JIM Original Article

doi: 10.1111/joim.13374

The impact of a tailored mindfulness-based program for resident physicians on distress and the quality of care: A randomised controlled trial

Johannes C. Fendel[1,2], Vanessa M. Aeschbach[1,2], Stefan Schmidt[2] & Anja S. Göritz[1]

From the [1]Department of Occupational and Consumer Psychology, Institute of Psychology, University of Freiburg, Freiburg, Germany; and [2]Department for Psychosomatic Medicine and Psychotherapy, Medical Faculty, Medical Centre, University of Freiburg, Freiburg, Germany

Abstract. Fendel JC, Aeschbach VM, Schmidt S, Göritz AS. The impact of a tailored mindfulness-based program for resident physicians on distress and the quality of care: A randomised controlled trial. *J Intern Med.* 2021; **290:** 1233–1248.

Background. Many resident physicians suffer from distress, which endangers their individual health and the quality of care.

Objective. To examine the impact of a tailored mindfulness-based program (MBP) for resident physicians on distress and the quality of care.

Methods. A single-centre, two-armed, longitudinal randomised controlled trial. The intervention group took part in an 8-week, tailored MBP that included a coursebook. The MBP was followed by a 4-month maintenance phase. The active control group received the coursebook for self-study. Assessments were at baseline (t0, 0 months), after the intervention (t1, 2 months), after the maintenance phase (t2, 6 months), and at follow-up (t3, 12 months). The primary outcome was a change in burnout at t2. Secondary outcomes included perceived stress, mental distress, perceived job strain, depression, anxiety, hair cortisol secretion, self-reported medical errors and third-party ratings by patients, supervisors and colleagues.

Results. Seventy-six participants were randomised to the intervention and 71 to the control group. The intervention group showed greater improvements in the primary outcome (burnout at t2, $d = 0.32$, $p = 0.046$), in perceived stress ($d = 0.31$, $p = 0.046$) and perceived job strain ($d = 0.33$, $p = 0.026$) at t1, and in supervisor rated empathy ($d = 0.71$, $p = 0.037$) and colleague rated attentiveness ($d = 0.85$, $p = .006$) at t2. There was no difference between groups in the other outcomes.

Conclusion. A tailored MBP for resident physicians improved burnout and might have improved other aspects of distress and the quality of care.

Keywords: burnout, cortisol, distress, mindfulness, quality of care, resident physicians

Background

Residency is a demanding period in a physician's career. Excessive workload, long working hours, administrative burdens, scarcity of supervisor support, restricted autonomy and economic pressure behind medical decisions contribute to resident physicians' high levels of distress [1–3]. More than half of resident physicians are affected by burnout [4], which exceeds the prevalence in practising physicians [4–7], medical students [4], faculty [8] and the general working population [4]. Burnout is often defined as a work-related syndrome that is characterised by emotional exhaustion, depersonalisation and reduced personal accomplishment [9,10]. It is often used as a surrogate measure of heightened distress in resident physicians [11,12]. Besides burnout, resident physicians report higher levels of depression than the general population [13] and are less satisfied with their life [14]. Burnout in resident physicians can have serious personal consequences such as substance abuse [15], suicidal thoughts [16], work-life conflicts [17] and increased odds of motor vehicle incidents [18]. Furthermore, burnout endangers the quality of care, as burned-out resident physicians seem to commit more medical errors [17,19–23], have less medical knowledge [24], show

reduced work engagement [25], express less empathy [17,26,27], inferior social skills [28] and adhere less to practice and safety standards [23]. On a societal level, physician burnout causes tremendous costs in Western healthcare systems through higher levels of absenteeism, job turnover and early retirement (e.g., $4.6 billion in the United States per year) [29]. In light of these findings, it is important to reduce rates of burnout and other symptoms of distress among resident physicians, for the sake of themselves, their patients and the quality of care in general [30].

Approaches to reduce resident physicians' distress can be classified into initiatives directed at organisations, that is, targeting the work environment, and initiatives directed at individuals, that is, targeting physicians [31]. A promising individual-directed approach is the implementation of mindfulness-based programs (MBPs). Mindfulness can be described as moment-to-moment awareness, cultivated by paying attention to the present moment, as non-judgmentally and openheartedly as possible [32]. A meta-analysis has shown that MBPs are effective in reducing physicians' burnout and stress [33]. This can be explained by the notion that, through training in mindfulness, resident physicians increase self-awareness, strengthen their ability to set priorities and limits regarding their work and learn to be more accepting of unpleasant and difficult experiences [34]. Moreover, MBPs for physicians have been shown to improve the quality of care, for example, in terms of increased empathy, dedication to work or hand hygiene adherence [35]. The positive effects of mindfulness practice on the quality of care can be explained by Epstein's notion that mindful physicians are more aware of their own physical and mental processes and better recognise bias in judgment [36]. This critical self-reflection would enable physicians to listen attentively to their patients, to communicate with greater awareness, and to act with compassion; in other words, to deliver a higher quality of care [36].

However, the evidence on MBPs for resident physicians is ambiguous and lacks methodological rigour. There are only three randomised controlled trials (RCTs) [34,37, 38], while most evidence is based on non-randomised trials [33]. Furthermore, effects have almost exclusively been measured by self-reports. Moreover, the complex interplay of resident physicians' specific work stressors, resources, and personality traits call for specifically tailored programs [35,39–42]. With the present study, we intended to improve on these shortcomings as we developed a tailored program that takes resident physicians' particular needs into account and puts a focus on how to integrate mindfulness into daily medical practice. For the assessment of this program, we conducted an RCT that entailed different types of outcome measures including self-report measures, hair cortisol secretion as a biomarker of stress, third-party ratings by patients, supervisors and colleagues, as well as information provided by human resource departments.

We hypothesised that participation in a novel, tailored MBP for resident physicians would result in reduced rates of burnout and other symptoms of distress, as well as improved markers of the quality of care.

Methods

Trial design and participants

We conducted a single-centre, two-armed, parallel, longitudinal RCT comparing a tailored MBP to an active control condition. The allocation ratio was 1:1. Measurements were taken at baseline (t0, 0 months), after the program (t1, 2 months), after a maintenance phase (t2, 6 months) and at follow-up 6 months after completion of the maintenance phase (t3, 12 months). Participants were recruited through emails, flyers, a study webpage, short presentations at division meetings, radio and lay media announcements. Eligible participants were physicians younger than 45, with an ongoing position as a resident physician at baseline, and minimum employment of 40%. The study took place at the University of Freiburg, Germany from September 2018 to May 2020. Participants received no financial compensation but could collect points for Continuing Medical Education. Participants provided written informed consent. The program trainers were three psychiatrists who are highly experienced mindfulness instructors, certified by the German Mindfulness-based Stress Reduction program (MBSR) teacher association [43]. The study protocol was pre-registered (trial registration: DRKS00014015), published [42], and approved by the Ethics Committee of the University of Freiburg, Germany. The report of the study follows the CONSORT statement for non-pharmacological treatment interventions [44].

Interventions

The intervention group engaged in eight (135 min one evening per week) guided group sessions (maximum 14 participants per group) as well as a full day 6-h, silent retreat. The sessions were followed by a 4-month maintenance phase consisting of three monthly booster sessions. We based the program on the validated MBSR program [45] and tailored it to resident physicians' particular needs and circumstances. The tailoring process, program content and feasibility findings have been described elsewhere [40,42]. Importantly, as proposed in the literature, we introduced mindfulness as a practice of self-care, in order to promote personal well-being, meaning and professional fulfilment rather than as a means to foster stress resistance [30,46,47]. We did this to prevent the mindfulness practice from being purposely or unwittingly functionalised for self- or performance optimisation [48]. Each session followed the same structure: (1) theoretical input (20 min), (2) formal mindfulness practice (45 min), (3) group inquiry (40 min), (4) integration into daily medical practice (25 min) and (5) practice-at-home assignments (5 min). An outline of the session themes and a summary of contents are in Table S1. After each session, participants received a coursebook containing detailed information and a description of practical exercises about mindfulness and its relationship with stress and quality of care as well as texts about the importance of self-care, acceptance and meaning in medicine. For treatment fidelity, the trainers received a curriculum guide that included a comprehensible schedule and detailed accompanying material.

