	3.9	0.7	3.5	0.6	4.633(83)	<.001	0.50
Providing a vision	3.5	0.9	3.3	0.8	1.592(82)	.115	-
Being a role model	4.1	0.6	3.8	0.6	2.833(82)	.006	0.31
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.6</b>	<b>3.3</b>	<b>0.5</b>	<b>1.291(81)</b>	<b>.200</b>	-
Goal setting	3.5	0.7	3.2	0.6	2.681(81)	.009	0.30
Management by exception	3.3	0.8	3.4	0.6	-0.470(82)	.640	-
<b>Negative leadership</b>	<b>1.5</b>	<b>0.4</b>	<b>1.6</b>	<b>0.4</b>	<b>-1.744(82)</b>	<b>.085</b>	-
Laissez-faire	1.6	0.6	1.7	0.5	-1.563(82)	.122	-
Destructive	1.4	0.5	1.6	0.5	-1.514(82)	.134	-
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	<b>26.8</b>	<b>3.5</b>	<b>3.275(79)</b>	<b>.002</b>	<b>0.37</b>

### Transformational and transactional leadership by practice type

271 The Kruskal-Wallis test was applied to analyze for differences in leadership by practice types. It showed  
 272 slight but non-significant differences in the raw values between practice types (single vs. group,  
 273 teaching vs. non-teaching practices), e.g., slightly higher ratings for transformational leadership in  
 274 single and non-teaching practices. These slight differences persisted when using reference T-values.  
 275 For details, see Table 4.

276

277 **Table 4: Comparison of leadership assessments by practice type: single / group practices; teaching /**  
 278 **non-teaching practices**

	Single (n=21)			Group (n=39)			Non-teaching (n=26)			Teaching (n=34)		
	M	T	n	M	T	N	M	T	n	M	T	n
<b>Practice owners</b>												
Transformational	4.0	47	21	3.9	45	63	3.8	43	37	4.0	47	47
Transactional	3.4	47	21	3.4	47	62	3.3	45	37	3.4	47	46
Negative	1.5	51	21	1.5	51	62	1.6	53	37	1.5	51	46
LMX-7	28.8	-	20	27.9	-	61	27.5	-	36	28.6	-	45
<b>Practice staff</b>												
Transformational	3.7	54	67	3.6	53	212	3.5	52	117	3.6	53	162
Transactional	3.4	53	67	3.3	51	212	3.3	51	117	3.2	50	162
Negative	1.7	47	70	1.7	47	212	1.7	47	117	1.6	46	165
LMX-7	27.5	-	61	26.3	-	190	25.6	-	105	27.3	-	146

279

## 280 Discussion

281 To our knowledge, this is the first study on GP practice leaders to assess transactional and  
 282 transformational leadership using 180° feedback approach. Practice owners and practice personnel  
 283 showed good relationships and a similar understanding of leadership behaviors regarding low levels of  
 284 negative leadership and moderate levels of transactional leadership. Yet, potential for optimization  
 285 was shown for transformational leadership, especially regarding the dimensions innovation, team  
 286 spirit, and individuality focus.

287

288 The 180° feedback approach, also called multi-rater assessment, was shown to provide a more realistic  
 289 picture of leader-team situations [20, 21]. In international studies, 360° or 180° feedback has been  
 290 found to be a good predictor for promotions [36] and leadership effectiveness [37]. This type of  
 291 feedback has the potential to identify differences in leadership expectations on behalf of the staff and  
 292 the leader. Even though ratings are not low on either side, there is a discrepancy between self- and  
 293 staff ratings, especially for transformational leadership as a more visionary approach. Such differences  
 294 in the perceptions of leadership are influenced by mutual expectations. They are associated with job  
 295 satisfaction, which is higher when the ratings are in agreement [38]. In our study, the gap in the

1  
2  
3 296 perception of leadership was larger for practice assistants than for employed physicians. This is likely  
4  
5 297 related to the fact that employed physicians are much closer to their physician leader regarding  
6  
7 298 training, roles, and duties compared to practice assistants. Also, practice assistants do not have the  
8  
9 299 prospects to become physician leaders themselves, which implies a fundamentally different  
10  
11 300 perspective. While such differences cannot be overcome, a better mutual understanding of leaders  
12  
13 301 and staff can improve workplaces. Interestingly, staff members who perceive themselves as more  
14  
15 302 similar to the leader also give better ratings [39]. Rowold and Poethke [26] describe how leaders can  
16  
17 303 learn to adapt when receiving this feedback and recommend implementing changes, such as regular  
18  
19 304 team meetings or improving leadership skills through training. Although multi-rater assessments were  
20  
21 305 shown to provide a higher accuracy of ratings in the medical setting and are applied in various levels  
22  
23 306 of medical training [22, 23], this approach has not been implemented routinely in physician leadership  
24  
25 307 training.

26  
27 308  
28  
29 309 Overall, the self- and staff ratings showed average or moderate levels when compared to a  
30  
31 310 representative sample of German leaders provided by the manual of the FIF questionnaire [26].  
32  
33 311 Notably, these reference values are already corrected for aspects such as self-protective response and  
34  
35 312 social desirability biases [40, 41]. These results are comparable to international studies on GP  
36  
37 313 leadership, where GPs were aware of their leadership role in general but failed to describe explicit  
38  
39 314 actions and perspectives for everyday practice situations [42]. Spehar et al. suggested that a lack of  
40  
41 315 leadership training and credentials may play a role [43].

42  
43 316  
44  
45 317 In line with a larger study in a German hospital setting [29], practice leaders' perception of their  
46  
47 318 leadership skills was moderate. Yet, the differences between self- and staff ratings were smaller in our  
48  
49 319 study. This might have been due to the sample selection, as leaders and employees in the current study  
50  
51 320 were analyzed at the practice team level while only aggregate analyses of leaders' and staff's  
52  
53 321 assessments were possible in the hospital setting. There were no differences in leaders' and staff's  
54  
55 322 assessments of transactional and transformational leadership between practice types, although  
56  
57 323 differences were observed when looking at mental health outcomes such as burnout [9].

58  
59 324  
60  
61 325 Applying the LMX-7 questionnaire, prior studies on relationship quality in the health sector showed  
62  
63 326 associations with enhanced commitment, reduced staff turnover, and better organizational behavior  
64  
65 327 [44]. Also, positive effects on employees' health and well-being were outlined [45, 46]. Our results  
66  
67 328 show a good relationship quality between practice owners and practice assistants when applying  
68  
69 329 reference values [34].

70 330

1  
2  
3 331 ***Strengths and limitations***

4 332 Our study presents an innovative implementation of the 180° feedback approach in the GP setting.  
5  
6 333 The relatively large sample size and the analysis at practice team levels are noteworthy. The results for  
7  
8 334 group practices with two or more leaders are limited as staff were asked to rate practice leaders as a  
9  
10 335 group, not individually. On the other hand, this approach reflects leadership in group practices typically  
11  
12 336 executed by a leadership team. Given the data's cross-sectional nature, cause-effect analyses are not  
13  
14 337 possible. The negative leadership scale showed a low reliability but was included according to the  
15  
16 338 respective manual [26]. For the small number of employed physicians, complete data on the Leader-  
17  
18 339 Member Exchange questionnaire were missing, yet the analysis of the available six rather than seven  
19  
20 340 questions yielded a high relationship quality similar to that documented for practice assistants.

21 341  
22 342 The stratified randomization, which took into account teaching and non-teaching practices as well as  
23  
24 343 solo and group practices, reflects a typical spectrum of German general practices. However, a selection  
25  
26 344 bias cannot be excluded as participating practice leaders might have had a greater interest in the topic.  
27  
28 345

28 346 ***Conclusion and practical implications***

29  
30 347 Compared to the reference populations in the literature, GPs and their practice teams have a good  
31  
32 348 relationship quality. However, transactional and transformational leadership skills show potential for  
33  
34 349 improvement, especially regarding the dimensions goal-setting, innovation, and individuality focus.  
35  
36 350 Recent developments call for leadership workshops at every level of medical training [47] in order to  
37  
38 351 help GP leaders better understand staff's needs. Future results from the IMPROVE<sub>job</sub> study will show  
39  
40 352 if the intervention changed leadership skills and their perception.

41 353  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 354 **Formalities**  
4

5 355

6 356 **Acknowledgment**  
7

8 357 All authors are part of the *IMPROVEjob* consortium and owe special thanks to all participating  
9  
10 358 physicians and practice assistants. In addition, special thanks go to the other members of the  
11  
12 359 *IMPROVEjob* consortium; current members: M. Brinkmann, S. Burgess, A. Herrmann-Werner, K.-H.  
13  
14 360 Jöckel, F. Junne, A. Siegel, F. Stuber, A. Wagner, and former members: J. Block, J.-M. Bois, A. Dreher,  
15  
16 361 M. Hippler, S. Hartmann, L. Imhoff, C. Kersting, C. Ose and E. Tsarouha. Additionally, we are grateful to  
17  
18 362 all members of the scientific advisory board: Stefanie Esper, Katja Götz, Andre Habrock, Albert  
19  
20 363 Nienhaus, Horst Christian Vollmar, and Andrea Wittich. The work of the Institute of Occupational and  
21  
22 364 Social Medicine and Health Services Research is supported by an unrestricted grant of the Employers'  
23  
24 365 Association of the Metal and Electrical Industry Baden-Wuerttemberg (Suedwestmetall). We  
25  
26 366 acknowledge administrative support by the DLR Project Management Agency.  
27

28 367

29 368 **Contributorship statement**

30 369 Conceptualization, B.M.W., M.A.R., T.S.-D., C.P., and B.W.; methodology, B.M.W. M.A.R., T.S.-D., M.S.,  
31  
32 370 B.W., M.G., E.R., C.P., K.-H.J., L.D., and J.G.; validation, K.-H.J.; formal analysis, M.S. and J.G.;  
33  
34 371 investigation, L.D., K.L., T.S.-D., B.W., E.R., C.P., J.G., M.A.R., and B.M.W.; resources, B.M.W., L.D., K.L.,  
35  
36 372 T.S.-D., B.W., M.G., C.P., and E.R.; data curation, K.-H.J.; writing—original draft preparation, M.S. and  
37  
38 373 B.M.W.; writing—review and editing, M.S., T.S.-D., B.W., M.G., E.R., C.P., K.-H.J., L.D., J.G., M.A.R. and  
39  
40 374 B.M.W.; visualization, M.S.; supervision, B.M.W., M.A.R., and B.W.; project administration, K.L., E.R.,  
41  
42 375 B.M.W. and M.A.R.; funding acquisition, B.M.W., M.A.R., C.P., and B.W. All authors have read and  
43  
44 376 approved the published version of the manuscript.  
45

46 377

47 378 **Competing interests**

48 379 The authors declare that they have no competing interests.  
49

50 380

51 381 **Funding**

52 382 This study is funded by the German Federal Ministry for Education and Research (grant numbers:  
53  
54 383 01GL1851D, 01GL1751B, 01GL1751A, 01GL1751C). The funder had no role in the study design, the data  
55  
56 384 collection, and analysis, the decision to publish, or the preparation of the manuscript.  
57

58 385

59 386 **Data sharing statement**

60 387 There are no plans to grant access to the full protocol, participant-level dataset, or statistical code as  
388 388 data contain potentially identifying information.

1  
2  
3 389 **Ethics approval**

4  
5 390 The study complies with the ethical principles of the World Medical Association Declaration of Helsinki.

6 391 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn

7  
8 392 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the

9  
10 393 Medical Association Nordrhein (Lfd-Nr.: 2019107), and of the Medical Faculty, University Hospital of

11 394 Tuebingen (project no.: 446/2019BO2) approved the study protocol.

12  
13 395 All participating practice team members received written information and signed informed consent

14  
15 396 forms, which are stored at the Institute for General Practice and Family Medicine, University of Bonn.

16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

## References

- 1 Boamah SA, Spence Laschinger HK, Wong C, et al. Effect of transformational leadership on job satisfaction and patient safety outcomes. *Nurs Outlook* 2018;66(2):180–89. doi:10.1016/j.outlook.2017.10.004 [published Online First: 23 November 2017].
- 2 Goetz K, Campbell S, Broge B, et al. Job satisfaction of practice assistants in general practice in Germany: an observational study. *Fam Pract* 2013;30(4):411–17. doi:10.1093/fampra/cmt015 [published Online First: 19 April 2013].
- 3 Kersting C, Zimmer L, Thielmann A, et al. Chronic stress, work-related daily challenges and medicolegal investigations: a cross-sectional study among German general practitioners. *BMC Fam Pract* 2019;143. doi:10.1186/s12875-019-1032-6 [published Online First: 24 October 2019].
- 4 Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *The Lancet* 2009;374(9702):1714–21.
- 5 Wong CA, Cummings GG, Ducharme L. The relationship between nursing leadership and patient outcomes: a systematic review update. *J Nurs Manag* 2013;21(5):709–24.
- 6 Saravo B, Netzel J, Kiesewetter J. The need for strong clinical leaders - Transformational and transactional leadership as a framework for resident leadership training. *PLoS One* 2017;12(8):e0183019.
- 7 Richter A, Lornudd C, Thiele Schwarz U von, et al. Evaluation of iLead, a generic implementation leadership intervention: mixed-method preintervention-postintervention design. *BMJ Open* 2020;10(1):e033227. doi:10.1136/bmjopen-2019-033227 [published Online First: 12 January 2020].
- 8 Viehmann A, Kersting C, Thielmann A, et al. Prevalence of chronic stress in general practitioners and practice assistants: Personal, practice and regional characteristics. *PLoS One* 2017;12(5):e0176658.
- 9 Dreher A, Theune M, Kersting C, et al. Prevalence of burnout among German general practitioners: Comparison of physicians working in solo and group practices. *PLoS One* 2019;14(2):e0211223.
- 10 Abrell C, Rowold J, Weibler J, et al. Evaluation of a long-term transformational leadership development program. *German Journal of Human Resource Management* 2011;25(3):205–24.
- 11 Kelloway EK, Barling J, Helleur J. Enhancing transformational leadership: The roles of training and feedback. *Leadership & Organization Development Journal* 2000;21(3):145–49.
- 12 Bass BM. Two Decades of Research and Development in Transformational Leadership. *European Journal of Work and Organizational Psychology* 1999;8(1):9–32.
- 13 Poethke U. Leadership, values and communication: a cross-cultural investigation of the extended full-range of leadership behaviors: Doctoral dissertation 2019.
- 14 Rowold J, Schlotz W. Transformational and transactional leadership and followers' chronic stress. *Leadership Review* 2009;9(1):35–48.
- 15 Judge TA, Piccolo RF. Transformational and transactional leadership: a meta-analytic test of their relative validity. *J Appl Psychol* 2004;89(5):755–68.
- 16 Gregersen S, Vincent-Höper S, Nienhaus A. Health-Relevant Leadership Behaviour: A Comparison of Leadership Constructs. *German Journal of Human Resource Management* 2014;28(1-2):117–38.
- 17 Kelloway EK, Barling J. Leadership development as an intervention in occupational health psychology. *Work & Stress* 2010;24(3):260–79.
- 18 Schyns B, Schilling J. How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly* 2013;24(1):138–58.