The control group received the same coursebook on the same weekly basis as the intervention group, except that the coursebook for the control group did not contain and a description of practical exercises. Thus, we paralleled the groups with respect to information (i.e., description-based learning) but contrasted them with respect to guided experience and practice of mindfulness (i.e., experience-based learning). Description and experience are both powerful ways of learning and adaptation but involve systematically distinct cognitions and behaviours [49]. For a full understanding of mindfulness, it has been suggested that this requires an introspective practical engagement in mindfulness, resulting in first-person experience. This experience, in turn, cannot be gained through the description (i.e., speaking, writing, theorising) alone [50]. We hypothesised that the combined acquisition of mindfulness from first-person experience and description in the intervention group would result in greater benefits than learning from the description alone in the control group. After the completion of their participation in the trial, participants of the control group were offered to take part in the MBP.

Outcomes

To administer self-report measures, we used the online survey platform Unipark EFS Survey by Questback GmbH, Cologne, Germany. Cronbach's alphas stem from the current sample at baseline. In addition to the study outcomes we report below, we assessed positive mental health, qualitative and implicit measures, as well as keyboard and mouse usage as an indirect measure of stress. These outcomes are reported elsewhere [42].

The primary outcome, as established in the study protocol, was a change in burnout levels between t0 and t2, measured by the 19-item Copenhagen Burnout Inventory (CBI, $\alpha = 0.93$) [9]. Participants rated to what extent they experienced exhaustion (i.e., the core component of burnout) [51], on three subscales for personal burnout (six items, $\alpha = 0.84$), work-related burnout (seven items, $\alpha = 0.85$), and client-related burnout (six items, $\alpha = 0.90$). All items are scored on a 5-point scale (from *never* to *very often*; range, 0–100). Secondary outcomes included measures of distress other than burnout as well as hair cortisol secretion, markers of the quality of care, satisfaction, and attendance.

Distress. General mental distress was measured by the 12-item General Health Questionnaire ($\alpha = 0.89$) [52]. Participants rated how often they experienced symptoms of psychological and psychiatric disorders during the past weeks on a 4-point scale (from *not at all* to *much more than usual*; range, 0–3). Stress was measured by the 10-item Perceived Stress Scale ($\alpha = 0.86$) [53]. Participants rated the frequency of stress-related feelings and thoughts during the past month on a 5-point scale (from *never* to *very often*; range, 0–4). Depression ($\alpha = 0.69$) and anxiety ($\alpha = 0.74$) were measured by the four-item Patient Health Questionnaire (PHQ-4) [54]. Participants rated how often they experienced symptoms of depression and anxiety during the past 2 weeks on a 4-point scale (from *not at*

Table 1. *Baseline demographics by group*

Variable	Total ($n = 147$)	Intervention ($n = 76$)	Control ($n = 71$)
Women, No. (%)	96 (65.31)	49 (64.47)	47 (66.20)
Age, mean (SD)	31.02 (3.43)	31.04 (3.39)	31.00 (3.49)
In a relationship, *n* (%)	109 (74.15)	56 (73.68)	53 (74.65)
One or more children, *n* (%)	22 (14.97)	9 (11.84)	13 (18.31)
Years in practice, mean (SD)	3.0 (1.66)	3.20 (1.67)	2.79 (1.63)
Hours worked per week, mean (SD)	48.96 (9.19)	48.82 (8.63)	49.10 (9.81)
Meditation experience, Yes. (%)	54 (36.73)	32 (42.11)	22 (30.99)
MBSR	6 (4.08)	3 (3.95)	3 (4.23)
Other mindfulness course	7 (4.76)	4 (5.26)	3 (4.23)
Regular personal practice	9 (6.12)	5 (6.58)	4 (5.63)
Retreat	5 (3.40)	3 (3.95)	2 (2.82)
Other experience	33 (22.45)	20 (26.32)	13 (18.31)
Specialty, No. (%)			
Internal medicine	34 (23.13)	19 (25.00)	15 (21.13)
Paediatrics	15 (10.20)	7 (9.21)	8 (11.27)
Psychiatry	14 (9.52)	7 (9.21)	7 (9.86)
Anaesthesiology	13 (8.84)	7 (9.21)	6 (8.45)
Neurology/neuropathology	12 (8.16)	2 (2.63)	10 (14.08)
Dentistry	10 (6.80)	5 (6.58)	5 (7.04)
Psychosomatic medicine	8 (5.44)	5 (6.58)	3 (4.23)
Gynaecology	7 (4.76)	4 (5.26)	3 (4.23)
Dermatology	6 (4.08)	2 (2.63)	4 (5.63)
Radiology	6 (4.08)	5 (6.58)	1 (1.41)
Urology	5 (3.40)	2 (2.63)	3 (4.23)
Nuclear medicine	4 (2.72)	2 (2.63)	2 (2.82)
Surgery	3 (2.04)	1 (1.32)	2 (2.82)
Ophthalmology	2 (1.36)	1 (1.32)	1 (1.41)
Orthodontics	2 (1.36)	2 (2.63)	0 (0.00)
Neurosurgery	2 (1.36)	2 (2.63)	0 (0.00)
Orthopaedics	2 (1.36)	2 (2.63)	0 (0.00)
Otorhinolaryngology	1 (0.68)	1 (1.32)	0 (0.00)
Oral and maxillofacial surgery	1 (0.68)	0 (0.00)	1 (1.41)

Abbreviation: MBSR, mindfulness-based stress reduction.

all to *almost every day*; range, 0–3). Perceived job strain was measured by the eight-item Irritation Scale ($\alpha = 0.87$) [55]. Participants rated to what extent statements about job strain apply to them on a 7-point scale (from *not at all* to *very much*; range, 1–7), divided into two subscales for cognitive strain (three items, $\alpha = 0.87$) and emotional strain (five items, $\alpha = 0.88$).

Hair cortisol. We took hair samples of 1 cm in length and 3 mm in diameter close to the scalp from a posterior vertex position at four on-site appointments. We did this to assess hair cortisol secretion as a biomarker of stress via the Immunoassay method [56]. Hair sampling provides an objective measure of longer term cortisol secretion. While transient cortisol secretion is an adaptive response to cope with acute stressors, elevated levels of longer term cortisol secretion indicate chronic stress, which is related to ill health [57]. The chosen approach enables the evaluation of the cumulative cortisol level of the 4 weeks prior to assessment [58]. Exclusion criteria were baldness, pregnancy, use of glucocorticoid medication and adrenocortical dysfunction (e.g., Cushing Syndrome, Morbus Addison) [59,60].

Quality of care. Medical errors were assessed by a six-item scale ($\alpha = 0.78$) to gauge the quality of patient care provided by resident physicians [19]. Participants indicated, on a 5-point scale, how often errors occurred (from *never occurs* to *occurs often*; range, 1–5). This scale was translated into German by the authors. To measure absenteeism, we asked human resource departments of the respective hospitals at t3 to indicate how many days participants had been absent due to illness during 12 months prior to the start of the study and during 12 months of the study period (open response format). In addition, at t0 and t2, participants selected one of their supervisors and one of their physician colleagues to rate on three items: how tense, empathic and attentive they appeared. Moreover, three patients rated on five items: how attentive, empathic and competent the respective resident physician appeared as well as how satisfied they were with the physician and the care. Ratings were given on a 7-point scale (from *not at all* to *very much*; range, 1–7). All third-party ratings were given anonymously and were not shown to the participating resident physicians. In contrast to pre-registration, we do not report assessments by the Jefferson Scale of Physician Empathy due to licensing issues [61].