- 1  
2  
3 444 19 Graen GB, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-  
4 445 member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-  
5 446 domain perspective. *Leadership Quarterly* 1995;6(2):219–47.
- 6 447 20 Bracken D, Timmreck C, Church A. The Handbook of Multisource Feedback The Comprehensive  
7 448 Resource for Designing and Implementing MSF Processes 2001.
- 8 449 21 Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ*  
9 450 *Health Prof* 2003;23(1):4–12.
- 10 451 22 Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with  
11 452 observed measures of competence: a systematic review. *JAMA* 2006;296(9):1094–102.
- 12 453 23 Donnon T, Al Ansari A, Al Alawi S, et al. The reliability, validity, and feasibility of multisource  
13 454 feedback physician assessment: a systematic review. *Acad Med* 2014;89(3):511–16.
- 14 455 24 Weltermann BM, Kersting C, Pieper C, et al. IMPROVEjob - Participatory intervention to  
15 456 improve job satisfaction of general practice teams: a model for structural and behavioural  
16 457 prevention in small and medium-sized enterprises - a study protocol of a cluster-randomised  
17 458 controlled trial. *Trials* 2020;21(1):532.
- 18 459 25 Degen L, Linden K, Seifried-Dübön T, et al. Job Satisfaction and Chronic Stress of General  
19 460 Practitioners and Their Teams: Baseline Data of a Cluster-Randomised Trial (IMPROVEjob). *Int J*  
20 461 *Environ Res Public Health* 2021;18(18). doi:10.3390/ijerph18189458 [published Online First: 8  
21 462 September 2021].
- 22 463 26 Rowold J, Poethke U. FIF: Fragebogen zur Integrativen Führung Manual, 1st edn. Göttingen:  
23 464 Hogrefe 2017.
- 24 465 27 Ingold PV. Fragebogen zur integrativen Führung. *Zeitschrift für Arbeits- und*  
25 466 *Organisationspsychologie A&O* 2019;63(2):100–06.
- 26 467 28 Lange S, Bormann KC, Rowold J. Mindful leadership: mindfulness as a new antecedent of  
27 468 destructive and transformational leadership behavior. *Gr Interakt Org* 2018;49(2):139–47.
- 28 469 29 Stuber F, Seifried-Dübön T, Rieger MA, et al. Investigating the Role of Stress-Preventive  
29 470 Leadership in the Workplace Hospital: The Cross-Sectional Determination of Relational Quality  
30 471 by Transformational Leadership. *Front Psychiatry* 2019;10:622. doi:10.3389/fpsy.2019.00622  
31 472 [published Online First: 3 September 2019].
- 32 473 30 Schyns B. Überprüfung einer deutschsprachigen Skala zum Leader-Member-Exchange-Ansatz.  
33 474 *Zeitschrift für Differentielle und Diagnostische Psychologie* 2002.
- 34 475 31 Schyns B, Paul T. Leader-Member Exchange Skala (LMX7): ZIS - GESIS Leibniz Institute for the  
35 476 Social Sciences 2014.
- 36 477 32 Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback:  
37 478 perceptions of credibility and usefulness. *Med Educ* 2005;39(5):497–504.
- 38 479 33 Hu J, Lee R, Mullin S, et al. How physicians change: Multisource feedback driven intervention  
39 480 improves physician leadership and teamwork. *Surgery* 2020;168(4):714–23.  
40 481 doi:10.1016/j.surg.2020.06.008 [published Online First: 6 August 2020].
- 41 482 34 Northouse PG. Leadership: Theory and practice, 6th edn. Thousand Oaks: Sage Publications  
42 483 2010.
- 43 484 35 Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Hoboken: Taylor and  
44 485 Francis 1988.
- 45 486 36 Halverson SK, Tonidandel S, Barlow C, et al. Self-other agreement on a 360-degree leadership  
46 487 evaluation. In: 17th Annual Conference of the Society for Industrial and Organizational  
47 488 Psychology, Toronto, Canada 2002.
- 48 489 37 Furnham A, Stringfield P. Congruence of self and subordinate ratings of managerial practices as  
49 490 a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*  
50 491 1994;67(1):57–67.

- 1  
2  
3 492 38 Erben J, Schneider FM, Maier M. In the Ear of the Beholder: Self-Other Agreement in  
4 493 Leadership Communication and Its Relationship With Subordinates' Job Satisfaction.  
5 494 *International Journal of Business Communication* 2019;56(4):505–29.  
6  
7 495 39 Fleenor JW, Smither JW, Atwater LE, et al. Self–other rating agreement in leadership: A review.  
8 496 *The Leadership Quarterly* 2010;21(6):1005–34.  
9 497 40 Sarros JC, Cooper BK, Hartican AM. Leadership and character. *Leadership & Organization*  
10 498 *Development Journal* 2006;27(8):682–99.  
11 499 41 Thornton GC. Psychometric properties of self-appraisals of job performance. *Personnel*  
12 500 *Psychology* 1980;33(2):263–71.  
13 501 42 Hana J, Rudebeck CE. Leadership in rural medicine: the organization on thin ice? *Scandinavian*  
14 502 *Journal of Primary Health Care* 2011;29(2):122–28. doi:10.3109/02813432.2011.577148  
15 503 [published Online First: 28 April 2011].  
16 504 43 Spehar I, Sjøvik H, Karevold KI, et al. General practitioners' views on leadership roles and  
17 505 challenges in primary health care: a qualitative study. *Scandinavian Journal of Primary Health*  
18 506 *Care* 2017;35(1):105–10. doi:10.1080/02813432.2017.1288819 [published Online First: 1 March  
19 507 2017].  
20 508 44 Chen C-HV, Wang S-J, Chang W-C, et al. The effect of leader-member exchange, trust,  
21 509 supervisor support on organizational citizenship behavior in nurses. *Journal of Nursing Research*  
22 510 2008;16(4):321–28.  
23 511 45 Gregersen S, Vincent-Höper S, Nienhaus A. Job-related resources, leader–member exchange  
24 512 and well-being – a longitudinal study. *Work & Stress* 2016;30(4):356–73.  
25 513 46 Vincent-Höper S, Stein M, Gregersen S, et al. Messung gesundheitsförderlicher Führung in  
26 514 Gesundheitsdienst und Wohlfahrtspflege. *ASU Arbeitsmed Sozialmed Umweltmed* 2018(53):46–  
27 515 53.  
28 516 47 Frank JR, Snell L, Sherbino J. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal  
29 517 College of Physicians and Surgeons of Canada 2015.

1  
2  
3 **1 Additional Files**4  
5 **2 Additional file 1: CONSORT 2010 checklist of information to include when reporting a randomised trial**6  
7 38  
9 4  
5**CONSORT 2010 checklist of information to include when reporting a randomised trial\***

Section/Topic	Item No	Checklist item	Reported on page No
<b>Title and abstract</b>			
	1a	Identification as a randomised trial in the title	n/a
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Abstract, 2
<b>Introduction</b>			
Background and objectives	2a	Scientific background and explanation of rationale	Background, 4-5
	2b	Specific objectives or hypotheses	Background, 4-5
<b>Methods</b>			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Methods, 5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	Methods, 5-6
	4b	Settings and locations where the data were collected	Methods, 5-6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	n/a
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Methods, 6-7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	Study protocol, Weltermann et al., Trials 2020
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Randomisation:

Sequence generation	8a	Method used to generate the random allocation sequence	Details in Study protocol, Weltermann et al., Trials 2020
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details in Study protocol, Weltermann et al., Trials 2020
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	n/a
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Details in Study protocol, Weltermann et al., Trials 2020
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	n/a
	11b	If relevant, description of the similarity of interventions	n/a
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Methods,7-8
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	n/a
<b>Results</b> 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	n/a
	13b	For each group, losses and exclusions after randomisation, together with reasons	n/a
Recruitment	14a	Dates defining the periods of recruitment and follow-up	Details in Study protocol, Weltermann et al., Trials 2020
	14b	Why the trial ended or was stopped	Details in Study protocol,

				Weltermann et al., Trials 2020
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group		Results, table 1, Page 8
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups		Results, 9
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)		Results, 9-11
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		N/A.
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		n/a
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)		n/a
<b>Discussion</b>				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		Discussion, Strengths and limitations, Page 12-13
Generalisability	21	Generalisability (external validity, applicability) of the trial findings		13
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Discussion, 11-13
<b>Other information</b>				
Registration	23	Registration number and name of trial registry		Trial registration, 2
Protocol	24	Where the full trial protocol can be accessed, if available		Study protocol, Weltermann et al., Trials 2020
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		Funding, 17-18



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

7 \*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  
8 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  
9 references relevant to this checklist, see [www.consort-statement.org](http://www.consort-statement.org).

For peer review only

# BMJ Open

## A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and staff assessments using data from the IMPROVEjob study

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2022-066298.R1
Article Type:	Original research
Date Submitted by the Author:	20-Dec-2022
Complete List of Authors:	Schmidt, Manuela; University Hospital Bonn, Institute of General Practice and Family Medicine Seifried-Dübon, Tanja; University Hospital Tuebingen Osianderstraße 5, Department of Psychosomatic Medicine and Psychotherapy Göbel, Julian; University Hospital Bonn, Institute of General Practice and Family Medicine Degen, Lukas; University Hospital Bonn, Institute of General Practice and Family Medicine Werners, Brigitte; Ruhr University Bochum Grot, Matthias; Ruhr University Bochum Rind, Esther; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Pieper, Claudia; University of Duisburg-Essen Jöckel, Karl-Heinz; Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg- Essen Linden, Karen; University Hospital Bonn, Institute of General Practice and Family Medicine Rieger, Monika A.; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Weltermann, Birgitta; University Hospital Bonn, Institute of General Practice and Family Medicine
<b>Primary Subject Heading</b>:	General practice / Family practice
Secondary Subject Heading:	General practice / Family practice
Keywords:	PRIMARY CARE, Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™  
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 1 **A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and**  
4 **staff assessments using data from the IMPROVE*job* study**  
5  
6  
7

8 4 **Authors:** Manuela Schmidt<sup>1\*</sup>, Tanja Seifried-Dübon<sup>2</sup>, Julian Göbel<sup>1</sup>, Lukas Degen<sup>1</sup>, Brigitte Werners<sup>3</sup>,  
9  
10 5 Matthias Grot<sup>3</sup>, Esther Rind<sup>4</sup>, Claudia Pieper<sup>5</sup>, Karl-Heinz Jöckel<sup>5,6</sup>, Karen Linden<sup>1</sup>, Monika A. Rieger<sup>4</sup>,  
11 6 Birgitta M. Weltermann<sup>1</sup> on behalf of the IMPROVE*job* consortium  
12  
13  
14 7

15  
16 8 <sup>1</sup> Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn,  
17 9 Venusberg-Campus 1, 53127 Bonn, Germany  
18  
19

20 10 <sup>2</sup> Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tuebingen,  
21 11 Osianderstraße 5, 72076 Tuebingen, Germany  
22  
23

24 12 <sup>3</sup> Institute of Management, Operations Research, Ruhr University Bochum, Universitätsstr. 150,  
25 13 44801 Bochum, Germany  
26  
27

28 14 <sup>4</sup> Institute of Occupational and Social Medicine and Health Services Research, University Hospital  
29 15 Tuebingen, Wilhelmstr. 27, 72074 Tuebingen, Germany  
30  
31

32 16 <sup>5</sup> Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University  
33 17 of Duisburg-Essen, Hufelandstr. 55, 45147 Essen, Germany  
34  
35

36 18 <sup>6</sup> Center for Clinical Trials, University Hospital Essen, University of Duisburg-Essen, Hufelandstr. 55,  
37 19 45147 Essen, Germany.  
38  
39

40 20  
41  
42 21 \*Correspondence to: [julian.goebel@ukbonn.de](mailto:julian.goebel@ukbonn.de)  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 22 **Abstract**  
4  
5 23

6 24 **Objectives:** Strong primary care leaders are needed to assure high quality services for patient  
7  
8 25 populations. This study analyzed general practitioners' (GP) leadership skills comparing practice-level  
9  
10 26 self- and staff assessments based on the Full Range of Leadership Model (FRLM) and the Leader-  
11  
12 27 Member Exchange (LMX).

13 28 **Setting:** The questionnaire survey was conducted among German general practice leaders and their  
14  
15 29 staff participating in the IMPROVE*job* trial.

16 30 **Participants:** The study population comprised 60 German general practices with 366 participants:  
17  
18 31 84 GP practice leaders and 282 employees (28 physicians and 254 practice assistants).

19  
20 32 **Primary and secondary outcome measures:** Leadership skills of the practice leaders were measured  
21  
22 33 using the Integrative Leadership Questionnaire (German *FIF*) and the Leader-Member Exchange (LMX-  
23  
24 34 7) questionnaire. Leaders rated themselves and practice staff rated their leaders. The data was  
25  
26 35 analyzed by paired mean comparisons on the practice level.

27 36 **Results:** For most leadership dimensions, practice leaders rated themselves higher than their  
28  
29 37 employees rated them. Differences were found for transformational leadership ( $p<.001$ ,  $d=.41$ ),  
30  
31 38 especially for the dimensions 'innovation' ( $p<.001$ ,  $d=.69$ ) and 'individuality focus' ( $p<.001$ ,  $d=.50$ ). For  
32  
33 39 transactional leadership, the dimension 'goal setting' differed significantly ( $p<.01$ ,  $d=.30$ ) but not the  
34  
35 40 other dimensions. Scores for negative leadership were low and showed no differences between  
36  
37 41 leaders and employees. Interestingly, employed physicians' rated their practice leaders higher on the  
38  
39 42 two transformational ('performance development', 'providing a vision') and all transactional  
40  
41 43 dimensions. The LMX-7 scale showed high quality relationships between leaders and employees.

42 44 **Conclusions:** This 180° analysis of GPs' leadership skills with self- and employee ratings indicated good  
43  
44 45 relationships. There is a potential to improve leadership regarding goal-setting, innovation and  
45  
46 46 focusing on individual team members. These results allow for the development of targeted  
47  
48 47 interventions.

49  
50 48 **Trial registration:** German Clinical Trials Register, DRKS00012677. Registered 16 October 2019.  
51  
52 49

53  
54 50 **Keywords:** leadership, leadership quality, general practitioner, practice staff, 180° feedback  
55  
56 51  
57 52  
58 53  
59 54  
60 55  
56

1  
2  
3 57 **Strengths & limitations:**  
4

- 5 58 • Leadership skills are crucial to secure a strong primary care workforce for patient care.  
6  
7 59 • High quality leadership can protect and promote the mental health of employees.  
8  
9 60 • This presentation of an innovative 180° feedback approach in the GP setting allows for a  
10 61 better understanding of leadership as viewed by the different professional groups working in  
11 62 practices.  
12  
13 63 • The data reflects a typical spectrum of German general practices including teaching and non-  
14 64 teaching practices as well as solo and group practices.  
15  
16 65 • In group practices, leadership teams, not individual leaders were rated by staff to capture  
17 66 leadership comprehensively on practice level.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 **67 Background:**

4  
5 68 Strong primary care leaders and a strong primary care workforce are important to assure the health of  
6  
7 69 populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al.  
8  
9 70 showed an association between stronger leadership and less burnout among different medical  
10  
11 71 professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job  
12  
13 72 satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed  
14  
15 73 correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-  
16  
17 74 day-mortality [11]. In addition to individual outcomes, leadership is important to promote  
18  
19 75 organizational changes (e.g., the implementation of IT-supported care) [12].  
20

21  
22 77 Scientifically, leadership is conceptualized in several theories. One of the most studied leadership  
23  
24 78 frameworks is the Full Range of Leadership Model (FRLM), which integrates transactional,  
25  
26 79 transformational, and negative leadership [13, 14]. Transactional leadership describes leaders'  
27  
28 80 structuring of work situations, the exchange of contingent rewards (e.g. work against salary), and the  
29  
30 81 management by exception [13–15]. In contrast, transformational leadership moves beyond leaders'  
31  
32 82 and staff's self-interests. It focuses on the staff's attitudes and values regarding overarching goals such  
33  
34 83 as self-actualization, organizational achievements, and the well-being of others and society as a whole  
35  
36 84 [13, 14]. Building on the FRLM, a recent further development, the so-called Implementation Leadership  
37  
38 85 Scale, focusses on the role of leadership for implementation of organizational changes [12]. Another  
39  
40 86 important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the  
41  
42 87 relationship between leaders and staff. It concentrates on the perceived quality of the dyadic  
43  
44 88 relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a  
45  
46 89 dyadic social exchange process ranging from low LMX, described by limited social transactions with  
47  
48 90 more transactional leadership to high LMX, which represents a transformational approach with a high  
49  
50 91 degree of social exchange and a mature leader-member partnership [17]. High-quality relationships  
51  
52 92 positively influence employees' work-related well-being and are associated with higher job satisfaction  
53  
54 93 of health care workers [18].  
55

56  
57 94  
58  
59 95 Based on these theories, various questionnaires were developed, e.g. the Leader-Member Exchange  
60  
61 96 questionnaire 'LMX-7' [16] and the German questionnaire 'Fragebogen für integrative Führung' (FIF; in  
62  
63 97 English: Questionnaire for Integrative Leadership) [15, 19]. These instruments allow for a multi-rater  
64  
65 98 perspective: the leader's and the staff's views on the leader's behavior are measured and compared  
66  
67 99 providing 180° feedback. This method is valuable because assessments from different perspectives  
68  
69 100 create a more comprehensive picture of the leaders' actual skills and performances [20]. Two recent  
70  
71 101 reviews of 60 studies from various medical settings showed that such approaches are increasingly

1  
2  
3 102 applied in medical education and graduate training [21, 22], but have not been used to evaluate GP  
4  
5 103 leaders and their teams. Effective interventions to improve leadership were developed and evaluated  
6  
7 104 in the hospital [23] and healthcare management setting [24]. For example, Saravo et al. showed an  
8  
9 105 improvement in transformational and transactional leadership performance of 57 medical residents in  
10  
11 106 hospital rotations after a 4-week intervention [23]. In addition, a 2018 study from Hill et al. highlighted  
12  
13 107 positive effects of a leadership training for surgical residents on teamwork and team involvement in  
14  
15 108 decision-making [25]. However, such interventions have not been implemented in German primary  
16  
17 109 care, although high chronic stress and burnout rates are reported for this workforce [26, 27]. The need  
18  
19 110 is even larger as about half of the German GPs who mainly work in GP-owned private practices [28],  
20  
21 111 will reach retirement age in the next ten years [29]. Based on the leadership frameworks mentioned  
22  
23 112 above, the publicly funded IMPROVE*job* study aimed to improve the job satisfaction of physician  
24  
25 113 leaders and practice personnel of German GP practices focusing on leadership, communication, and  
26  
27 114 work processes [30, 31]. At baseline, GPs' leadership skills were evaluated comparing GP leaders' self  
28  
29 115 and staff ratings on practice level.  
30  
31 116

## 32 117 **Methods**

33 118 This analysis draws on the baseline data of the IMPROVE*job* study, which is designed as a cluster-  
34  
35 119 randomized controlled trial (cRCT) to improve job satisfaction among practice personnel. The details  
36  
37 120 are described in the study protocol [30].

38 121 In short, a total of 60 GP practices in the North Rhine region in Germany were recruited by the Institute  
39  
40 122 of General Practice and Family Medicine of the University of Bonn. The sample comprised single  
41  
42 123 (owned by one practice leader) and group practices (owned by more than one practice leader), some  
43  
44 124 of which were also involved as teaching practices (affiliated to a university). The study aimed to recruit  
45  
46 125 practice teams, including physician leaders, employed physicians, and practice assistants. A total of 84  
47  
48 126 GP practice leaders, 28 employed physicians and 254 practice assistants were recruited. In Germany,  
49  
50 127 primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician,  
51  
52 128 practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years.  
53  
54 129 Similar to other regions worldwide, the size of group practices is increasing.  
55  
56 130

## 57 131 ***Patient and public involvement***

58 132 The study targeted general practice staff. Therefore, GPs and practice assistants were involved in all  
59  
60 133 phases of the study. As the study did not target patients, no patients or members of the public were  
61  
62 134 involved in the design or conduct of the study.  
63  
64 135

## 65 136 ***Ethics***



1  
2  
3 137 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn  
4  
5 138 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the  
6  
7 139 Medical Association North-Rhine (Lfd-Nr.: 2019107) and of the Medical Faculty, University Hospital of  
8  
9 140 Tuebingen (project no.: 446/2019BO2) approved the study protocol. The study was performed in  
10  
11 141 accordance with the Declaration of Helsinki. All participating practice team members received a study  
12  
13 142 information and signed informed consent forms.  
14

143

#### 144 **Measures**

145 Practice leaders answered a short questionnaire on practice characteristics and the questionnaire for  
146 practice leaders. Employed physicians and practice assistants completed different versions of the same  
147 employee questionnaire. Details of the methods and the characteristics of the study population are  
148 published [30, 31].

149 All participants provided sociodemographic, professional, and work-related characteristics which are  
150 published [31]. In addition, GP leaders and practice staff filled the following two leadership  
151 questionnaires:

##### 152 1. Integrative Leadership Questionnaire (FIF)

153 Transformational, transactional, and negative leadership were measured using the FIF questionnaire.  
154 Its scales' validity and internal consistency are confirmed for different populations [19, 32]. The FIF has  
155 been used in non-medical and hospital settings [33], but not in primary care.

156 All 40 items of the FIF are answered on a 5-point Likert scale and are worded to reflect either the  
157 leader's or the staff's position.

158 The measures comprise:

- 159 - the transformational leadership scale consisting of six dimensions: innovation, team spirit,  
160 performance development, individuality focus, providing a vision, and being a role model;
- 161 - the transactional leadership scale with two dimensions: goal setting and management by  
162 exception;
- 163 - the negative leadership scale with two dimensions: laissez-faire and destructive leadership.

##### 164 2. Leader-Member Exchange (LMX-7)

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

168 The multi-rater, 180° approach is applied to the two leadership scales. Results of such assessments are  
169 usually shared with the ratee, yet previous studies showed mixed reactions in the medical setting [36,  
170 37]. Therefore, the results of the 180° feedback in our study were not shared with the participating  
171 practices but are used on an aggregated level for research purposes only.

1  
2  
3 1724  
5 173 **Statistical analysis**

6 174 Statistical analyses were conducted using SPSS Statistics 27 (IBM Cooperation, Armonk, Ny, USA, 2020).

7  
8 175 All analyses were carried out at the participant and the practice level.9  
10 176 The FIF data were analyzed according to the official manual [19]. Mean scores for transformational,  
11 177 transactional, and negative leadership were summarized both for the respective main scale and all  
12 178 dimensions: for transactional and transformational leadership, they ranged from 1 (worst rating) to 5  
13 179 (best rating); for negative leadership, they ranged from 1 (best rating) to 5 (worst rating). For  
14 180 comparison, scores were standardized using T-scaling tables from reference populations as defined by  
15 181 Rowold & Poethke [19]. These T-values are based on a normal distribution around 50 (SD 10). Thus,  
16 182 values above 70 only reflect about 2% of the reference population from German-speaking countries  
17 183 [19].18 184 The LMX-7 was analyzed per standard protocol by creating a sum score of all seven items without  
19 185 transformation [35]. The LMX-7 score can range from 7 to 35 with five standard categories which were  
20 186 interpreted as follows: score 7 to 14 = very low; 15 to 19 = low; 20 to 24 = moderate; 25 to 29 = high,  
21 187 30 to 35 = very high [38]. Inadvertently, question seven was missing on all employed physicians'  
22 188 questionnaires, which reduced the answered questions to six. As the LMX-7 manual does not suggest  
23 189 a standard approach for missing values, we excluded employed physicians from the further analyses.24  
25  
26  
27  
28  
29  
30  
31  
32  
33 19034  
35 191 For the 180° feedback approach on practice level, the combined mean scores of employed physicians  
36 192 and practice assistants per practice were compared to the self-assessment of their respective leaders  
37 193 using paired t-tests, as the data satisfied the condition of a normal distribution with the Kolmogorov-  
38 194 Smirnov test. Cohen's d was applied to determine the effect size of mean comparisons with the  
39 195 following standard interpretations: small effects from  $d=.2$ , medium from  $d=.5$ , and high from  $d=.8$   
40 196 [39]. In single practices, the staff ratings were compared to the leader's assessment. In practices with  
41 197 more than one owner (group practices), each leader's self-rating was compared with the respective  
42 198 ratings of the practice personnel, who were asked to rate the leadership team of the practice, not  
43 199 stratified by individual leaders. This approach was chosen because practice owners of German  
44 200 practices typically work as a leadership team. In addition, the ratings of the transformational and  
45 201 transactional leadership scales were compared stratified by practice type (single vs. group and teaching  
46 202 vs. non-teaching practices) using the Kruskal-Wallis tests because the data for practice comparisons  
47 203 did not satisfy the conditions for parametric tests.48  
49  
50  
51  
52  
53  
54  
55  
56  
57 20458 205 **Results**59  
60 206 **Population**

207 The baseline data of the IMPROVE<sub>job</sub> study included 366 participants from 60 practices, consisting of  
 208 84 practice leaders, 28 employed physicians, and 254 practice assistants. The mean age of the  
 209 participants was 44.4 years, with a mean of 54.3 years for practice leaders, 44.8 for employed  
 210 physicians, and 41.0 for practice assistants. Among the practice assistants, 99.6% were female, as were  
 211 76.6% of the employed physicians and half of the practice leaders (52.4%). Most practice leaders  
 212 worked full-time (90.5%), as did about a quarter of the employed physicians (28.6%) and 41.5% of the  
 213 practice assistants (see Table 1). The details on the sociodemographic descriptions are published [31].

214

215 **Table 1: Sociodemographic description of participants at baseline [31]**

Variable	Total sample N=366	Practice leaders N=84	Employed physicians N=28	Practice assistants N=254
Female, %	87.1	52.4	78.6	99.6
Age in years, mean (SD)	44.4 (12.8)	54.3 (6.2)	44.8 (9.8)	41.0 (13.0)
Years in current practice, mean (SD)	10.0 (9.1)	15.3 (8.4)	3.9 (5.4)	8.8 (8.9)
Working full-time, %	52.0	90.5	28.6	41.5

216

217 On average, practice leaders had been accredited for 26.6 years and licensed for the statutory health  
 218 insurance for 16.4 years. Seven (25%) of the employed physicians were in GP training. Practice  
 219 assistants had graduated on average 19.9 years ago, while 7.5% were still in training. Of the  
 220 60 practices, 21 (35%) were single and 39 (65%) were group practices; of these, 34 were teaching (57%)  
 221 and 26 (43%) were non-teaching practices. On average, practices were in the same location for  
 222 20.4 years (SD 14 years).

223

224 **Leadership**

225 The transactional and transformational leadership scales showed a high internal consistency with  
 226 Cronbach's  $\alpha = .74$  to  $.93$  for the staff members' assessment and Cronbach's  $\alpha = .72$  to  $.87$  for the  
 227 leaders' assessment. For negative leadership, the scales showed a sufficient internal consistency for  
 228 staff members' (Cronbach's  $\alpha = .73$  to  $.80$ ) but not for leaders' assessments (Cronbach's  $\alpha = .47$  to  $.68$ ).  
 229 The mean results of the FIF were within the one standard deviation range of the reference population  
 230 [19]. Based on raw values, employed physicians rated their leaders consistently better than practice  
 231 assistants and better than the leaders themselves for some items. While practice assistants rated their  
 232 leaders more poorly than the practice leaders in raw values, reference T-values showed only minor  
 233 differences. The details are outlined in Table 2.

234

235 The LMX-7 scale showed an internal consistency of Cronbach's  $\alpha = .88$  for staff members (practice  
 236 assistants) and  $\alpha = .71$  for leaders. Both groups showed a high relationship quality, scoring 28 for

237 practice leaders and 26 for practice assistants. As the seventh question was missing for employed  
 238 physicians, they were excluded from the analysis. However, the sum score of the remaining six  
 239 questions also showed a high score of 24.9 out of 30. The details are shown in Table 2.