Satisfaction and attendance. At t3, participants rated on five self-constructed items their overall program satisfaction, satisfaction with trainers, perceived professional as well as private benefit/harm and willingness to recommend the program to peers (7-point scale; range, 1–7; higher values indicate greater satisfaction, benefit and willingness). In line with previous studies, study completion was defined by having attended four or more sessions [34,62].

Sample size

We planned to enrol 178 participants to detect effects on mental health variables of 0.45 standard deviations with 80% power [63] and with an anticipated dropout of 30%.

Randomisation

Participants contacted us via telephone and email were assessed for eligibility and received confirmation of study enrolment via email from the study team. We used minimisation to allocate participants into one of two groups using the software Qminim [64]. Through this approach, we minimised the imbalance between the groups with regard to gender (male, female) and baseline levels of personal burnout (CBI cut-off values 0–37.4 = low, 37.5–62.4 = medium, 62.5–100 = high burnout) [65]. We applied a weighted random allocation with a probability of 0.8 to minimise imbalance. A researcher with no contact with participants carried out minimisation and group assignment after the completion of baseline assessments.

Blinding

Due to the nature of the interventions, participants and trainers were aware of the allocated arm. To minimise bias, self-report measures were administered online. Moreover, the outcome assessors of on-site assessments (i.e., hair sampling and instructions for third-party ratings) were blinded to group allocation and were not involved in data analyses.

Statistical methods

We conducted the analyses according to the intent-to-treat principle. We performed two-tailed tests and considered findings with $p < 0.05$ as statistically significant. We calculated *t*-tests and chi-squared tests to conduct baseline comparisons between groups. Missing outcome data were handled by linear mixed modelling using maximum likelihood estimation. Through graphical analyses and statistical tests provided in the R-package *MissMech*, we determined all missing data to be missing at random. The main analyses were performed by linear mixed modelling the group by point in time interaction, thus taking possible baseline differences between groups into account. The model included group, point in time and the group by point in time interaction as fixed effects. Moreover, the model included a random intercept to take into account inter-individual differences as well as an autoregressive residual covariance structure to take into account correlations that arise from repeated measures. For a comparison between groups across all points in time, an overall interaction effect was calculated. For a comparison between groups at a particular point in time, dummy-coded treatment contrasts were calculated. For additional within-group analyses, the model included the same parameters except for the group and interaction effect. We adjusted all models to allow for different residual variances if the assumption of homogeneity was not met [66]. We replaced outliers (i.e., values ≥ 3 SDs from the mean) with the highest/lowest value that excluded the outliers (Winsor method). By this

correction, less than 1% of the data were modified. We did not adjust the level of significance for multiple testing [67]. Therefore, the analysis of all outcomes except the primary outcome should be considered exploratory and interpreted with caution. We calculated Cohen's *d* with the adjusted model-based differences in change from baseline divided by the standard deviations at baseline [68]. All analyses were carried out using R version 4.0.3.

Results

Recruitment and participant flow

Recruitment took place between July 2018 and May 2019. The recruitment ended when the scheduled closure date were reached. The final sample for the primary outcome consisted of 76 resident physicians in the intervention group and 71 in the control group (shown in Fig. 1). Participants worked in at least 25 different clinics including a university hospital, hospitals in private and public ownership, as well as church-funded hospitals, from both urban and rural areas (some participants did not provide information on their employers). We knew from previous studies that anonymity and data protection were important issues for many resident physicians. Therefore, participants generated an individual code for the data assessments, which only they could link to their data. Due to these efforts to ensure anonymity and data protection, we were not able to ascertain the reasons for dropout. The treatment for both groups started within 1 month after randomisation.

Baseline measures

At baseline, there were no statistically significant differences between groups with regard to demographics, meditation experience (Table 1), distress, and quality of care outcomes (Table 2), except for a difference in how attentive the resident physicians were, as judged by their colleagues ($p = .041$). The raw outcome values without any model adjustments are shown in Table 2.

Primary outcome

From baseline to t2, the dummy-coded contrast indicated that the intervention group showed statistically significantly greater reductions in burnout levels than the control group ($d = 0.32$; $p = 0.046$; MD = –4.81; 95% CI = –9.52, –0.11; Table 3). This effect was small. There were no statistically significant differences between groups with regard to changes in burnout levels from baseline to t1, from baseline to follow-up or across all time periods (all *p*s > 0.05). Parameter estimates for all group comparisons are shown in Table 3. The plots are shown in Figs S1 and S2.

Secondary outcomes

Distress and hair cortisol. From baseline to t1, the intervention group showed statistically significantly greater reductions in perceived stress ($d = 0.31$; $p = 0.046$; MD = –1.76; 95% CI [–3.49, –0.04]) and perceived job strain ($d = 0.29$; $p = 0.044$; MD = –0.36; 95% CI [–0.70, –0.01]) compared to the control group (Table 3). These differences were small. We found no statistically significant differences between groups in change in the other distress outcomes and hair cortisol at any single point in time, or across all time periods (all *p*s > 0.05). Within-group analyses revealed that scores in most distress outcomes and hair cortisol decreased statistically significantly in both groups to a small to a medium degree (Tables S2 and S3).

Quality of care. From baseline to t2, the intervention group showed a statistically significantly greater gain in how empathic they were, as judged by their supervisors ($d = 0.71$; $p = 0.037$; MD = 0.73; 95% CI [0.06, 1.40]), and in how attentive they were, as judged by their colleagues ($d = 0.85$; $p = 0.006$; MD = 0.95; 95% CI [0.30, 1.60]) compared to the control group (Table 3). These differences were medium and large. There were no statistically significant differences in change in medical errors, patients' ratings and absenteeism between the groups at any point in time (all *p*s > .05). Within-group analyses revealed statistically significant reductions in self-reported medical errors across all time periods in both groups (Tables S2 and S3). Third-party ratings improved only in the intervention group, with a statistically significant, small increase in how empathic participants were, and a statistically significant medium increase in how attentive they were, as judged by their colleagues.

Additional outcomes: Satisfaction and attendance

The mean ratings of the intervention group for satisfaction with the program and the trainers, for professional and private benefit, and for the willingness to recommend the program were high (Table 4). Of the 77 participants randomised to the intervention group, 71 completed the 8-week pro-