240

241 **Table 2: Leadership assessment by employment group:** main scales (in bold) and dimensions (LMX  
 242 values can range from 7 to 35, FIF scales from 1 to 5)

	Practice leaders (N=84)				Employed physicians (N=28)				Practice assistants (N=254)			
	M	SD	T*	n	M	SD	T*	N	M	SD	T*	n
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.6</b>	<b>45</b>	<b>84</b>	<b>3.9</b>	<b>0.7</b>	<b>56</b>	<b>27</b>	<b>3.5</b>	<b>0.8</b>	<b>52</b>	<b>237</b>
Innovation	4.2	0.6	49	84	4.0	1.0	55	28	3.7	0.9	52	247
Team spirit	4.1	0.7	49	84	3.8	1.0	54	28	3.6	1.1	52	251
Performance development	3.6	0.8	44	84	4.1	0.7	57	27	3.5	1.0	51	247
Individuality focus	3.9	0.7	47	84	3.7	1.0	54	28	3.5	1.1	53	249
Providing a vision	3.5	0.9	45	84	3.6	0.9	55	28	3.2	1.1	51	245
Being a role model	4.1	0.6	45	84	4.0	0.8	55	27	3.7	1.0	52	246
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.7</b>	<b>47</b>	<b>83</b>	<b>3.5</b>	<b>0.7</b>	<b>54</b>	<b>27</b>	<b>3.2</b>	<b>0.8</b>	<b>50</b>	<b>244</b>
Goal setting	3.5	0.7	44	83	3.7	0.9	56	27	3.1	1.0	50	246
Management by exception	3.3	0.8	51	83	3.4	0.8	52	27	3.3	0.9	51	245
<b>Negative leadership</b>	<b>1.5</b>	<b>0.5</b>	<b>51</b>	<b>83</b>	<b>1.5</b>	<b>0.6</b>	<b>45</b>	<b>28</b>	<b>1.7</b>	<b>0.7</b>	<b>47</b>	<b>248</b>
Laissez-faire	1.6	0.6	52	83	1.6	0.8	45	28	1.7	0.8	46	249
Destructive	1.4	0.5	51	83	1.4	0.6	46	28	1.6	0.7	48	248
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	-	<b>81</b>	<b>n/a</b>	<b>n/a</b>	-	<b>n/a</b>	<b>26.7</b>	<b>4.8</b>	-	<b>222</b>

243 \*Reference T-values range from 0 to 100, as defined by Rowold & Poethke 2017

244

### 245 **180° leadership feedback**

246 Practice leaders self-rated their leadership skills slightly better than their staff for all dimensions except  
 247 for 'management by exception'. There were no statistically significant differences for negative  
 248 leadership. For transactional leadership, goal setting differed significantly with a low effect size  
 249 ( $p=.009$ ,  $d=.30$ ). Leaders' scores on transformational leadership were significantly higher than the  
 250 scores of the teams, with the dimension for innovation reaching the strongest effect size ( $p<.001$ ,  
 251  $d=.69$ ), followed by individuality focus with a medium effect size ( $p<0.001$ ,  $d=.50$ ). The scores for team  
 252 spirit and being a role model were slightly lower, but significant. The main scale for transformational  
 253 leadership also showed a significant difference with a medium effect size ( $p<.001$ ,  $d=.41$ ). The details  
 254 are outlined in Table 3.

255

256 **Table 3: Comparison of leaders' self- and staff ratings (n=84 leader-team pairs):** main scales (in bold)  
 257 and dimensions

	Practice leaders		Practice staff		Paired t-test		
	M	SD	M	SD	t(df)	p	d
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.5</b>	<b>3.6</b>	<b>0.6</b>	<b>3.721(82)</b>	<b>&lt;.001</b>	<b>0.41</b>
Innovation	4.2	0.6	3.8	0.6	6.359(83)	<.001	0.69
Team spirit	4.1	0.7	3.8	0.7	3.462(82)	.001	0.38
Performance development	3.6	0.8	3.7	0.6	-0.208(83)	.836	-
Individuality focus	3.9	0.7	3.5	0.6	4.633(83)	<.001	0.50
Providing a vision	3.5	0.9	3.3	0.8	1.592(82)	.115	-
Being a role model	4.1	0.6	3.8	0.6	2.833(82)	.006	0.31
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.6</b>	<b>3.3</b>	<b>0.5</b>	<b>1.291(81)</b>	<b>.200</b>	<b>-</b>
Goal setting	3.5	0.7	3.2	0.6	2.681(81)	.009	0.30
Management by exception	3.3	0.8	3.4	0.6	-0.470(82)	.640	-
<b>Negative leadership</b>	<b>1.5</b>	<b>0.4</b>	<b>1.6</b>	<b>0.4</b>	<b>-1.744(82)</b>	<b>.085</b>	<b>-</b>
Laissez-faire	1.6	0.6	1.7	0.5	-1.563(82)	.122	-
Destructive	1.4	0.5	1.6	0.5	-1.514(82)	.134	-
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	<b>26.8</b>	<b>3.5</b>	<b>3.275(79)</b>	<b>.002</b>	<b>0.37</b>

258

### 259 *Transformational and transactional leadership by practice type*

260 The Kruskal-Wallis test was applied to analyze for differences in leadership by practice types. It showed  
 261 slight but non-significant differences in the raw values between practice types (single vs. group,  
 262 teaching vs. non-teaching practices), e.g., slightly higher ratings for transformational leadership in  
 263 single and non-teaching practices. These slight differences persisted when using reference T-values.  
 264 For details, see Table 4.

265

266 **Table 4: Comparison of leadership assessments by practice type: single versus group practices and**  
 267 **teaching versus non-teaching practices**

	Single (n=21)			Group (n=39)			Non-teaching (n=26)			Teaching (n=34)		
	M	T	n	M	T	N	M	T	n	M	T	n
<b>Practice leaders</b>												
Transformational	4.0	47	21	3.9	45	63	3.8	43	37	4.0	47	47
Transactional	3.4	47	21	3.4	47	62	3.3	45	37	3.4	47	46
Negative	1.5	51	21	1.5	51	62	1.6	53	37	1.5	51	46
LMX-7	28.8	-	20	27.9	-	61	27.5	-	36	28.6	-	45
<b>Practice staff</b>												
Transformational	3.7	54	67	3.6	53	212	3.5	52	117	3.6	53	162
Transactional	3.4	53	67	3.3	51	212	3.3	51	117	3.2	50	162
Negative	1.7	47	70	1.7	47	212	1.7	47	117	1.6	46	165
LMX-7	27.5	-	61	26.3	-	190	25.6	-	105	27.3	-	146

268

269

## 270 **Discussion**

1  
2  
3 271 Using a 180° feedback approach of leadership in GP practices, this study showed good relationships  
4  
5 272 between leaders and staff with low levels of negative leaderships. Practice staff rated their leaders  
6  
7 273 slightly higher on all transformational and transactional dimensions than the 234 German leaders and  
8  
9 274 713 employees from the FIF questionnaire reference population [19]. Also, agreement between GP  
10  
11 275 leaders and staff was higher than in a study of 1,137 German hospital employees (315 leaders, 822  
12  
13 276 staff members) from different occupational groups (e.g., physicians, nurses, administration,  
14  
15 277 information technology), which used the same methodology [33]. Interestingly, hospital and GP  
16  
17 278 leaders rated themselves approximately similar [33].

18  
19 279 The benefit of 180° and 360° feedback is shown in studies from various settings. In a sample of more  
20  
21 280 than 2,000 U.S. military leaders, 360° feedback (leaders, subordinates, peers) was identified as a good  
22  
23 281 predictor for promotions [40]. This is in line with a 180° feedback (leaders, employees) study among  
24  
25 282 396 managers from different departments of an international airline: congruence between managers  
26  
27 283 self-ratings and employees ratings predicted managerial behavior such as innovation, decision making,  
28  
29 284 leading, and motivation [41]. In a sample of 1,190 physicians from the U.S. and Canada, the 180°  
30  
31 285 feedback approach, which is also called multi-rater assessment, provided a more realistic picture of  
32  
33 286 leader-team situations as shown by an improvement in a leadership teamwork index [37]. In our study,  
34  
35 287 leadership ratings of employed physicians were markedly higher in most dimensions than those by  
36  
37 288 non-physician practice personnel. This likely reflects that employed physicians are much closer to their  
38  
39 289 physician leaders regarding training, roles, and duties compared to practice assistants. In addition,  
40  
41 290 practice assistants do not have the perspective to become physician leaders themselves, which implies  
42  
43 291 a fundamentally different perspective. This finding is in line with a 2010 review identifying several  
44  
45 292 studies which showed that staff members who perceive themselves as more similar to the leader give  
46  
47 293 better performance ratings [42]. This effect was shown for example among 406 rater and 396 ratees  
48  
49 294 in an insurance company [43].

50  
51 295 Multi-rater assessments can provide the basis for analyzing and at best improving the psychological  
52  
53 296 well-being at workplaces by a better mutual understanding of leaders and staff [7, 42, 44]. A 2016  
54  
55 297 study of 110 insurance managers and their teams showed higher job satisfaction with higher mutual  
56  
57 298 ratings. Job satisfaction among employees (assessed on a 1 to 5 scale) was lowest when leaders rated  
58  
59 299 their leadership skills higher than their subordinates did (Mean 3.89 of 5 compared to 4.53 of 5 in  
60  
300 agreement) [45]. Rowold and Poethke who developed the FIF questionnaire conclude from their  
301  
302 studies that leaders can learn to adapt when receiving the leadership ratings as feedback. In addition,  
303  
304 they recommend to implement for example regular team meetings and improving leadership skills  
305  
306 through training [19]. Results from the DIALHS collaboration from South Africa point at the need for  
307  
308 accountability strategies such as standard operation procedures, facility audits and target setting [46].  
309  
310 While other studies followed this approach to share the assessment results with the ratee, we

1  
2  
3 306 abstained from this because previous studies in the medical field showed mixed reactions. In a 2005  
4 307 study, 15 family physicians rated multisource feedback extremely different, from negative to positive.  
5 308 This evaluation was affected by the perceived usefulness, accuracy, and credibility [36].  
6  
7  
8 309 Using the LMX questionnaire, a 2008 study with 200 nurses from six smaller and larger hospitals  
9  
10 310 showed positive associations of high mutual relationship scores with enhanced commitment, reduced  
11 311 staff turnover, and better organizational behavior [47]. Also, positive effects on employees' health and  
12 312 well-being are described in association with good relationships between leaders and employees. Lower  
13 313 levels of emotional exhaustion were associated with higher leader-member exchange quality in a  
14 314 sample of 343 employees working in the German healthcare sector after 11 months [48]. In addition,  
15 315 a hierarchical regression model showed that the LMX was a good predictor for the health of 412  
16 316 employees in health and social services in Germany [49]. Compared to the LMX reference values based  
17 317 on 113 participants, our study showed an overall better relationship quality between practice leaders  
18 318 and practice assistants (mean value of 28.1 of 35 for practice leaders and 26.8 for practice assistants  
19 319 vs. 22.9 in the LMX reference population) [35]. Higher scores in the practice setting are likely influenced  
20 320 by the fact that GP leaders recruit personnel themselves, while personnel recruitment and placement  
21 321 in larger institutions is not necessarily in the hands of the direct team leaders.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

322

### 323 ***Strengths and limitations***

324 Novel for the German GP setting, we investigated GP leadership in a large sample with analysis on  
325 325 practice level. Our data provide leadership ratings for each solo practice leaders, but not for each group  
326 326 practice leader, as we had asked staff to rate their leadership team to reflect current small team  
327 327 leadership situations. LMX data were missing for one of seven questions for the small number of  
328 328 employed physicians. However, the analysis of the available data yielded a high relationship quality  
329 329 with leaders like the results for practice assistants. A selection bias cannot be excluded as participating  
330 330 practices might have had a greater interest in the topic.  
331

331

### 332 ***Conclusion and practical implications***

333 Overall, our data from the IMPROVE<sub>job</sub> study show trustful relationships between GP leaders and their  
334 334 staff. Future GPs' trainings should enable GP leaders to implement goal-setting, innovation, and  
335 335 individuality focus more effectively. Our results support recent calls for leadership workshops on every  
336 336 level of the medical training for strengthening the GP and other health services workforce.  
337

337

### 338 **Acknowledgment**

339 All authors are part of the IMPROVE*job* consortium and owe special thanks to all participating  
340 physicians and practice assistants. In addition, special thanks go to the other members of the  
341 IMPROVE*job* consortium; current members: M. Brinkmann, S. Burgess, A. Herrmann-Werner, K.-H.  
342 Jöckel, F. Junne, A. Siegel, F. Stuber, A. Wagner, and former members: J. Block, J.-M. Bois, A. Dreher,  
343 M. Hippler, S. Hartmann, L. Imhoff, C. Kersting, C. Ose and E. Tsarouha. Additionally, we are grateful to  
344 all members of the scientific advisory board: Stefanie Esper, Katja Götz, Andre Habrock, Albert  
345 Nienhaus, Horst Christian Vollmar, and Andrea Wittich. The work of the Institute of Occupational and  
346 Social Medicine and Health Services Research is supported by an unrestricted grant of the Employers'  
347 Association of the Metal and Electrical Industry Baden-Wuerttemberg (Suedwestmetall). We  
348 acknowledge administrative support by the DLR Project Management Agency.

349

### 350 **Contributorship statement**

351 Conceptualization, B.M.W., M.A.R., T.S.-D., C.P., and B.W.; methodology, B.M.W. M.A.R., T.S.-D., M.S.,  
352 B.W., M.G., E.R., C.P., K.-H.J., L.D., and J.G.; validation, K.-H.J.; formal analysis, M.S. and J.G.;  
353 investigation, L.D., K.L., T.S.-D., B.W., E.R., C.P., J.G., M.A.R., and B.M.W.; resources, B.M.W., L.D., K.L.,  
354 T.S.-D., B.W., M.G., C.P., and E.R.; data curation, K.-H.J.; writing—original draft preparation, M.S. and  
355 B.M.W.; writing—review and editing, M.S., T.S.-D., B.W., M.G., E.R., C.P., K.-H.J., L.D., J.G., M.A.R. and  
356 B.M.W.; visualization, M.S.; supervision, B.M.W., M.A.R., and B.W.; project administration, K.L., E.R.,  
357 B.M.W. and M.A.R.; funding acquisition, B.M.W., M.A.R., C.P., and B.W. All authors have read and  
358 approved the published version of the manuscript.

359

### 360 **Competing interests**

361 The authors declare that they have no competing interests.

362

### 363 **Funding**

364 This study is funded by the German Federal Ministry for Education and Research (grant numbers:  
365 01GL1851D, 01GL1751B, 01GL1751A, 01GL1751C). The funder had no role in the study design, the data  
366 collection, and analysis, the decision to publish, or the preparation of the manuscript.

367

### 368 **Data sharing statement**

369 There are no plans to grant access to the full protocol, participant-level dataset, or statistical code as  
370 data contain potentially identifying information.