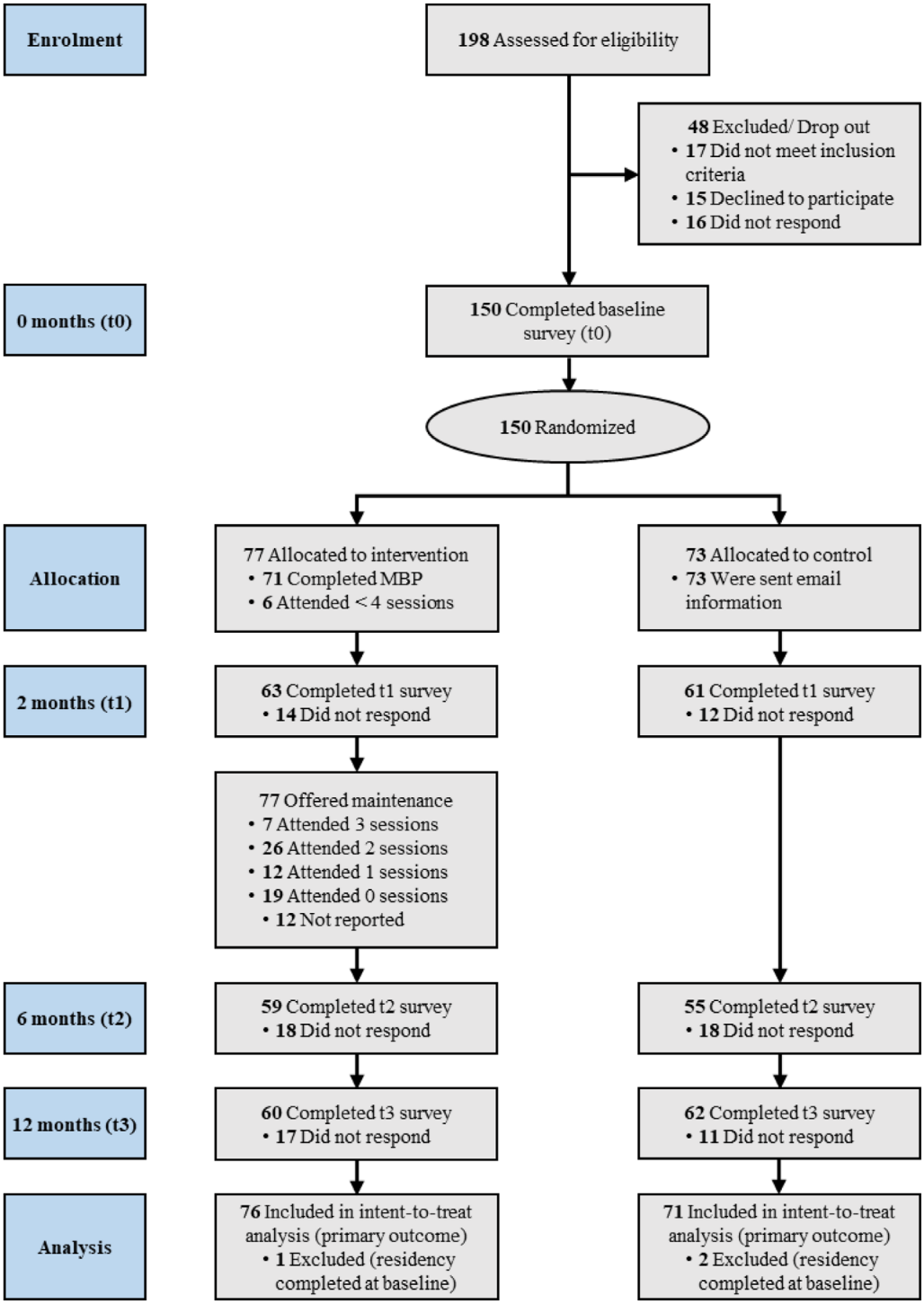

Fig. 1 *Study flowchart.*

Table 2. *Means, standard deviations and sample sizes by the group across the four points in time (raw values)*

Measure		Baseline (t0)						After 2 months (t1)						After 6 months (t2)						After 12 months (t3)					
		Intervention			Control			Intervention			Control			Intervention			Control			Intervention			Control		
		M	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*	*M*	SD	*N*
Distress	Burnout	41.72	13.88	76	41.98	16.95	71	39.79	13.14	63	40.22	15.77	60	37.89	15.46	58	42.86	17.92	54	36.05	14.29	60	39.69	14.48	61
	Personal	50.22	16.17	76	53.23	17.69	71	46.56	14.07	63	50.62	17.97	60	45.04	15.87	58	51.70	19.35	54	42.15	16.42	60	48.77	17.39	61
	Work related	44.88	15.07	76	44.32	19.16	71	42.86	15.25	63	41.79	18.30	60	39.22	17.86	58	44.51	20.15	54	37.08	15.22	60	40.81	15.76	61
	Client related	29.55	18.13	76	27.99	21.57	71	29.43	17.26	63	27.99	19.75	60	29.17	18.20	58	32.10	21.88	54	28.75	18.05	60	29.30	19.60	61
	Mental distress	14.13	5.94	76	14.56	5.58	71	9.71	5.47	63	11.65	4.27	60	11.22	5.90	58	12.80	5.91	54	10.55	6.17	60	12.41	6.05	61
	Perceived stress	19.58	5.95	76	19.87	5.73	71	17.40	5.63	63	18.52	5.66	60	17.53	6.52	58	17.94	6.43	54	16.57	6.39	60	17.92	6.16	61
	Depression	1.72	1.16	76	1.73	1.24	71	1.35	0.86	63	1.42	1.09	60	1.26	1.09	58	1.35	1.18	54	1.32	1.24	60	1.26	1.08	61
	Anxiety	2.26	1.48	76	2.14	1.54	71	1.67	1.18	63	1.65	1.13	60	1.79	1.51	58	1.78	1.00	54	1.83	1.21	60	1.79	1.13	61
	Perceived job strain	3.77	1.20	76	4.02	1.27	71	3.17	1.20	63	3.69	1.18	60	3.31	1.26	58	3.56	1.30	54	3.20	1.20	60	3.56	1.32	61
	Cognitive	4.45	1.49	76	4.68	1.63	71	3.46	1.43	63	4.15	1.62	60	3.74	1.62	58	4.08	1.70	54	3.59	1.62	60	4.17	1.72	61
	Emotional	3.37	1.28	76	3.62	1.46	71	3.00	1.21	63	3.41	1.22	60	3.06	1.26	58	3.26	1.35	54	2.97	1.23	60	3.19	1.34	61
	Hair cortisol	6.71	5.45	53	5.75	3.12	52	5.53	3.30	51	5.33	3.04	44	5.21	2.93	57	4.44	2.17	49	5.16	4.79	53	5.51	3.35	50
Quality of care	Medical errors	2.77	0.73	76	2.75	0.84	71	2.61	0.73	63	2.48	0.72	60	2.45	0.77	58	2.49	0.78	54	2.34	0.79	60	2.32	0.83	61
	Absenteeism	3.24	4.60	41	3.53	4.60	39													6.09	9.53	43	5.38	9.57	40
	Patients' ratings																								
	Empathic	6.44	0.68	51	6.52	0.64	47							6.62	0.47	27	6.44	0.63	26						
	Attentive	6.34	0.72	51	6.54	0.63	47							6.47	0.63	27	6.35	0.67	26						
	Competent	6.46	0.63	51	6.53	0.59	47							6.60	0.58	27	6.50	0.48	26						
	Satisf. physician	6.49	0.61	51	6.56	0.54	47							6.59	0.59	27	6.50	0.47	26						
	Satisf. care	6.52	0.61	51	6.65	0.46	47							6.62	0.59	27	6.60	0.45	26						
	Supervisors' ratings																								
	Empathic	6.12	0.91	51	6.04	1.14	45							6.29	0.75	29	5.50	1.53	30						
	Attentive	6.10	1.12	51	6.07	1.10	45							6.43	0.68	29	5.87	1.57	30						
	Tense	3.02	1.53	51	2.98	1.44	45							2.62	1.50	29	3.13	1.50	30						
	Colleagues' ratings																								
	Empathic	6.02	1.17	54	5.90	1.12	49							6.39	0.70	32	5.76	1.39	33						
	Attentive	**5.72**	1.27	54	**6.20**	0.93	49							6.38	0.87	32	5.94	1.34	33						
	Tense	3.31	1.38	54	3.37	1.63	49							3.28	1.84	32	3.73	1.61	33						

Bold indicates statistical significance (two-tailed, $p < 0.05$).