### 371 **Ethics approval**

372 The study complies with the ethical principles of the World Medical Association Declaration of Helsinki.



1  
2  
3 373 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn  
4  
5 374 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the  
6  
7 375 Medical Association Nordrhein (Lfd-Nr.: 2019107), and of the Medical Faculty, University Hospital of  
8  
9 376 Tuebingen (project no.: 446/2019BO2) approved the study protocol.  
10  
11 377 All participating practice team members received written information and signed informed consent  
12  
13 378 forms, which are stored at the Institute for General Practice and Family Medicine, University of Bonn.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

## REFERENCES

- 1  
2  
3  
4 379
- 5  
6 380 1 Kringos D, Boerma W, Bourgueil Y, et al. The strength of primary care in Europe: an  
7 381 international comparative study. *British Journal of General Practice* 2013;63(616):e742-e750.  
8 382 2 Rohde J, Cousens S, Chopra M, et al. 30 years after Alma-Ata: has primary health care worked in  
9 383 countries? *The Lancet* 2008;372(9642):950–61.  
10 384 3 Swanwick T, Varnam R. Leadership development and primary care. *BMJ Leader* 2019;3(2).  
11 385 4 Meredith LS, Bouskill K, Chang J, et al. Predictors of burnout among US healthcare providers: a  
12 386 systematic review. *BMJ Open* 2022;12(8):e054243.  
13 387 5 Boamah SA, Spence Laschinger HK, Wong C, et al. Effect of transformational leadership on job  
14 388 satisfaction and patient safety outcomes. *Nurs Outlook* 2018;66(2):180–89.  
15 389 doi:10.1016/j.outlook.2017.10.004 [published Online First: 23 November 2017].  
16 390 6 Goetz K, Campbell S, Broge B, et al. Job satisfaction of practice assistants in general practice in  
17 391 Germany: an observational study. *Fam Pract* 2013;30(4):411–17. doi:10.1093/fampra/cmt015  
18 392 [published Online First: 19 April 2013].  
19 393 7 i Solà GJ, i Badia JG, Hito PD, et al. Self-perception of leadership styles and behaviour in primary  
20 394 health care. *BMC Health Services Research* 2016;16.  
21 395 8 Kersting C, Zimmer L, Thielmann A, et al. Chronic stress, work-related daily challenges and  
22 396 medicolegal investigations: a cross-sectional study among German general practitioners. *BMC*  
23 397 *Fam Pract* 2019;143. doi:10.1186/s12875-019-1032-6 [published Online First: 24 October  
24 398 2019].  
25 399 9 Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *The Lancet*  
26 400 2009;374(9702):1714–21.  
27 401 10 Wong CA, Cummings GG, Ducharme L. The relationship between nursing leadership and patient  
28 402 outcomes: a systematic review update. *J Nurs Manag* 2013;21(5):709–24.  
29 403 11 Sfantou DF, Laliotis A, Patelarou AE, et al. Importance of Leadership Style towards Quality of  
30 404 Care Measures in Healthcare Settings: A Systematic Review. *Healthcare (Basel, Switzerland)*  
31 405 2017;5(4):73.  
32 406 12 Söling S, Pfaff H, Karbach U, et al. How is leadership behavior associated with organization-  
33 407 related variables? Translation and psychometric evaluation of the implementation leadership  
34 408 scale in German primary healthcare. *BMC Health Services Research* 2022;22(1):1–13.  
35 409 13 Bass BM. Two Decades of Research and Development in Transformational Leadership.  
36 410 *European Journal of Work and Organizational Psychology* 1999;8(1):9–32.  
37 411 14 Poethke U. Leadership, values and communication: a cross-cultural investigation of the  
38 412 extended full-range of leadership behaviors: Doctoral dissertation 2019.  
39 413 15 Rowold J, Schlotz W. Transformational and transactional leadership and followers' chronic  
40 414 stress. *Leadership Review* 2009;9(1):35–48.  
41 415 16 Graen GB, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-  
42 416 member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-  
43 417 domain perspective. *Leadership Quarterly* 1995;6(2):219–47.  
44 418 17 Graen GB, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-  
45 419 member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-  
46 420 domain perspective. *The Leadership Quarterly* 1995;6(2):219–47.  
47 421 18 Gregersen S, Vincent-Höper S, Nienhaus A. Health-Relevant Leadership Behaviour: A  
48 422 Comparison of Leadership Constructs. *German Journal of Human Resource Management*  
49 423 2014;28(1-2):117–38.  
50 424 19 Rowold J, Poethke U. FIF: Fragebogen zur Integrativen Führung Manual, 1st edn. Göttingen:  
51 425 Hogrefe 2017.

- 1  
2  
3 426 20 Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof* 2003;23(1):4–12.
- 4 427  
5 428 21 Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with  
6 429 observed measures of competence: a systematic review. *JAMA* 2006;296(9):1094–102.
- 7 430 22 Donnon T, Al Ansari A, Al Alawi S, et al. The reliability, validity, and feasibility of multisource  
8 431 feedback physician assessment: a systematic review. *Acad Med* 2014;89(3):511–16.
- 9 432 23 Saravo B, Netzel J, Kiesewetter J. The need for strong clinical leaders - Transformational and  
10 433 transactional leadership as a framework for resident leadership training. *PLoS One*  
11 434 2017;12(8):e0183019.
- 12 435 24 Richter A, Lornudd C, Thiele Schwarz U von, et al. Evaluation of iLead, a generic implementation  
13 436 leadership intervention: mixed-method preintervention-postintervention design. *BMJ Open*  
14 437 2020;10(1):e033227. doi:10.1136/bmjopen-2019-033227 [published Online First: 12 January  
15 438 2020].
- 16 439 25 Hill DA, Jimenez J-C, Cohn SM, et al. How to be a leader: a course for residents. *Cureus*  
17 440 2018;10(7).
- 18 441 26 Viehmann A, Kersting C, Thielmann A, et al. Prevalence of chronic stress in general practitioners  
19 442 and practice assistants: Personal, practice and regional characteristics. *PLoS One*  
20 443 2017;12(5):e0176658.
- 21 444 27 Dreher A, Theune M, Kersting C, et al. Prevalence of burnout among German general  
22 445 practitioners: Comparison of physicians working in solo and group practices. *PLoS One*  
23 446 2019;14(2):e0211223.
- 24 447 28 Freund T, Everett C, Griffiths P, et al. Skill mix, roles and remuneration in the primary care  
25 448 workforce: who are the healthcare professionals in the primary care teams across the world?  
26 449 *International journal of nursing studies* 2015;52(3):727–43.
- 27 450 29 infas im Auftrag der KBV und des NAV-Virchow-Bund. Tabellenband Ärztemonitor 2018,  
28 451 Ergebnisse für Haus- und Fachärzte. Bonn 2018.
- 29 452 30 Weltermann BM, Kersting C, Pieper C, et al. IMPROVEjob - Participatory intervention to  
30 453 improve job satisfaction of general practice teams: a model for structural and behavioural  
31 454 prevention in small and medium-sized enterprises - a study protocol of a cluster-randomised  
32 455 controlled trial. *Trials* 2020;21(1):532.
- 33 456 31 Degen L, Linden K, Seifried-Dübön T, et al. Job Satisfaction and Chronic Stress of General  
34 457 Practitioners and Their Teams: Baseline Data of a Cluster-Randomised Trial (IMPROVEjob). *Int J*  
35 458 *Environ Res Public Health* 2021;18(18). doi:10.3390/ijerph18189458 [published Online First: 8  
36 459 September 2021].
- 37 460 32 Ingold PV. Fragebogen zur integrativen Führung. *Zeitschrift für Arbeits- und*  
38 461 *Organisationspsychologie A&O* 2019;63(2):100–06.
- 39 462 33 Stuber F, Seifried-Dübön T, Rieger MA, et al. Investigating the Role of Stress-Preventive  
40 463 Leadership in the Workplace Hospital: The Cross-Sectional Determination of Relational Quality  
41 464 by Transformational Leadership. *Front Psychiatry* 2019;10:622. doi:10.3389/fpsy.2019.00622  
42 465 [published Online First: 3 September 2019].
- 43 466 34 Schyns B. Überprüfung einer deutschsprachigen Skala zum Leader-Member-Exchange-Ansatz.  
44 467 *Zeitschrift für Differentielle und Diagnostische Psychologie* 2002.
- 45 468 35 Schyns B, Paul T. Leader-Member Exchange Skala (LMX7): ZIS - GESIS Leibniz Institute for the  
46 469 Social Sciences 2014.
- 47 470 36 Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback:  
48 471 perceptions of credibility and usefulness. *Med Educ* 2005;39(5):497–504.

- 1  
2  
3 472 37 Hu J, Lee R, Mullin S, et al. How physicians change: Multisource feedback driven intervention  
4 473 improves physician leadership and teamwork. *Surgery* 2020;168(4):714–23.  
5 474 doi:10.1016/j.surg.2020.06.008 [published Online First: 6 August 2020].  
6 475 38 Northouse PG. *Leadership: Theory and practice*, 6th edn. Thousand Oaks: Sage Publications  
7 476 2010.  
8 477 39 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hoboken: Taylor and  
9 478 Francis 1988.  
10 479 40 Halverson SK, Tonidandel S, Barlow C, et al. Self-other agreement on a 360-degree leadership  
11 480 evaluation. In: 17th Annual Conference of the Society for Industrial and Organizational  
12 481 Psychology, Toronto, Canada 2002.  
13 482 41 Furnham A, Stringfield P. Congruence of self and subordinate ratings of managerial practices as  
14 483 a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*  
15 484 1994;67(1):57–67.  
16 485 42 Fleenor JW, Smither JW, Atwater LE, et al. Self–other rating agreement in leadership: A review.  
17 486 *The Leadership Quarterly* 2010;21(6):1005–34.  
18 487 43 Antonioni D, Park H. The effects of personality similarity on peer ratings of contextual work  
19 488 behaviors. *Personnel Psychology* 2001;54(2):331–60.  
20 489 44 Alrubaysh MA, Alshehri MH, Alsuhaibani EA, et al. The leadership styles of primary healthcare  
21 490 center managers and center performance outcomes in Riyadh, Saudi Arabia: A correlational  
22 491 study. *Journal of Family & Community Medicine* 2022;29(1):56.  
23 492 45 Erben J, Schneider FM, Maier M. In the Ear of the Beholder: Self-Other Agreement in  
24 493 Leadership Communication and Its Relationship With Subordinates’ Job Satisfaction.  
25 494 *International Journal of Business Communication* 2019;56(4):505–29.  
26 495 46 Cleary S, Du Toit A, Scott V, et al. Enabling relational leadership in primary healthcare settings:  
27 496 lessons from the DIALHS collaboration. *Health Policy and Planning* 2018;33(suppl\_2):ii65–ii74.  
28 497 47 Chen C-HV, Wang S-J, Chang W-C, et al. The effect of leader-member exchange, trust,  
29 498 supervisor support on organizational citizenship behavior in nurses. *Journal of Nursing Research*  
30 499 2008;16(4):321–28.  
31 500 48 Gregersen S, Vincent-Höper S, Nienhaus A. Job-related resources, leader–member exchange  
32 501 and well-being – a longitudinal study. *Work & Stress* 2016;30(4):356–73.  
33 502 49 Vincent-Höper S, Stein M, Gregersen S, et al. Messung gesundheitsförderlicher Führung in  
34 503 Gesundheitsdienst und Wohlfahrtspflege. *ASU Arbeitsmed Sozialmed Umweltmed* 2018(53):46–  
35 504 53.  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

1 **Additional Files**

2 **Additional file 1: CONSORT 2010 checklist of information to include when reporting a randomised trial**

3

4  
5



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

Section/Topic	Item No	Checklist item	Reported on page No
<b>Title and abstract</b>			
	1a	Identification as a randomised trial in the title	n/a
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Abstract, 2
<b>Introduction</b>			
Background and objectives	2a	Scientific background and explanation of rationale	Background, 4-5
	2b	Specific objectives or hypotheses	Background, 4-5
<b>Methods</b>			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Methods, 5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	Methods, 5-6
	4b	Settings and locations where the data were collected	Methods, 5-6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	n/a
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Methods, 6-7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	Study protocol, Weltermann et al., Trials 2020
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a

## Randomisation:

Sequence generation	8a	Method used to generate the random allocation sequence	Details in Study protocol, Weltermann et al., Trials 2020
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details in Study protocol, Weltermann et al., Trials 2020
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	n/a
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Details in Study protocol, Weltermann et al., Trials 2020
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	n/a
	11b	If relevant, description of the similarity of interventions	n/a
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Methods,7-8
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	n/a
<b>Results</b>			
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	n/a
	13b	For each group, losses and exclusions after randomisation, together with reasons	n/a
Recruitment	14a	Dates defining the periods of recruitment and follow-up	Details in Study protocol, Weltermann et al., Trials 2020
	14b	Why the trial ended or was stopped	Details in Study protocol,

				Weltermann et al., Trials 2020
	Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Results, table 1, Page 8
	Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Results, 9
	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)	Results, 9-11
		17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	N/A.
	Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	n/a
	Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	n/a
	<b>Discussion</b>			
	Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	Discussion, Strengths and limitations, Page 12-13
	Generalisability	21	Generalisability (external validity, applicability) of the trial findings	13
	Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Discussion, 11-13
	<b>Other information</b>			
	Registration	23	Registration number and name of trial registry	Trial registration, 2
	Protocol	24	Where the full trial protocol can be accessed, if available	Study protocol, Weltermann et al., Trials 2020
	Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	Funding, 17-18

1  
2  
3 7 \*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  
4 8 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  
5 9 references relevant to this checklist, see [www.consort-statement.org](http://www.consort-statement.org).  
6 10  
7 11  
8 12

For peer review only



# BMJ Open

**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**

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2022-066298.R2
Article Type:	Original research
Date Submitted by the Author:	17-Jul-2023
Complete List of Authors:	Schmidt, Manuela; University Hospital Bonn, Institute of General Practice and Family Medicine Seifried-Dübon, Tanja; University Hospital Tuebingen Osianderstraße 5, Department of Psychosomatic Medicine and Psychotherapy Göbel, Julian; University Hospital Bonn, Institute of General Practice and Family Medicine Degen, Lukas; University Hospital Bonn, Institute of General Practice and Family Medicine Werners, Brigitte; Ruhr University Bochum Grot, Matthias; Ruhr University Bochum Rind, Esther; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Pieper, Claudia; University of Duisburg-Essen Jöckel, Karl-Heinz; Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg- Essen Linden, Karen; University Hospital Bonn, Institute of General Practice and Family Medicine Rieger, Monika A.; Institute for Occupational Medicine, Social Medicine and Health Services Research , University Hospital Tuebingen Weltermann, Birgitta; University Hospital Bonn, Institute of General Practice and Family Medicine
<b>Primary Subject Heading</b>:	General practice / Family practice
Secondary Subject Heading:	General practice / Family practice
Keywords:	PRIMARY CARE, Human resource management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

1  
2  
3 1 **A 180° view on general practitioners' leadership skills: practice-level comparisons of leader and**  
4 **staff assessments using data from the cluster-randomized controlled IMPROVE*job* study**  
5 2  
6 3

8 4 **Authors:** Manuela Schmidt<sup>1\*</sup>, Tanja Seifried-Dübon<sup>2</sup>, Julian Göbel<sup>1</sup>, Lukas Degen<sup>1</sup>, Brigitte Werners<sup>3</sup>,  
9  
10 5 Matthias Grot<sup>3</sup>, Esther Rind<sup>4</sup>, Claudia Pieper<sup>5</sup>, Karl-Heinz Jöckel<sup>5,6</sup>, Karen Linden<sup>1</sup>, Monika A. Rieger<sup>4</sup>,  
11 6 Birgitta M. Weltermann<sup>1</sup> on behalf of the IMPROVE*job* consortium  
12  
13  
14 7

16 8 <sup>1</sup> Institute of General Practice and Family Medicine, University Hospital Bonn, University of Bonn,  
17 9 Venusberg-Campus 1, 53127 Bonn, Germany

20 10 <sup>2</sup> Department of Psychosomatic Medicine and Psychotherapy, University Hospital Tuebingen,  
21 11 Osianderstraße 5, 72076 Tuebingen, Germany

24 12 <sup>3</sup> Institute of Management, Operations Research, Ruhr University Bochum, Universitätsstr. 150,  
25 13 44801 Bochum, Germany

28 14 <sup>4</sup> Institute of Occupational and Social Medicine and Health Services Research, University Hospital  
29 15 Tuebingen, Wilhelmstr. 27, 72074 Tuebingen, Germany

32 16 <sup>5</sup> Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University  
33 17 of Duisburg-Essen, Hufelandstr. 55, 45147 Essen, Germany

36 18 <sup>6</sup> Center for Clinical Trials, University Hospital Essen, University of Duisburg-Essen, Hufelandstr. 55,  
37 19 45147 Essen, Germany.  
38  
39  
40 20

42 21 \*Correspondence to: [julian.goebel@ukbonn.de](mailto:julian.goebel@ukbonn.de)  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Abstract

**Objectives:** 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).

**Setting:** The questionnaire survey was conducted among German general practice leaders and their staff participating in the IMPROVE*job* trial.

**Participants:** The study population comprised 60 German general practices with 366 participants: 84 GP practice leaders and 282 employees (28 physicians and 254 practice assistants).

**Primary and secondary outcome measures:** 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.

**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.

**Trial registration:** German Clinical Trials Register, DRKS00012677. Registered 16 October 2019.

**Keywords:** leadership, leadership quality, general practitioner, practice staff, 180° feedback

1  
2  
3 57 **Strengths & limitations:**  
4

- 5 58 • This 180° feedback approach in the GP setting allows for a better understanding of  
6 59 leadership from the perspective of different practice professionals.  
7  
8 60 • The data reflect a typical spectrum of German general practices with solo and group  
9 61 practices, but results may differ in other settings.  
10  
11 62 • Leadership teams, not individual leaders, were rated in group practices to capture leadership  
12 63 at the practice level.  
13  
14  
15 64  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

1  
2  
3 **65 Background:**  
4

5 **66** Strong primary care leaders and a strong primary care workforce are important to assure the health of  
6 **67** populations and primary care teams [1–3]. A recent systematic review of 20 studies by Meredith et al.  
7 **68** showed an association between stronger leadership and less burnout among different medical  
8 **69** professionals in the United States [4]. In contrast, poor leadership skills have a negative impact on job  
9 **70** satisfaction [5–7], staff well-being [8] and the quality of patient care [5, 9, 10]. A review showed  
10 **71** correlations between better leadership and various quality of care indicators, e.g. pain, safety and 30-  
11 **72** day-mortality [11]. In addition to individual outcomes, leadership is important to promote  
12 **73** organizational changes (e.g., the implementation of IT-supported care) [12].  
13  
14  
15  
16  
17  
18  
19

20 **75** Scientifically, leadership is conceptualized in several theories. One of the most studied leadership  
21 **76** frameworks is the Full Range of Leadership Model (FRLM), which integrates transactional,  
22 **77** transformational, and negative leadership [13, 14]. Transactional leadership describes leaders'  
23 **78** structuring of work situations, the exchange of contingent rewards (e.g. work against salary), and the  
24 **79** management by exception [13–15]. In contrast, transformational leadership moves beyond leaders'  
25 **80** and staff's self-interests. It focuses on the staff's attitudes and values regarding overarching goals such  
26 **81** as self-actualization, organizational achievements, and the well-being of others and society as a whole  
27 **82** [13, 14]. Building on the FRLM, a recent further development, the so-called Implementation Leadership  
28 **83** Scale, focusses on the role of leadership for implementation of organizational changes [12]. Another  
29 **84** important leadership theory, the Leader-Member Exchange (LMX), specifically addresses the  
30 **85** relationship between leaders and staff. It concentrates on the perceived quality of the dyadic  
31 **86** relationship between a staff member and the immediate leader [13, 16]. The relationship reflects a  
32 **87** dyadic social exchange process ranging from low LMX, described by limited social transactions with  
33 **88** more transactional leadership to high LMX, which represents a transformational approach with a high  
34 **89** degree of social exchange and a mature leader-member partnership [17]. High-quality relationships  
35 **90** positively influence employees' work-related well-being and are associated with higher job satisfaction  
36 **91** of health care workers [18].  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

50 **93** Based on these theories, various questionnaires were developed, e.g. the Leader-Member Exchange  
51 **94** questionnaire 'LMX-7' [16] and the German questionnaire 'Fragebogen für integrative Führung' (FIF; in  
52 **95** English: Questionnaire for Integrative Leadership) [15, 19]. These instruments allow for a multi-rater  
53 **96** perspective: the leader's and the staff's views on the leader's behavior are measured and compared  
54 **97** providing 180° feedback. This method is valuable because assessments from different perspectives  
55 **98** create a more comprehensive picture of the leaders' actual skills and performances [20]. Two recent  
56 **99** reviews of 60 studies from various medical settings showed that such approaches are increasingly  
57  
58  
59  
60

1  
2  
3 100 applied in medical education and graduate training [21, 22], but have not been used to evaluate GP  
4  
5 101 leaders and their teams. Effective interventions to improve leadership were developed and evaluated  
6  
7 102 in the hospital [23] and healthcare management setting [24]. For example, Saravo et al. showed an  
8  
9 103 improvement in transformational and transactional leadership performance of 57 medical residents in  
10  
11 104 hospital rotations after a 4-week intervention [23]. In addition, a 2018 study from Hill et al. highlighted  
12  
13 105 positive effects of a leadership training for surgical residents on teamwork and team involvement in  
14  
15 106 decision-making [25]. However, such interventions have not been implemented in German primary  
16  
17 107 care, although high chronic stress and burnout rates are reported for this workforce [26, 27]. The need  
18  
19 108 is even larger as about half of the German GPs who mainly work in GP-owned private practices [28],  
20  
21 109 will reach retirement age in the next ten years [29]. Based on the leadership frameworks mentioned  
22  
23 110 above, the publicly funded IMPROVE*job* study aimed to improve the job satisfaction of physician  
24  
25 111 leaders and practice personnel of German GP practices focusing on leadership, communication, and  
26  
27 112 work processes [30, 31]. At baseline, GPs' leadership skills were evaluated comparing GP leaders' self  
28  
29 113 and staff ratings on practice level.  
30  
31 114

## 32 115 **Methods**

33 116 This analysis draws on the baseline data of the IMPROVE*job* study, which is designed as a cluster-  
34  
35 117 randomized controlled trial (cRCT) to improve job satisfaction among practice personnel. The details  
36  
37 118 are described in the study protocol [30].

38 119 In short, a total of 60 GP practices in the North Rhine region in Germany were recruited by the Institute  
39  
40 120 of General Practice and Family Medicine of the University of Bonn. The sample comprised single  
41  
42 121 (owned by one practice leader) and group practices (owned by more than one practice leader), some  
43  
44 122 of which were also involved as teaching practices (affiliated to a university). The study aimed to recruit  
45  
46 123 practice teams, including physician leaders, employed physicians, and practice assistants. A total of 84  
47  
48 124 GP practice leaders, 28 employed physicians and 254 practice assistants were recruited. In Germany,  
49  
50 125 primary care is typically provided by GP-owned practices with 1 to 3 physicians. For each physician,  
51  
52 126 practices employ about 1 to 2 certified practice assistants who finished a vocational training of 3 years.  
53  
54 127 Similar to other regions worldwide, the size of group practices is increasing.  
55  
56 128

## 57 129 ***Patient and public involvement***

58 130 The study did not target patients, but general practice personnel. Therefore, no patients or members  
59  
60 131 of the public were involved.  
61  
62 132

## 63 133 ***Ethics***



1  
2  
3 134 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn  
4  
5 135 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the  
6  
7 136 Medical Association North-Rhine (Lfd-Nr.: 2019107) and of the Medical Faculty, University Hospital of  
8  
9 137 Tuebingen (project no.: 446/2019BO2) approved the study protocol. The study was performed in  
10  
11 138 accordance with the Declaration of Helsinki. All participating practice team members received a study  
12  
13 139 information and signed informed consent forms.  
14

140

### 141 **Measures**

142 Practice leaders answered a short questionnaire on practice characteristics and the questionnaire for  
143 practice leaders. Employed physicians and practice assistants completed different versions of the same  
144 employee questionnaire. Details of the methods and the characteristics of the study population are  
145 published [30, 31].

146 All participants provided sociodemographic, professional, and work-related characteristics which are  
147 published [31]. In addition, GP leaders and practice staff filled the following two leadership  
148 questionnaires:

#### 149 1. Integrative Leadership Questionnaire (FIF)

150 Transformational, transactional, and negative leadership were measured using the FIF questionnaire.  
151 Its scales' validity and internal consistency are confirmed for different populations [19, 32]. The FIF has  
152 been used in non-medical and hospital settings [33], but not in primary care.

153 All 40 items of the FIF are answered on a 5-point Likert scale and are worded to reflect either the  
154 leader's or the staff's position.

155 The measures comprise:

- 156 - the transformational leadership scale consisting of six dimensions: innovation, team spirit,  
157 performance development, individuality focus, providing a vision, and being a role model;
- 158 - the transactional leadership scale with two dimensions: goal setting and management by  
159 exception;
- 160 - the negative leadership scale with two dimensions: laissez-faire and destructive leadership.

#### 161 2. Leader-Member Exchange (LMX-7)

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

165 The multi-rater, 180° approach is applied to the two leadership scales. Results of such assessments are  
166 usually shared with the ratee, yet previous studies showed mixed reactions in the medical setting [36,  
167 37]. Therefore, the results of the 180° feedback in our study were not shared with the participating  
168 practices but are used on an aggregated level for research purposes only.

169

**Statistical analysis**

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

All analyses were carried out at the participant and the practice level.

The FIF data were analyzed according to the official manual [19]. Mean scores for transformational, transactional, and negative leadership were summarized both for the respective main scale and all dimensions: for transactional and transformational leadership, they ranged from 1 (worst rating) to 5 (best rating); for negative leadership, they ranged from 1 (best rating) to 5 (worst rating). For comparison, scores were standardized using T-scaling tables from reference populations as defined by Rowold & Poethke [19]. These T-values are based on a normal distribution around 50 (SD 10). Thus, values above 70 only reflect about 2% of the reference population from German-speaking countries [19].

The LMX-7 was analyzed per standard protocol by creating a sum score of all seven items without transformation [35]. The LMX-7 score can range from 7 to 35 with five standard categories which were interpreted as follows: score 7 to 14 = very low; 15 to 19 = low; 20 to 24 = moderate; 25 to 29 = high, 30 to 35 = very high [38]. Inadvertently, question seven was missing on all employed physicians' questionnaires, which reduced the answered questions to six. As the LMX-7 manual does not suggest a standard approach for missing values, we excluded employed physicians from the further analyses.

For the 180° feedback approach on practice level, the combined mean scores of employed physicians and practice assistants per practice were compared to the self-assessment of their respective leaders using paired t-tests, as the data satisfied the condition of a normal distribution with the Kolmogorov-Smirnov test. Cohen's d was applied to determine the effect size of mean comparisons with the following standard interpretations: small effects from  $d=.2$ , medium from  $d=.5$ , and high from  $d=.8$  [39]. In single practices, the staff ratings were compared to the leader's assessment. 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. In addition, the ratings of the transformational and transactional leadership scales were compared stratified by practice type (single vs. group and teaching vs. non-teaching practices) using the Kruskal-Wallis tests because the data for practice comparisons did not satisfy the conditions for parametric tests.

**Results****Population**

204 The baseline data of the IMPROVE<sub>job</sub> study included 366 participants from 60 practices, consisting of  
 205 84 practice leaders, 28 employed physicians, and 254 practice assistants. The mean age of the  
 206 participants was 44.4 years, with a mean of 54.3 years for practice leaders, 44.8 for employed  
 207 physicians, and 41.0 for practice assistants. Among the practice assistants, 99.6% were female, as were  
 208 76.6% of the employed physicians and half of the practice leaders (52.4%). Most practice leaders  
 209 worked full-time (90.5%), as did about a quarter of the employed physicians (28.6%) and 41.5% of the  
 210 practice assistants (see Table 1). The details on the sociodemographic descriptions are published [31].

211

212 **Table 1: Sociodemographic description of participants at baseline [31]**

Variable	Total sample N=366	Practice leaders N=84	Employed physicians N=28	Practice assistants N=254
Female, %	87.1	52.4	78.6	99.6
Age in years, mean (SD)	44.4 (12.8)	54.3 (6.2)	44.8 (9.8)	41.0 (13.0)
Years in current practice, mean (SD)	10.0 (9.1)	15.3 (8.4)	3.9 (5.4)	8.8 (8.9)
Working full-time, %	52.0	90.5	28.6	41.5

213

214 On average, practice leaders had been accredited for 26.6 years and licensed for the statutory health  
 215 insurance for 16.4 years. Seven (25%) of the employed physicians were in GP training. Practice  
 216 assistants had graduated on average 19.9 years ago, while 7.5% were still in training. Of the  
 217 60 practices, 21 (35%) were single and 39 (65%) were group practices; of these, 34 were teaching (57%)  
 218 and 26 (43%) were non-teaching practices. On average, practices were in the same location for  
 219 20.4 years (SD 14 years).

220

221 **Leadership**

222 The transactional and transformational leadership scales showed a high internal consistency with  
 223 Cronbach's  $\alpha = .74$  to  $.93$  for the staff members' assessment and Cronbach's  $\alpha = .72$  to  $.87$  for the  
 224 leaders' assessment. For negative leadership, the scales showed a sufficient internal consistency for  
 225 staff members' (Cronbach's  $\alpha = .73$  to  $.80$ ) but not for leaders' assessments (Cronbach's  $\alpha = .47$  to  $.68$ ).  
 226 The mean results of the FIF were within the one standard deviation range of the reference population  
 227 [19]. Based on raw values, employed physicians rated their leaders consistently better than practice  
 228 assistants and better than the leaders themselves for some items. While practice assistants rated their  
 229 leaders more poorly than the practice leaders in raw values, reference T-values showed only minor  
 230 differences. The details are outlined in Table 2.