Table 3. *Adjusted between-group effect estimates as the interaction of group and time*

		After 2 months (t1)			After 6 months (t2)			After 12 months (t3)			Overall		
	Measure	MD (95% CI)[a]	*p*	*d*[b]	MD (95% CI)[a]	*p*	*d*[b]	MD (95% CI)[a]	*p*	*d*[b]	*F*	df	*p*[c]
Distress	Burnout	−2.53 [−6.43, 1.38]	0.206	0.17	−4.81 [−9.52, −0.11]	**0.046**	0.32	−3.32 [−8.22, 1.58]	0.185	0.22	1.36	(3,350)	0.256
	Personal	−3.16 [−7.49, 1.18]	0.155	0.19	−4.02 [−9.11, 1.07]	0.123	0.23	−3.74 [−8.95, 1.46]	0.159	0.22	1.05	(3,350)	0.369
	Work-related	−2.25 [−6.92, 2.41]	0.345	0.13	−5.38 [−10.92, 0.17]	0.058	0.30	−4.07 [−9.77, 1.64]	0.164	0.24	1.24	(3,350)	0.294
	Client-related	−1.90 [−7.77, 3.96]	0.525	0.10	−4.79 [−11.68, 2.11]	0.175	0.25	−2.03 [−9.08, 5.02]	0.573	0.11	0.65	(3,350)	0.585
	Mental distress	−1.86 [−4.14, 0.41]	0.109	0.35	−1.14 [−3.61, 1.34]	0.369	0.20	−1.42 [−3.86, 1.02]	0.254	0.25	0.94	(3,350)	0.419
	Perceived stress	−1.76 [−3.49, −0.04]	**0.046**	0.31	−0.44 [−2.48, 1.60]	0.672	0.07	−1.20 [−3.28, 0.89]	0.262	0.20	1.71	(3,350)	0.164
	Depression	−0.22 [−0.57, 0.13]	0.216	0.20	−0.10 [−0.51, 0.30]	0.618	0.09	0.01 [−0.40, 0.42]	.962	−0.01	0.65	(3,350)	0.581
	Anxiety	−0.25 [−0.69, 0.20]	0.277	0.18	−0.12 [−0.63, 0.38]	0.636	0.08	−0.08 [−0.57, 0.41]	0.762	0.05	0.47	(3,350)	0.706
	Perceived job strain	−0.36 [−0.70, −0.01]	**0.044**	0.29	−0.06 [−0.44, 0.31]	0.737	0.05	−0.14 [−0.51, 0.24]	0.474	0.11	1.53	(3,350)	0.205
	Cognitive	−0.50 [−0.95, −0.06]	**0.026**	0.33	−0.08 [−0.58, 0.42]	0.765	0.05	−0.32 [−0.83, 0.18]	0.207	0.20	2.20	(3,350)	0.087
	Emotional	−0.27 [−0.64, 0.10]	0.153	0.20	−0.05 [−0.45, 0.35]	0.790	0.04	−0.02 [−0.42, 0.37]	0.914	0.02	0.82	(3,350)	0.482
	Hair cortisol	−0.62 [−2.56, 1.32]	0.531	0.14	−0.19 [−1.90, 1.52]	0.829	0.04	−1.12 [−3.07, 0.82]	0.260	0.24	0.75	(3,269)	0.523
Quality of Care	Medical errors	0.04 [−0.18, 0.26]	0.690	−0.05	−0.03 [−0.28, 0.21]	0.797	0.04	−0.04 [−0.29, 0.21]	0.750	0.05	0.19	(3,350)	0.907
	Patients' ratings												
	Empathic				0.24 [−0.08, 0.55]	0.150	0.38						
	Attentive				0.31 [−0.09, 0.71]	0.134	0.46						

(*Continued*)

Table 3. *Continued*

	After 2 months (t1)			After 6 months (t2)			After 12 months (t3)			Overall		
Measure	MD (95% CI)[a]	*p*	*d*[b]	MD (95% CI)[a]	*p*	*d*[b]	MD (95% CI)[a]	*p*	*d*[b]	*F*	df	*p*[c]
Competent				0.13 [−0.23, 0.49]	0.485	0.22						
Satisf. physician				0.13 [−0.22, 0.48]	0.465	0.23						
Satisf. care				0.15 [−0.19, 0.48]	0.404	0.27						
Supervisors' ratings												
Empathic				0.73 [0.06, 1.40]	**0.037**	0.71						
Attentive				0.54 [−0.17, 1.25]	0.138	0.49						
Tense				−0.49 [−1.31, 0.33]	0.245	0.33						
Colleagues' ratings												
Empathic				0.40 [−0.21, 1.01]	0.203	0.36						
Attentive				0.95 [0.30, 1.60]	**0.006**	0.85						
Tense				−0.39 [−1.37, 0.59]	0.439	0.24						
Absenteeism							0.88 [−2.94, 4.70]	0.654	0.12			

Bold indicates statistical significance (two-tailed, $p < .05$).

[a]MD = Adjusted mean difference between groups in change from baseline (i.e., the difference of differences).

[b]Cohen's *d* as MD divided by the pooled standard deviation at baseline. Values indicating improvement are coded to be positive; values indicating worsening are coded to be negative.

[c]*p*-Values for the overall interaction effect include all measurements from baseline.

Table 4. *Program satisfaction*

Satisfaction domains and items ($n = 59$)	Mean (SD)	Min–Max
Overall satisfaction (1 = very dissatisfied, 7 = very satisfied)		
How satisfied were you with the program?	6.05 (1.21)	1–7
How satisfied were you with the trainers?	6.44 (0.86)	4–7
Benefit (1 = great harm, 7 = great benefit)		
How do you rate your professional benefit from the program?	5.97 (0.87)	4–7
How do you rate your private benefit from the program?	6.2 (0.85)	4–7
Recommendation (1 = no, 7 = yes)		
Would you recommend the program?	6.22 (1.02)	3–7

gram. The mean attendance was 6.64 out of a total of nine sessions (SD = 1.75). The mean attendance at the three maintenance sessions was 1.33 sessions (SD = 1.02). However, for 12 participants, information on attendance at the maintenance sessions was not available. Harms from the MBP were not reported.

Discussion

This RCT compared an 8-week mindfulness-based program that was tailored to the needs of resident physicians with active control. After a 4-month maintenance phase at t2 (i.e., the primary endpoint), the intervention group that participated in the MBP showed greater reductions in levels of burnout than the control group. This effect was small. Moreover, at t1, the intervention group showed greater reductions in perceived stress and perceived job strain. These differences were also small. At t2, the intervention group showed greater improvements in supervisor rated empathy and colleague rated attentiveness. These differences were medium and large. There was no significant difference between the groups in the other outcomes of distress, hair cortisol and the quality of care at any single point in time, or across all time periods. The results corroborated other findings on the feasibility of the tailored MBP [40] as attendance was high and participants were satisfied with the program and the trainers, perceived benefits from the program, and expressed their willingness to recommend it.

Regarding burnout reduction, we did not find an advantage of the tailored MBP over the control directly after the 8-week program at t1, but we found a significant advantage after a 4-month maintenance phase at t2. The experience-based approach to mindfulness, as practised in the intervention group, may need time and repetition for the true benefits in reducing burnout to be seen compared to the description-based approach, as undergone by the control group. However, the advantage in reducing burnout of the MBP over the control at t2 was small and no longer significant at t3 (i.e., 12 months after baseline). It is possible that reducing burnout in resident physicians is particularly difficult [34]. One reason might be that burnout, once present in resident physicians, tends to persist [69]. Likewise, interventions aiming at reducing physician burnout are generally found to have only modest effects [70]. Moreover, it has been argued that individual-directed interventions such as the tailored MBP are less effective in reducing physician burnout than organization-directed interventions because physician burnout is assumed to be primarily rooted in system-level problems [31]. However, the observed difference between groups in burnout reduction at t2 equals a 2.6-point difference when converted into emotional exhaustion on Maslach Burnout Inventory [10]. It is known from the literature on physicians' health that each 1-point increase in emotional exhaustion on Maslach Burnout Inventory is associated with a 7% greater likelihood of reporting suicidal ideation [71], a 5%–7% increase in the likelihood of reporting a medical error [20,72,73], and a 43% greater likelihood of reductions in a professional effort [74]. Therefore, even small reductions in burnout as observed in this study may be considered as meaningful differences. Other controlled trials that evaluated the effectiveness of MBPs for resident physicians are scarce. In accordance with our results, two studies found no advantage of an MBP over control in reducing burnout directly after the program [34,37], while a third study found a medium advantage [38]. The inconsistencies across studies may partly be attributed to differences in study design, but this warrants further

research to determine moderating and mediating variables of burnout reduction in MBPs for resident physicians [75].