231

232 The LMX-7 scale showed an internal consistency of Cronbach's  $\alpha = .88$  for staff members (practice  
 233 assistants) and  $\alpha = .71$  for leaders. Both groups showed a high relationship quality, scoring 28 for

234 practice leaders and 26 for practice assistants. As the seventh question was missing for employed  
 235 physicians, they were excluded from the analysis. However, the sum score of the remaining six  
 236 questions also showed a high score of 24.9 out of 30. The details are shown in Table 2.

237  
 238 **Table 2: Leadership assessment by employment group:** main scales (in bold) and dimensions (LMX  
 239 values can range from 7 to 35, FIF scales from 1 to 5)

	Practice leaders (N=84)				Employed physicians (N=28)				Practice assistants (N=254)			
	M	SD	T*	n	M	SD	T*	N	M	SD	T*	n
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.6</b>	<b>45</b>	<b>84</b>	<b>3.9</b>	<b>0.7</b>	<b>56</b>	<b>27</b>	<b>3.5</b>	<b>0.8</b>	<b>52</b>	<b>237</b>
Innovation	4.2	0.6	49	84	4.0	1.0	55	28	3.7	0.9	52	247
Team spirit	4.1	0.7	49	84	3.8	1.0	54	28	3.6	1.1	52	251
Performance development	3.6	0.8	44	84	4.1	0.7	57	27	3.5	1.0	51	247
Individuality focus	3.9	0.7	47	84	3.7	1.0	54	28	3.5	1.1	53	249
Providing a vision	3.5	0.9	45	84	3.6	0.9	55	28	3.2	1.1	51	245
Being a role model	4.1	0.6	45	84	4.0	0.8	55	27	3.7	1.0	52	246
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.7</b>	<b>47</b>	<b>83</b>	<b>3.5</b>	<b>0.7</b>	<b>54</b>	<b>27</b>	<b>3.2</b>	<b>0.8</b>	<b>50</b>	<b>244</b>
Goal setting	3.5	0.7	44	83	3.7	0.9	56	27	3.1	1.0	50	246
Management by exception	3.3	0.8	51	83	3.4	0.8	52	27	3.3	0.9	51	245
<b>Negative leadership</b>	<b>1.5</b>	<b>0.5</b>	<b>51</b>	<b>83</b>	<b>1.5</b>	<b>0.6</b>	<b>45</b>	<b>28</b>	<b>1.7</b>	<b>0.7</b>	<b>47</b>	<b>248</b>
Laissez-faire	1.6	0.6	52	83	1.6	0.8	45	28	1.7	0.8	46	249
Destructive	1.4	0.5	51	83	1.4	0.6	46	28	1.6	0.7	48	248
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	-	<b>81</b>	<b>n/a</b>	<b>n/a</b>	-	<b>n/a</b>	<b>26.7</b>	<b>4.8</b>	-	<b>222</b>

240 \*Reference T-values range from 0 to 100, as defined by Rowold & Poethke 2017

### 241 242 **180° leadership feedback**

243 Practice leaders self-rated their leadership skills slightly better than their staff for all dimensions except  
 244 for 'management by exception'. There were no statistically significant differences for negative  
 245 leadership. For transactional leadership, goal setting differed significantly with a low effect size  
 246 ( $p=.009$ ,  $d=.30$ ). Leaders' scores on transformational leadership were significantly higher than the  
 247 scores of the teams, with the dimension for innovation reaching the strongest effect size ( $p<.001$ ,  
 248  $d=.69$ ), followed by individuality focus with a medium effect size ( $p<0.001$ ,  $d=.50$ ). The scores for team  
 249 spirit and being a role model were slightly lower, but significant. The main scale for transformational  
 250 leadership also showed a significant difference with a medium effect size ( $p<.001$ ,  $d=.41$ ). The details  
 251 are outlined in Table 3.

252  
 253 **Table 3: Comparison of leaders' self- and staff ratings (n=84 leader-team pairs):** main scales (in bold)  
 254 and dimensions

	Practice leaders		Practice staff		Paired t-test		
	M	SD	M	SD	t(df)	p	d
<b>Transformational leadership</b>	<b>3.9</b>	<b>0.5</b>	<b>3.6</b>	<b>0.6</b>	<b>3.721(82)</b>	<b>&lt;.001</b>	<b>0.41</b>
Innovation	4.2	0.6	3.8	0.6	6.359(83)	<.001	0.69
Team spirit	4.1	0.7	3.8	0.7	3.462(82)	.001	0.38
Performance development	3.6	0.8	3.7	0.6	-0.208(83)	.836	-
Individuality focus	3.9	0.7	3.5	0.6	4.633(83)	<.001	0.50
Providing a vision	3.5	0.9	3.3	0.8	1.592(82)	.115	-
Being a role model	4.1	0.6	3.8	0.6	2.833(82)	.006	0.31
<b>Transactional leadership</b>	<b>3.4</b>	<b>0.6</b>	<b>3.3</b>	<b>0.5</b>	<b>1.291(81)</b>	<b>.200</b>	<b>-</b>
Goal setting	3.5	0.7	3.2	0.6	2.681(81)	.009	0.30
Management by exception	3.3	0.8	3.4	0.6	-0.470(82)	.640	-
<b>Negative leadership</b>	<b>1.5</b>	<b>0.4</b>	<b>1.6</b>	<b>0.4</b>	<b>-1.744(82)</b>	<b>.085</b>	<b>-</b>
Laissez-faire	1.6	0.6	1.7	0.5	-1.563(82)	.122	-
Destructive	1.4	0.5	1.6	0.5	-1.514(82)	.134	-
<b>LMX-7</b>	<b>28.1</b>	<b>2.6</b>	<b>26.8</b>	<b>3.5</b>	<b>3.275(79)</b>	<b>.002</b>	<b>0.37</b>

255

### 256 *Transformational and transactional leadership by practice type*

257 The Kruskal-Wallis test was applied to analyze for differences in leadership by practice types. It showed  
 258 slight but non-significant differences in the raw values between practice types (single vs. group,  
 259 teaching vs. non-teaching practices), e.g., slightly higher ratings for transformational leadership in  
 260 single and non-teaching practices. These slight differences persisted when using reference T-values.  
 261 For details, see Table 4.

262

263 **Table 4: Comparison of leadership assessments by practice type: single versus group practices and**  
 264 **teaching versus non-teaching practices**

	Single (n=21)			Group (n=39)			Non-teaching (n=26)			Teaching (n=34)		
	M	T	n	M	T	N	M	T	n	M	T	n
<b>Practice leaders</b>												
Transformational	4.0	47	21	3.9	45	63	3.8	43	37	4.0	47	47
Transactional	3.4	47	21	3.4	47	62	3.3	45	37	3.4	47	46
Negative	1.5	51	21	1.5	51	62	1.6	53	37	1.5	51	46
LMX-7	28.8	-	20	27.9	-	61	27.5	-	36	28.6	-	45
<b>Practice staff</b>												
Transformational	3.7	54	67	3.6	53	212	3.5	52	117	3.6	53	162
Transactional	3.4	53	67	3.3	51	212	3.3	51	117	3.2	50	162
Negative	1.7	47	70	1.7	47	212	1.7	47	117	1.6	46	165
LMX-7	27.5	-	61	26.3	-	190	25.6	-	105	27.3	-	146

265

266

## 267 **Discussion**

268  
269  
270

1  
2  
3 268 Using a 180° feedback approach of leadership in GP practices, this study showed good relationships  
4  
5 269 between leaders and staff with low levels of negative leaderships. Practice staff rated their leaders  
6  
7 270 slightly higher on all transformational and transactional dimensions than the 234 German leaders and  
8  
9 271 713 employees from the FIF questionnaire reference population [19]. Also, agreement between GP  
10  
11 272 leaders and staff was higher than in a study of 1,137 German hospital employees (315 leaders, 822  
12  
13 273 staff members) from different occupational groups (e.g., physicians, nurses, administration,  
14  
15 274 information technology), which used the same methodology [33]. Interestingly, hospital and GP  
16  
17 275 leaders rated themselves approximately similar [33].

16  
17 276 The benefit of 180° and 360° feedback is shown in studies from various settings. In a sample of more  
18  
19 277 than 2,000 U.S. military leaders, 360° feedback (leaders, subordinates, peers) was identified as a good  
20  
21 278 predictor for promotions [40]. This is in line with a 180° feedback (leaders, employees) study among  
22  
23 279 396 managers from different departments of an international airline: congruence between managers  
24  
25 280 self-ratings and employees ratings predicted managerial behavior such as innovation, decision making,  
26  
27 281 leading, and motivation [41]. In a sample of 1,190 physicians from the U.S. and Canada, the 180°  
28  
29 282 feedback approach, which is also called multi-rater assessment, provided a more realistic picture of  
30  
31 283 leader-team situations as shown by an improvement in a leadership teamwork index [37]. In our study,  
32  
33 284 leadership ratings of employed physicians were markedly higher in most dimensions than those by  
34  
35 285 non-physician practice personnel. This likely reflects that employed physicians are much closer to their  
36  
37 286 physician leaders regarding training, roles, and duties compared to practice assistants. In addition,  
38  
39 287 practice assistants do not have the perspective to become physician leaders themselves, which implies  
40  
41 288 a fundamentally different perspective. This finding is in line with a 2010 review identifying several  
42  
43 289 studies which showed that staff members who perceive themselves as more similar to the leader give  
44  
45 290 better performance ratings [42]. This effect was shown for example among 406 rater and 396 ratees  
46  
47 291 in an insurance company [43].

43  
44 292 Multi-rater assessments can provide the basis for analyzing and at best improving the psychological  
45  
46 293 well-being at workplaces by a better mutual understanding of leaders and staff [7, 42, 44]. A 2016  
47  
48 294 study of 110 insurance managers and their teams showed higher job satisfaction with higher mutual  
49  
50 295 ratings. Job satisfaction among employees (assessed on a 1 to 5 scale) was lowest when leaders rated  
51  
52 296 their leadership skills higher than their subordinates did (Mean 3.89 of 5 compared to 4.53 of 5 in  
53  
54 297 agreement) [45]. Rowold and Poethke who developed the FIF questionnaire conclude from their  
55  
56 298 studies that leaders can learn to adapt when receiving the leadership ratings as feedback. In addition,  
57  
58 299 they recommend to implement for example regular team meetings and improving leadership skills  
59  
60 300 through training [19]. Results from the DIALHS collaboration from South Africa point at the need for  
61  
62 301 accountability strategies such as standard operation procedures, facility audits and target setting [46].  
63  
64 302 While other studies followed this approach to share the assessment results with the ratee, we

1  
2  
3 303 abstained from this because previous studies in the medical field showed mixed reactions. In a 2005  
4 304 study, 15 family physicians rated multisource feedback extremely different, from negative to positive.  
5  
6 305 This evaluation was affected by the perceived usefulness, accuracy, and credibility [36].  
7  
8 306 Using the LMX questionnaire, a 2008 study with 200 nurses from six smaller and larger hospitals  
9  
10 307 showed positive associations of high mutual relationship scores with enhanced commitment, reduced  
11 308 staff turnover, and better organizational behavior [47]. Also, positive effects on employees' health and  
12 309 well-being are described in association with good relationships between leaders and employees. Lower  
13 310 levels of emotional exhaustion were associated with higher leader-member exchange quality in a  
14 311 sample of 343 employees working in the German healthcare sector after 11 months [48]. In addition,  
15 312 a hierarchical regression model showed that the LMX was a good predictor for the health of 412  
16 313 employees in health and social services in Germany [49]. Compared to the LMX reference values based  
17 314 on 113 participants, our study showed an overall better relationship quality between practice leaders  
18 315 and practice assistants (mean value of 28.1 of 35 for practice leaders and 26.8 for practice assistants  
19 316 vs. 22.9 in the LMX reference population) [35]. Higher scores in the practice setting are likely influenced  
20 317 by the fact that GP leaders recruit personnel themselves, while personnel recruitment and placement  
21 318 in larger institutions is not necessarily in the hands of the direct team leaders.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

319

### 320 ***Strengths and limitations***

321 Novel for the German GP setting, we investigated GP leadership in a large sample with analysis on  
322 practice level. Our data provide leadership ratings for each solo practice leaders, but not for each group  
323 practice leader, as we had asked staff to rate their leadership team to reflect current small team  
324 leadership situations. LMX data were missing for one of seven questions for the small number of  
325 employed physicians. However, the analysis of the available data yielded a high relationship quality  
326 with leaders like the results for practice assistants. A selection bias cannot be excluded as participating  
327 practices might have had a greater interest in the topic.  
328

328

### 329 ***Conclusion and practical implications***

330 Overall, our data from the IMPROVE<sub>job</sub> study show trustful relationships between GP leaders and their  
331 staff. Future GPs' trainings should enable GP leaders to implement goal-setting, innovation, and  
332 individuality focus more effectively. Our results support recent calls for leadership workshops on every  
333 level of the medical training for strengthening the GP and other health services workforce.  
334

334

### 335 **Acknowledgment**

336 All authors are part of the IMPROVE*job* consortium and owe special thanks to all participating  
337 physicians and practice assistants. In addition, special thanks go to the other members of the  
338 IMPROVE*job* consortium; current members: M. Brinkmann, S. Burgess, A. Herrmann-Werner, K.-H.  
339 Jöckel, F. Junne, A. Siegel, F. Stuber, A. Wagner, and former members: J. Block, J.-M. Bois, A. Dreher,  
340 M. Hippler, S. Hartmann, L. Imhoff, C. Kersting, C. Ose and E. Tsarouha. Additionally, we are grateful to  
341 all members of the scientific advisory board: Stefanie Esper, Katja Götz, Andre Habrock, Albert  
342 Nienhaus, Horst Christian Vollmar, and Andrea Wittich. The work of the Institute of Occupational and  
343 Social Medicine and Health Services Research is supported by an unrestricted grant of the Employers'  
344 Association of the Metal and Electrical Industry Baden-Wuerttemberg (Suedwestmetall). We  
345 acknowledge administrative support by the DLR Project Management Agency.

### 347 **Contributorship statement**

348 Conceptualization, B.M.W., M.A.R., T.S.-D., C.P., and B.W.; methodology, B.M.W. M.A.R., T.S.-D., M.S.,  
349 B.W., M.G., E.R., C.P., K.-H.J., L.D., and J.G.; validation, K.-H.J.; formal analysis, M.S. and J.G.;  
350 investigation, L.D., K.L., T.S.-D., B.W., E.R., C.P., J.G., M.A.R., and B.M.W.; resources, B.M.W., L.D., K.L.,  
351 T.S.-D., B.W., M.G., C.P., and E.R.; data curation, K.-H.J.; writing—original draft preparation, M.S. and  
352 B.M.W.; writing—review and editing, M.S., T.S.-D., B.W., M.G., E.R., C.P., K.-H.J., L.D., J.G., M.A.R. and  
353 B.M.W.; visualization, M.S.; supervision, B.M.W., M.A.R., and B.W.; project administration, K.L., E.R.,  
354 B.M.W. and M.A.R.; funding acquisition, B.M.W., M.A.R., C.P., and B.W. All authors have read and  
355 approved the published version of the manuscript.