Regarding secondary distress outcomes, the results suggest that the tailored MBP had a significant small advantage over the control in reducing perceived stress directly after the 8-week program at t1. This finding matches the conceptual basis of the tailored MBP in MBSR [45] and aligns with two other studies [37,38]. The greater reductions in perceived job strain observed at t1 in the intervention group could be explained by participants being better able to detach from work problems while away from work [55]. However, the differences between groups in reducing perceived stress and perceived job strain were no longer significant at t2 and t3, or across all time periods.

Regarding the quality of care, to the best of our knowledge, this was the first study of an MBP for resident and practising physicians to assess third-party ratings of interpersonal aspects of the quality of care. The results suggest that the tailored MBP was more effective in promoting supervisor-rated empathy and colleague-rated attentiveness at t2 than the control. This advantage of the tailored MBP might have resulted from the fact that the intervention group practised the interpersonal aspects of care as part of undergoing the MBP (i.e., role-playing for mindful patient communication). The advantages in third-party ratings of empathy that were observed correspond with a review showing that MBPs are capable of improving self-reported empathy in physicians [35]. Empathy is a key component of the physician–patient relationship and is associated with patient satisfaction, more patient enablement and better clinical outcomes [76,77].

In summary, regarding effect sizes after the maintenance phase, the advantages of the MBP in maintaining and improving interpersonal aspects of quality of care over the control ranged from small to large, whereas the advantages in reducing symptoms of distress were small. There was no significant difference between the two groups in burnout reduction directly after the 8-week program, at follow-up, or across all time periods. Furthermore, there was no significant difference between the groups in other secondary distress outcomes, hair cortisol medical errors, patients' ratings and absenteeism at any single point in time, or across all time periods. Within-group analyses revealed that both groups improved in many of these outcomes. Hence, the non-significant differences between groups in many outcomes and at many points in time might be due to the control group having surprisingly improved in these outcomes more than expected (Tables S2 and S3).

Overall, we found effects of the tailored MBP on some outcomes of distress and quality of care, while we failed to find effects on many others. Accordingly, when balancing the general effectiveness of the tailored MBP, several conclusions are possible. On the one hand, one could argue that, given the greater efforts in the intervention group compared to the control group, the advantages of the tailored MBP appear meagre. On the other hand, one could argue that, despite being limited in number and scale, these reductions in distress are meaningful and were reached through individual effort, that is, without changing the harsh working conditions that contribute to the high prevalence of distress among resident physicians in the first place. Moreover, participants were satisfied with the program, perceived benefits from it, and expressed their willingness to recommend it to others. Hence, an MBP for resident physicians might be beneficial in more general terms of well-being [34]. Finally, the effects of the tailored MBP were noted by the participants' supervisors and colleagues who rated those in the intervention group to be more empathic and attentive compared to those in the control group.

Strengths and limitations

This trial has several methodological strengths including an active control group, a longitudinal design spanning four measurement points, a published and pre-registered study protocol [42] and a multi-method assessment including hair cortisol and third-party ratings by patients, supervisors and colleagues. The trial included a broad sample from different specialisations and clinics and had a low drop-out rate of 18% at follow-up for the primary outcome.

This trial has several methodological limitations. First, we were unable to determine whether the within-group reductions in many distress outcomes in both groups could be attributed to the treatments being effective in substance or to unspecific effects such as Hawthorne, maturation or parallel external events affecting both groups. Future studies should control for such unspecific

effects by including an untreated control next to a treated control group. In particular, it is possible that the mere fact that someone reached out to the resident physicians, publicly acknowledged their problematic work situation and offered possible solutions contributed to the improvements for both groups. Participants in both groups may have taken this attention as an acknowledgement and impetus to take care of their well-being, especially because the medical culture usually supports neglect of self-care and indifference to personal well-being [78]. Second, we did not include a group that engaged in a standard MBP such as MBSR [45]. Future studies might include such a group to assess the effects of our specific program tailoring to meet the needs of resident physicians. Third, participants were self-selected, and the results may have been influenced by selection bias. However, MBPs are believed to be most effective if individuals choose to participate [79]. Therefore, self-selection, in this case, does not infringe on the ecological validity of the trial, and it is a standard and preferred practice for this type of program [80]. In the same vein, blinding of participants to treatment conditions was not possible, and the ensuing treatment diffusion might have led to over or under-reporting. Fourth, we did not control whether the control group actually read and comprehended the course book that was sent to them. Finally, the third-party ratings of interpersonal aspects of the quality of care were possibly biased as resident physicians might have selected patients whose treatment was successful and colleagues and supervisors with whom they shared a good relationship. However, this possibility does not represent a threat to internal validity, as it can be assumed that it affected both groups.

This trial also has several limitations regarding the interpretation and the implications of the findings. First, we did not adjust the alpha levels for multiple testing [67]. This increases the risk of false-positive findings. Hence, the analysis of all outcomes except from the primary outcome should be regarded as exploratory and interpreted with caution, and future studies are needed to confirm these findings. Second, due to researcher allegiance, the advantage of the MBP over the control group might have been overestimated [81,82]. Third, the MBP is time-intense, and resident physicians may be dissuaded from choosing to participate. Nevertheless, mindfulness is not tied to a specific time or place, it is non-invasive in nature, and, once learned, it can easily and flexibly be implemented into daily life, making it attractive to busy practitioners such as resident physicians [38]. Finally, our program focuses on individual efforts to reduce distress and to improve markers of the quality of care. Although this seems beneficial to some individuals, mindfulness is not a panacea, and structural changes are needed to address the systemic problems rooted in medical training that contribute to the high prevalence of distress among resident physicians in the first place [4]. Otherwise, individual programs such as MBPs run the risk of stabilising the stress-generating current healthcare system by making its individuals more stress-resistant and personally responsible to deal with the consequences of stressors inherent in the current healthcare system. If dealing with distress is solely regarded as a personal responsibility, affected physicians may not be supported but blamed for not being resilient enough [31]. Nevertheless, structural changes often lie beyond the sphere of personal influence and take effect at a slower pace. Therefore, it is important to provide resident physicians with individual strategies such as mindfulness to prevent or mitigate distress.

Conclusion

Many resident physicians suffer from distress, which endangers not only their individual health but also the quality of care they provide. The results of this RCT suggest that a tailored MBP for resident physicians is more effective in reducing burnout on a medium time scale than an active control and might be more effective in reducing certain symptoms of distress and improving certain interpersonal aspects of the quality of care. However, more research is needed to confirm results, to determine mediating and moderating variables of the MBP's effectiveness as well as its sustainability and to disentangle program effects from unspecific effects.

Acknowledgements

We thank the German Research Foundation (project number 197396619 – SFB 1015) and the Collaborative Research Center 1015 *Otium. Boundaries, Chronotopes, Practices* for funding and supporting the project. We would especially thank Susanne Bregulla-Kuhn, Klaus Kuhn and Marc Loewer for their creativity and exceptional guidance during the design process and their outstanding commitment during the implementation of the program and the subsequent final adaptations. We also thank Paul Freihaut, Susanne Rausch, Elina

Kraemer, Dominik Ludwig, Johannes Bürkle, Charlotte Schulze-Marmeling, Sophie Lentz, Lisa Ullrich, Melina von Wernsdorff, Nadine Henrici and Christoph Rockstroh for their valuable assistance.

Conflict of interest

The authors have no conflict of interest to declare.

Author contributions

All authors contributed to the design of the program and the study. SS and ASG developed the initial concept, secured funding and supervised all stages of the trial. JCF and VMA implemented the trial, handled the participants, recorded all data and conducted the data analyses. JSF wrote the manuscript. ASG, SS and VMA made valuable revisions to the manuscript. All authors approved the final version of the manuscript.

References

1 Prins JT, Gazendam-Donofrio SM, Tubben BJ, van der Heijden FMMA, van de Wiel HBM, Hoekstra-Weebers JEHM. Burnout in medical residents: a review. *Med Educ.* 2007;**41**(8):788–800.
2 West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. *J Intern Med.* 2018;**283**(6):516–29.
3 Wehkamp K-H, Naegler H. The commercialization of patient-related decision making in hospitals. Dtsch Aerzteblatt Online [Internet]. 2017 Nov 24 [cited 2020 Nov 20]; https://www.aerzteblatt.de/10.3238/arztebl.2017.0797
4 Dyrbye LN, West CP, Satele D, Boone S, Tan L, Sloan J, et al. Burnout among U.S. medical students, residents, and early career physicians relative to the general U.S. population. *Acad Med.* 2014;**89**(3):443–51.
5 Stucky ER, Dresselhaus TR, Dollarhide A, Shively M, Maynard G, Jain S, et al. Intern to attending: assessing stress among physicians. *Acad Med.* 2009;**84**(2):251–7.
6 Panagopoulou E, Montgomery A, Benos A. Burnout in internal medicine physicians: differences between residents and specialists. *Eur J Intern Med.* 2006;**17**(3):195–200.
7 Heinke W, Dunkel P, Brähler E, Nübling M, Riedel-Heller S, Kaisers UX. Burn-out in der Anästhesie und Intensivmedizin. Gibt es ein Problem in Deutschland? *Anaesthesist* 2011;**60**(12):1109–18.
8 Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Stress and coping among orthopaedic surgery residents and faculty. *J Bone Joint Surg.* 2004;**86**(7):1579–86.
9 Kristensen TS, Borritz M, Villadsen E, Christensen KB. The Copenhagen Burnout Inventory: a new tool for the assessment of burnout. *Work Stress.* 2005;**19**(3):192–207.
10 Maslach C, Jackson SE, Leiter MP. *The Maslach Burnout Inventory.* 3rd ed. Palo Alto, CA: Consulting Psychologists Press; 1996.
11 Lebares CC, Guvva EV, Ascher NL, O'Sullivan PS, Harris HW, Epel ES. Burnout and stress among US surgery residents: psychological distress and resilience. *J Am Coll Surg.* 2018;**226**(1):80–90.
12 Lebares CC, Guvva EV, Desai A, Herschberger A, Ascher NL, Harris HW, et al. Key factors for implementing mindfulness-based burnout interventions in surgery. *Am J Surg.* 2019;**219**(2):328–34.
13 Mata DA, Ramos MA, Bansal N, Khan R, Guille C, Angelantonio ED, et al. Prevalence of depression and depressive symptoms among resident physicians. A systematic review and meta-analysis. *JAMA.* 2015;**314**(22):2373–83.
14 Tyssen R, Hem E, Gude T, Grønvold NT, Ekeberg Ø, Vaglum P. Lower life satisfaction in physicians compared with a general population sample: a 10-year longitudinal, nationwide study of course and predictors. *Soc Psychiatry Psychiatr Epidemiol.* 2009;**44**(1):47–54.
15 Lebensohn P, Dodds S, Benn R, Brooks AJ, Birch M, Cook P, et al. Resident wellness behaviors: relationship to stress, depression, and burnout. *Fam Med.* 2013;**45**(8):541–9.
16 van der Heijden F, Dillingh G, Bakker A, Prins J. Suicidal thoughts among medical residents with burnout. *Arch Suicide Res.* 2008;**12**(4):344–6.
17 Kemper KJ, Schwartz A, Wilson PM, Mahan JD, Schubert CJ, Staples BB, et al. Burnout in pediatric residents: three years of national survey data. *Pediatrics* 2020;**145**(1):e20191030.
18 West CP, Tan AD, Shanafelt TD. Association of resident fatigue and distress with occupational blood and body fluid exposures and motor vehicle incidents. *Mayo Clin Proc.* 2012;**87**(12):1138–44.
19 Prins JT, van der Heijden FM, Hoekstra-Weebers JE, Bakker AB, van de Wiel HB, Jacobs B, et al. Burnout, engagement and resident physicians' self-reported errors. *Psychol Health Med.* 2009;**14**(6):654–66.
20 West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, Habermann TM, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA.* 2006;**296**(9):1071–1078.
21 Baer TE, Feraco AM, Sagalowsky ST, Williams D, Litman HJ, Vinci RJ. Pediatric resident burnout and attitudes toward patients. *Pediatrics* 2017;**139**(3):e20162163.
22 Kang E-K, Lihm H-S, Kong E-H. Association of intern and resident burnout with self-reported medical errors. *Korean J Fam Med.* 2013;**34**(1):36–42.
23 de Oliveira Jr. GS, Chang R, Fitzgerald PC, Almeida MD, Castro-Alves LS, Ahmad S, et al. The prevalence of burnout and depression and their association with adherence to safety and practice standards: a survey of United States anesthesiology trainees. *Anesth Analg.* 2013;**117**(1):182–93.
24 West CP, Shanafelt TD, Kolars JC. Quality of life, burnout, educational debt, and medical knowledge among internal medicine residents. *JAMA.* 2011;**306**(9):952–960.
25 Prins JT, Hoekstra-Weebers JEHM, Gazendam-Donofrio SM, Dillingh GS, Bakker AB, Huisman M, et al. Burnout and engagement among resident doctors in the Netherlands: a national study. *Med Educ.* 2010;**44**(3):236–47.
26 Park C, Lee YJ, Hong M, Jung C-H, Synn Y, Kwack Y-S, et al. A multicenter study investigating empathy and burnout characteristics in medical residents with various specialties. *J Korean Med Sci.* 2016;**31**(4):590–7.
27 Ferreira S, Afonso P, Ramos M do R. Empathy and burnout: a multicentre comparative study between residents and specialists. *J Eval Clin Pract.* 2020;**26**(1):216–22.