### 357 **Competing interests**

358 The authors declare that they have no competing interests.

### 360 **Funding**

361 This study is funded by the German Federal Ministry for Education and Research (grant numbers:  
362 01GL1851D, 01GL1751B, 01GL1751A, 01GL1751C). The funder had no role in the study design, the data  
363 collection, and analysis, the decision to publish, or the preparation of the manuscript.

### 365 **Data sharing statement**

366 There are no plans to grant access to the full protocol, participant-level dataset, or statistical code as  
367 data contain potentially identifying information.

### 368 **Ethics approval**

369 The study complies with the ethical principles of the World Medical Association Declaration of Helsinki.



1  
2  
3 370 The study was approved first by the Ethics Committee of the Medical Faculty of the University of Bonn  
4  
5 371 (reference number: 057/19, date of approval: 20/02/2019). In addition, the Ethics Committees of the  
6  
7 372 Medical Association Nordrhein (Lfd-Nr.: 2019107), and of the Medical Faculty, University Hospital of  
8  
9 373 Tuebingen (project no.: 446/2019BO2) approved the study protocol.  
10  
11 374 All participating practice team members received written information and signed informed consent  
12  
13 375 forms, which are stored at the Institute for General Practice and Family Medicine, University of Bonn.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For peer review only

## REFERENCES

- 1  
2  
3  
4 376
- 5 377 1 Kringos D, Boerma W, Bourgueil Y, et al. The strength of primary care in Europe: an  
6 378 international comparative study. *British Journal of General Practice* 2013;63(616):e742-e750.
  - 7 379 2 Rohde J, Cousens S, Chopra M, et al. 30 years after Alma-Ata: has primary health care worked in  
8 380 countries? *The Lancet* 2008;372(9642):950–61.
  - 9 381 3 Swanwick T, Varnam R. Leadership development and primary care. *BMJ Leader* 2019;3(2).
  - 10 382 4 Meredith LS, Bouskill K, Chang J, et al. Predictors of burnout among US healthcare providers: a  
11 383 systematic review. *BMJ Open* 2022;12(8):e054243.
  - 12 384 5 Boamah SA, Spence Laschinger HK, Wong C, et al. Effect of transformational leadership on job  
13 385 satisfaction and patient safety outcomes. *Nurs Outlook* 2018;66(2):180–89.  
14 386 doi:10.1016/j.outlook.2017.10.004 [published Online First: 23 November 2017].
  - 15 387 6 Goetz K, Campbell S, Broge B, et al. Job satisfaction of practice assistants in general practice in  
16 388 Germany: an observational study. *Fam Pract* 2013;30(4):411–17. doi:10.1093/fampra/cmt015  
17 389 [published Online First: 19 April 2013].
  - 18 390 7 i Solà GJ, i Badia JG, Hito PD, et al. Self-perception of leadership styles and behaviour in primary  
19 391 health care. *BMC Health Services Research* 2016;16.
  - 20 392 8 Kersting C, Zimmer L, Thielmann A, et al. Chronic stress, work-related daily challenges and  
21 393 medicolegal investigations: a cross-sectional study among German general practitioners. *BMC*  
22 394 *Fam Pract* 2019;143. doi:10.1186/s12875-019-1032-6 [published Online First: 24 October  
23 395 2019].
  - 24 396 9 Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *The Lancet*  
25 397 2009;374(9702):1714–21.
  - 26 398 10 Wong CA, Cummings GG, Ducharme L. The relationship between nursing leadership and patient  
27 399 outcomes: a systematic review update. *J Nurs Manag* 2013;21(5):709–24.
  - 28 400 11 Sfantou DF, Laliotis A, Patelarou AE, et al. Importance of Leadership Style towards Quality of  
29 401 Care Measures in Healthcare Settings: A Systematic Review. *Healthcare (Basel, Switzerland)*  
30 402 2017;5(4):73.
  - 31 403 12 Söling S, Pfaff H, Karbach U, et al. How is leadership behavior associated with organization-  
32 404 related variables? Translation and psychometric evaluation of the implementation leadership  
33 405 scale in German primary healthcare. *BMC Health Services Research* 2022;22(1):1–13.
  - 34 406 13 Bass BM. Two Decades of Research and Development in Transformational Leadership.  
35 407 *European Journal of Work and Organizational Psychology* 1999;8(1):9–32.
  - 36 408 14 Poethke U. Leadership, values and communication: a cross-cultural investigation of the  
37 409 extended full-range of leadership behaviors: Doctoral dissertation 2019.
  - 38 410 15 Rowold J, Schlotz W. Transformational and transactional leadership and followers' chronic  
39 411 stress. *Leadership Review* 2009;9(1):35–48.
  - 40 412 16 Graen GB, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-  
41 413 member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-  
42 414 domain perspective. *Leadership Quarterly* 1995;6(2):219–47.
  - 43 415 17 Graen GB, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-  
44 416 member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-  
45 417 domain perspective. *The Leadership Quarterly* 1995;6(2):219–47.
  - 46 418 18 Gregersen S, Vincent-Höper S, Nienhaus A. Health-Relevant Leadership Behaviour: A  
47 419 Comparison of Leadership Constructs. *German Journal of Human Resource Management*  
48 420 2014;28(1-2):117–38.
  - 49 421 19 Rowold J, Poethke U. FIF: Fragebogen zur Integrativen Führung Manual, 1st edn. Göttingen:  
50 422 Hogrefe 2017.

- 1  
2  
3 423 20 Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof* 2003;23(1):4–12.
- 4 424
- 5 425 21 Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with  
6 426 observed measures of competence: a systematic review. *JAMA* 2006;296(9):1094–102.
- 7 427 22 Donnon T, Al Ansari A, Al Alawi S, et al. The reliability, validity, and feasibility of multisource  
8 428 feedback physician assessment: a systematic review. *Acad Med* 2014;89(3):511–16.
- 9 429 23 Saravo B, Netzel J, Kiesewetter J. The need for strong clinical leaders - Transformational and  
10 430 transactional leadership as a framework for resident leadership training. *PLoS One*  
11 431 2017;12(8):e0183019.
- 12 432 24 Richter A, Lornudd C, Thiele Schwarz U von, et al. Evaluation of iLead, a generic implementation  
13 433 leadership intervention: mixed-method preintervention-postintervention design. *BMJ Open*  
14 434 2020;10(1):e033227. doi:10.1136/bmjopen-2019-033227 [published Online First: 12 January  
15 435 2020].
- 16 436 25 Hill DA, Jimenez J-C, Cohn SM, et al. How to be a leader: a course for residents. *Cureus*  
17 437 2018;10(7).
- 18 438 26 Viehmann A, Kersting C, Thielmann A, et al. Prevalence of chronic stress in general practitioners  
19 439 and practice assistants: Personal, practice and regional characteristics. *PLoS One*  
20 440 2017;12(5):e0176658.
- 21 441 27 Dreher A, Theune M, Kersting C, et al. Prevalence of burnout among German general  
22 442 practitioners: Comparison of physicians working in solo and group practices. *PLoS One*  
23 443 2019;14(2):e0211223.
- 24 444 28 Freund T, Everett C, Griffiths P, et al. Skill mix, roles and remuneration in the primary care  
25 445 workforce: who are the healthcare professionals in the primary care teams across the world?  
26 446 *International journal of nursing studies* 2015;52(3):727–43.
- 27 447 29 infas im Auftrag der KBV und des NAV-Virchow-Bund. Tabellenband Ärztemonitor 2018,  
28 448 Ergebnisse für Haus- und Fachärzte. Bonn 2018.
- 29 449 30 Weltermann BM, Kersting C, Pieper C, et al. IMPROVEjob - Participatory intervention to  
30 450 improve job satisfaction of general practice teams: a model for structural and behavioural  
31 451 prevention in small and medium-sized enterprises - a study protocol of a cluster-randomised  
32 452 controlled trial. *Trials* 2020;21(1):532.
- 33 453 31 Degen L, Linden K, Seifried-Dübon T, et al. Job Satisfaction and Chronic Stress of General  
34 454 Practitioners and Their Teams: Baseline Data of a Cluster-Randomised Trial (IMPROVEjob). *Int J*  
35 455 *Environ Res Public Health* 2021;18(18). doi:10.3390/ijerph18189458 [published Online First: 8  
36 456 September 2021].
- 37 457 32 Ingold PV. Fragebogen zur integrativen Führung. *Zeitschrift für Arbeits- und*  
38 458 *Organisationspsychologie A&O* 2019;63(2):100–06.
- 39 459 33 Stuber F, Seifried-Dübon T, Rieger MA, et al. Investigating the Role of Stress-Preventive  
40 460 Leadership in the Workplace Hospital: The Cross-Sectional Determination of Relational Quality  
41 461 by Transformational Leadership. *Front Psychiatry* 2019;10:622. doi:10.3389/fpsy.2019.00622  
42 462 [published Online First: 3 September 2019].
- 43 463 34 Schyns B. Überprüfung einer deutschsprachigen Skala zum Leader-Member-Exchange-Ansatz.  
44 464 *Zeitschrift für Differentielle und Diagnostische Psychologie* 2002.
- 45 465 35 Schyns B, Paul T. Leader-Member Exchange Skala (LMX7): ZIS - GESIS Leibniz Institute for the  
46 466 Social Sciences 2014.
- 47 467 36 Sargeant J, Mann K, Ferrier S. Exploring family physicians' reactions to multisource feedback:  
48 468 perceptions of credibility and usefulness. *Med Educ* 2005;39(5):497–504.
- 49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 469 37 Hu J, Lee R, Mullin S, et al. How physicians change: Multisource feedback driven intervention  
4 470 improves physician leadership and teamwork. *Surgery* 2020;168(4):714–23.  
5 471 doi:10.1016/j.surg.2020.06.008 [published Online First: 6 August 2020].  
6 472 38 Northouse PG. Leadership: Theory and practice, 6th edn. Thousand Oaks: Sage Publications  
7 473 2010.  
8 474 39 Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Hoboken: Taylor and  
9 475 Francis 1988.  
10 476 40 Halverson SK, Tonidandel S, Barlow C, et al. Self-other agreement on a 360-degree leadership  
11 477 evaluation. In: 17th Annual Conference of the Society for Industrial and Organizational  
12 478 Psychology, Toronto, Canada 2002.  
13 479 41 Furnham A, Stringfield P. Congruence of self and subordinate ratings of managerial practices as  
14 480 a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*  
15 481 1994;67(1):57–67.  
16 482 42 Fleenor JW, Smither JW, Atwater LE, et al. Self–other rating agreement in leadership: A review.  
17 483 *The Leadership Quarterly* 2010;21(6):1005–34.  
18 484 43 Antonioni D, Park H. The effects of personality similarity on peer ratings of contextual work  
19 485 behaviors. *Personnel Psychology* 2001;54(2):331–60.  
20 486 44 Alrubaysh MA, Alshehri MH, Alsuhaibani EA, et al. The leadership styles of primary healthcare  
21 487 center managers and center performance outcomes in Riyadh, Saudi Arabia: A correlational  
22 488 study. *Journal of Family & Community Medicine* 2022;29(1):56.  
23 489 45 Erben J, Schneider FM, Maier M. In the Ear of the Beholder: Self-Other Agreement in  
24 490 Leadership Communication and Its Relationship With Subordinates' Job Satisfaction.  
25 491 *International Journal of Business Communication* 2019;56(4):505–29.  
26 492 46 Cleary S, Du Toit A, Scott V, et al. Enabling relational leadership in primary healthcare settings:  
27 493 lessons from the DIALHS collaboration. *Health Policy and Planning* 2018;33(suppl\_2):ii65–ii74.  
28 494 47 Chen C-HV, Wang S-J, Chang W-C, et al. The effect of leader-member exchange, trust,  
29 495 supervisor support on organizational citizenship behavior in nurses. *Journal of Nursing Research*  
30 496 2008;16(4):321–28.  
31 497 48 Gregersen S, Vincent-Höper S, Nienhaus A. Job-related resources, leader–member exchange  
32 498 and well-being – a longitudinal study. *Work & Stress* 2016;30(4):356–73.  
33 499 49 Vincent-Höper S, Stein M, Gregersen S, et al. Messung gesundheitsförderlicher Führung in  
34 500 Gesundheitsdienst und Wohlfahrtspflege. *ASU Arbeitsmed Sozialmed Umweltmed* 2018(53):46–  
35 501 53.  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 **1 Additional Files**4  
5 **2 Additional file 1: CONSORT 2010 checklist of information to include when reporting a randomised trial**6  
7 38  
9 410  
11 5
**CONSORT 2010 checklist of information to include when reporting a randomised trial\***

Section/Topic	Item No	Checklist item	Reported on page No
<b>Title and abstract</b>			
	1a	Identification as a randomised trial in the title	n/a
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	Abstract, 2
<b>Introduction</b>			
Background and objectives	2a	Scientific background and explanation of rationale	Background, 4-5
	2b	Specific objectives or hypotheses	Background, 4-5
<b>Methods</b>			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Methods, 5
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	Methods, 5-6
	4b	Settings and locations where the data were collected	Methods, 5-6
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	n/a
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	Methods, 6-7
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	Study protocol, Weltermann et al., Trials 2020
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

Randomisation:

Sequence generation	8a	Method used to generate the random allocation sequence	Details in Study protocol, Weltermann et al., Trials 2020
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details in Study protocol, Weltermann et al., Trials 2020
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	n/a
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Details in Study protocol, Weltermann et al., Trials 2020
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	n/a
	11b	If relevant, description of the similarity of interventions	n/a
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	Methods,7-8
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	n/a
<b>Results</b> 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	n/a
	13b	For each group, losses and exclusions after randomisation, together with reasons	n/a
Recruitment	14a	Dates defining the periods of recruitment and follow-up	Details in Study protocol, Weltermann et al., Trials 2020
	14b	Why the trial ended or was stopped	Details in Study protocol,

				Weltermann et al., Trials 2020
	Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Results, table 1, Page 8
	Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Results, 9
	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)	Results, 9-11
		17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	N/A.
	Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	n/a
	Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	n/a
	<b>Discussion</b>			
	Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	Discussion, Strengths and limitations, Page 12-13
	Generalisability	21	Generalisability (external validity, applicability) of the trial findings	13
	Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	Discussion, 11-13
	<b>Other information</b>			
	Registration	23	Registration number and name of trial registry	Trial registration, 2
	Protocol	24	Where the full trial protocol can be accessed, if available	Study protocol, Weltermann et al., Trials 2020
	Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	Funding, 17-18

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

7 \*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  
8 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  
9 references relevant to this checklist, see [www.consort-statement.org](http://www.consort-statement.org).

For peer review only