28 Pereira-Lima K, Loureiro SR. Burnout, anxiety, depression, and social skills in medical residents. *Psychol Health Med.* 2015;**20**(3):353–62.
29 Han S, Shanafelt TD, Sinsky CA, Awad KM, Dyrbye LN, Fiscus LC, et al. Estimating the attributable cost of physician burnout in the United States. *Ann Intern Med.* 2019;**170**(11):784–90.
30 Shanafelt TD. Enhancing meaning in work. A prescription for preventing physician burnout and promoting patient-centered care. *JAMA.* 2009;**302**(12):1338–40.
31 Panagioti M, Panagopoulou E, Bower P, Lewith G, Kontopantelis E, Chew-Graham C, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med.* 2016;**177**(2):195–205.
32 Kabat-Zinn J. Coming to our senses: healing ourselves and the world through mindfulness. London: Hachette; 2005. p. 736
33 Fendel JC, Bürkle JJ, Göritz AS. Mindfulness-based interventions to reduce burnout and stress in physicians: a systematic review and meta-analysis. *Acad Med.* 2021;**96**(5):751–64.
34 Verweij H, van Ravesteijn H, van Hooff MLM, Lagro-Janssen ALM, Speckens AEM. Mindfulness-based stress reduction for residents: a randomized controlled trial. *J Gen Intern Med.* 2017;**33**(4):429–36.
35 Scheepers RA, Emke H, Epstein RM, Lombarts KMJMH. The impact of mindfulness-based interventions on doctors' well-being and performance: a systematic review. *Med Educ.* 2020;**54**(2):138–49.
36 Epstein RM. Mindful practice. *JAMA.* 1999;**282**(9):833–9.
37 Lebares CC, Guvva EV, Olaru M, Sugrue LP, Staffaroni AM, Delucchi KL, et al. Efficacy of mindfulness-based cognitive training in surgery: additional analysis of the mindful surgeon pilot randomized clinical trial. *JAMA Netw Open.* 2019;**2**(5):e194108.
38 Ireland MJ, Clough B, Gill K, Langan F, O'Connor A, Spencer L. A randomized controlled trial of mindfulness to reduce stress and burnout among intern medical practitioners. *Med Teach.* 2017;**39**(4):409–14.
39 Lases SS, Lombarts MJ, Slootweg IA, Arah OA, Pierik EGJM, Heineman E. Evaluating mind fitness training and its potential effects on surgical residents' well-being. A mixed methods pilot study. *World J Surg.* 2016;**40**(1):29–37.
40 Fendel JC, Aeschbach VM, Göritz AS, Schmidt S. A mindfulness program to improve resident physicians' personal and work-related well-being: a feasibility study. *Mindfulness* 2020;**11**(6):1511–9.
41 Dobkin PL, Hickman S, Monshat K. Holding the heart of mindfulness-based stress reduction. Balancing fidelity and imagination when adapting MBSR. *Mindfulness* 2014;**5**(6):710–8.
42 Aeschbach VM-J, Fendel JC, Göritz AS, Schmidt S. Mindfulness-based programme for residents: study protocol of a randomised controlled trial. *BMJ Open.* 2020;**10**(3):e035025.
43 MBSR-MBCT Verband [Internet]. [cited 2020 Nov 27]. Available from: https://www.mbsr-verband.de/
44 Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, CONSORT NPT Group. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Ann Intern Med.* 2017;**167**(1):40.
45 Kabat-Zinn J. *Full catastrophe living: the program of the Stress Reduction Clinic at the University of Massachusetts Medical Center*. New York: Delta; 1990.
46 Berg DD, Divakaran S, Stern RM, Warner LN. Fostering meaning in residency to curb the epidemic of resident burnout: recommendations from four chief medical residents. *Acad Med.* 2019;**94**(11):1675–8.
47 Shanafelt T, Trockel M, Ripp J, Murphy ML, Sandborg C, Bohman B. Building a program on well-being: key design considerations to meet the unique needs of each organization. *Acad Med.* 2019;**94**(2):156–61.
48 Gimmel J, Keiling T. *Konzepte der Muße*. Tübingen: Mohr Siebeck; 2016. p. 104.
49 Hertwig R, Hogarth RM, Lejarraga T. Experience and description: exploring two paths to knowledge. *Curr Dir Psychol Sci.* 2018;**27**(2):123–8.
50 Schmidt S. Mindfulness in East and West–Is it the same? In: Walach H, Schmidt S, Jonas WB, editors. *Neuroscience, consciousness and spirituality*. Dordrecht: Springer; 2011. p. 23–38.
51 Hanebuth D, Aydin D, Scherf T. Burnout and related conditions in managers: a five-year longitudinal study. *Psychol Alltagshandelns.* 2012;**5**(3):4–39.
52 Goldberg DP, Williams P. *User's guide to the general health questionnaire*. Windsor: NFER-Nelson; 1988.
53 Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav.* 1983;**24**(4):385–96.
54 Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ–4. *Psychosomatics.* 2009;**50**(6):613–21.
55 Mohr G, Rigotti T, Müller A. Irritation - ein Instrument zur Erfassung psychischer Beanspruchung im Arbeitskontext. Skalen- und Itemparameter aus 15 Studien. *Z Für Arb- Organ AO.* 2005;**49**(1):44–8.
56 Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EFC. Hair cortisol, stress exposure, and mental health in humans: a systematic review. *Psychoneuroendocrinology* 2013;**38**(8):1220–35.
57 Stalder T, Steudte-Schmiedgen S, Alexander N, Klucken T, Vater A, Wichmann S, et al. Stress-related and basic determinants of hair cortisol in humans: a meta-analysis. *Psychoneuroendocrinology* 2017;**77**:261–74.
58 Kirschbaum C, Tietze A, Skoluda N, Dettenborn L. Hair as a retrospective calendar of cortisol production–increased cortisol incorporation into hair in the third trimester of pregnancy. *Psychoneuroendocrinology* 2009;**34**(1):32–7.
59 Dettenborn L, Tietze A, Kirschbaum C, Stalder T. The assessment of cortisol in human hair: associations with sociodemographic variables and potential confounders. *Stress* 2012;**15**(6):578–88.
60 Stalder T, Steudte S, Alexander N, Miller R, Gao W, Dettenborn L, et al. Cortisol in hair, body mass index and stress-related measures. *Biol Psychol.* 2012;**90**(3):218–23.
61 Hojat M, Mangione S, Nasca TJ, Cohen MJ, Gonnella JS, Erdmann JB, et al. The Jefferson scale of physician empathy: development and preliminary psychometric data. *Educ Psychol Meas.* 2001;**61**(2):349–65.
62 Teasdale JD, Segal ZV, Williams JM, Ridgeway VA, Soulsby JM, Lau MA. Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *J Consult Clin Psychol.* 2000;**68**(4):615–23.

63 Khoury B, Sharma M, Rush SE, Fournier C. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. *J Psychosom Res*. 2015;**78**(6):519–28.
64 Saghaei M. QMinim Online Minimization [Internet]. [cited 2020 Feb 11]. http://rct.mui.ac.ir/q/index.php
65 Altman DG, Bland JM. Treatment allocation by minimization. *BMJ*. 2005;**330**:843.
66 Pinheiro J, Bates D. *Mixed-effects models in S and S-PLUS*. New York: Springer; 2000.
67 Feise RJ. Do multiple outcome measures require p-value adjustment? *BMC Med Res Methodol*. 2002;**2**(1):8.
68 Morris SB. Estimating effect sizes from pretest-posttest-control group designs. *Organ Res Methods*. 2008;**11**(2):364–86.
69 Kocalevent R, Pinnschmidt H, Selch S, Nehls S, Meyer J, Boczor S, et al. Burnout is associated with work-family conflict and gratification crisis among German resident physicians. *BMC Med Educ*. 2020;**20**(1):145.
70 West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *Lancet* 2016;**388**(10057):2272–81.
71 Shanafelt TD, Balch CM, Dyrbye L, Bechamps G, Russell T, Satele D, et al. Special report: suicidal ideation among American surgeons. *Arch Surg*. 2011;**146**(1):54–62.
72 West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. *JAMA*. 2009;**302**(12):1294–1300.
73 Shanafelt TD, Balch CM, Bechamps G, Russell T, Dyrbye L, Satele D, et al. Burnout and medical errors among American surgeons. *Ann Surg*. 2010;**251**(6):995–1000.
74 Shanafelt TD, Mungo M, Schmitgen J, Storz KA, Reeves D, Hayes SN, et al. Longitudinal study evaluating the association between physician burnout and changes in professional work effort. *Mayo Clin Proc*. 2016;**91**(4):422–31.
75 Kalani SD, Azadfallah P, Oreyzi H, Adibi P. Interventions for physician burnout: a systematic review of systematic reviews. *Int J Prev Med*. 2018;**9**(1):81.
76 Lamothe M, Rondeau É, Malboeuf-Hurtubise C, Duval M, Sultan S. Outcomes of MBSR or MBSR-based interventions in health care providers. A systematic review with a focus on empathy and emotional competencies. *Complement Ther Med*. 2016;**24**:19–28.
77 Derksen F, Bensing J, Lagro-Janssen A. Effectiveness of empathy in general practice: a systematic review. *Br J Gen Pract*. 2013;**63**(606):e76–84.
78 Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;**374**(9702):1714–1721.
79 Walach H, Schmidt S, Esch T. Meditation intervention reviews. *JAMA Intern Med*. 2014;**174**(7):1193–4.
80 Burton A, Burgess C, Dean S, Koutsopoulou GZ, Hugh-Jones S. How effective are mindfulness-based interventions for reducing stress among healthcare professionals? A systematic review and meta-analysis. *Stress Health*. 2017;**33**(1):3–13.
81 Munder T, Brütsch O, Leonhart R, Gerger H, Barth J. Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clin Psychol Rev*. 2013;**33**(4):501–11.
82 Goldberg SB, Tucker RP. Allegiance effects in mindfulness-based interventions for psychiatric disorders: a meta-reanalysis. *Psychother Res*. 2020;**30**(6):753–62.

Correspondence: Johannes Caspar Fendel, Medical Faculty, Department for Psychosomatic Medicine and Psychotherapy, Medical Centre, University of Freiburg, Hauptstrasse 8, 79104 Freiburg, Germany.
Email: johannes.fendel@uniklinik-freiburg.de

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supporting Information ■

