Medical App Treatment of Non-Specific Low Back Pain in the 12-month Cluster-Randomized Controlled Trial Rise-uP: Where Clinical Superiority Meets Cost Savings

Janosch A Priebe, Linda Kerkemeyer, Katharina K Haas, Katharina Achtert, Leida F Moreno Sanchez, Paul Stockert, Maximilian Spannagl, Julia Wendlinger, Reinhard Thoma, Siegfried Ulrich Jedamzik, Jan Reichmann, Sebastian Franke, Leonie Sundmacher, Volker E Amelung, Thomas R Toelle

Purpose: Non-specific low back pain (NLBP) exerts a profound impact on global health and economics. In the era of Web 3.0, digital therapeutics offer the potential to improve NLBP management. The Rise-uP trial introduces a digitally anchored, general practitioner (GP)-focused back pain management approach with the Kaia back pain app as the key intervention. Here, we present the 12-months evaluation of the Rise-uP trial including clinical and economic outcomes, patient satisfaction and behavioral tracking analysis.

Methods: The cluster-randomized controlled study (registration number: DRKS00015048) enrolled 1237 patients, with 930 receiving treatment according to the Rise-uP approach and 307 subjected to standard of care treatment. Assessments of pain, psychological state, functional capacity, and well-being (patient-reported outcome measures; PROMs) were collected at baseline, and at 3-, 6-, and 12-months follow-up intervals. Health insurance partners AOK, DAK, and BARMER provided individual healthcare cost data. An artificial intelligence (AI)-driven behavioral tracking analysis identified distinct app usage clusters that presented all with about the same clinical outcome. Patient satisfaction (patient-reported experience measures; PREMs) was captured at the end of the trial.

Results: Intention-to-treat (ITT) analysis demonstrated that the Rise-uP group experienced significantly greater pain reduction at 12 months compared to the control group (IG: −46% vs CG: −24%; p < 0.001) with only the Rise-uP group achieving a pain reduction that was clinically meaningful. Improvements in all other PROMs were notably superior in patients of the Rise-uP group. The AI analysis of app usage discerned four usage clusters. Short- to long-term usage, all produced about the same level of pain reduction. Cost-effectiveness analysis indicated a substantial economic benefit for Rise-uP.

Conclusion: The Rise-uP approach with a medical multimodal back pain app as the central element of digital treatment demonstrates both, clinical and economic superiority compared to standard of care in the management of NLBP.

Keywords: digital medicine, medical apps, non-specific low back pain, multimodal pain therapy, healthcare costs, behavioral tracking analysis

Introduction

Digital solutions, particularly medical apps, have emerged as effective treatments for non-specific low back pain (NLBP). NLBP is defined as back pain without any identifiable pathoanatomical cause. Our team has demonstrated the efficacy of the Kaia back pain app through retrospective cohort studies and a randomized controlled trial (RCT), highlighting the beneficial role of digital interventions in managing back pain.
The Rise-uP approach (Rücken-innovative Schmerztherapie mit eHealth für unsere Patienten) was developed to establish a comprehensive digital treatment approach of NLBP for real-world settings, incorporates digital components such as a shared electronic case report form (eCRF), teleconsultations between general practitioners (GPs) and pain specialists for patients at risk for development of chronic back pain, and the multimodal Kaia back pain app that provides physical exercises, mindfulness training, and educational content. Three-month data from our cluster-randomized controlled trial have shown that patients treated with the Rise-uP approach experience less pain, greater reductions in psychological symptoms and improved functional ability and wellbeing compared to those receiving standard care, especially for those at high risk for developing chronic pain after 3 months.10

These advancements are particularly significant in light of the challenges posed by NLBP, which not only affects patients’ quality of life but also incurs considerable costs on global healthcare systems due to unnecessary imaging, pharmacological treatments, and surgeries which is against recommendation of guidelines. These guidelines demand conservative approaches with physical exercise, psychological intervention and education in order to empower the patient.15–17 Despite the common belief that NLBP mostly resolves spontaneously within 2–3 months, research shows that a substantial number of patients continue to suffer from pain after one year, underscoring the need for effective, guideline-conform treatment strategies that address biological, psychological, and social factors involved in the development and maintenance of pain.11–14,18,19

Furthermore, during the pandemic and in times of social distancing, remote treatment options for back pain patients were often the only alternative, as pain centers were shut down. Digital strategies, such as the Rise-uP approach and the empowerment of patients through the Kaia back pain app, offer a way to mitigate the treatment shortfall during such times.20,21

In conclusion, our final evaluation of the Rise-uP trial aims to confirm the sustained superiority of the Rise-uP approach over time based on the patient-reported outcome measures (PROMs), assess its cost-effectiveness compared to standard care, evaluate patient satisfaction through patient-reported experience measures (PREMs), and explore app usage patterns by skilled behavioral tracking analyses.

**Methods**

**Study Design and Participants**
- The Rise-uP approach and trial has been described elsewhere in detail.10 Rise-uP is a two-armed cluster-randomized controlled trial with the first arm “Rise-uP” (intervention group, IG) and second arm “control group” (CG). Patients were recruited via two channels: (1) Recruitment by participating GPs and (2) recruitment via Facebook advertising.10

**Inclusion Criteria of the Study Were**
- NLPB (ICD 10 Code M40-M54) in an acute stage (up to 6 weeks) or subacute stage (6–12 weeks).
- six episodes of recurrent NLBP at max lasting no longer than 12 weeks (end of preceding episode at least 6 months before the current episode).
- age between 18 and 65 years.
- email access for completing the follow-up questionnaires.
- access to a smartphone or tablet for using the Kaia back pain app in the Rise-uP group
- adequate knowledge of the German language.
- member of the statutory health insurances AOK Bayern, BARMER or DAK.

**Exclusion Criteria**
- other kinds of back pain, eg, chronic LBP.
- history of back- or vertebral surgery.
- LBP of any specific cause requiring treatment (for example, fractures or tumors).
- any other serious medical conditions (for example, severe heart failure).
- acute psychiatric disorders (for example, acute schizophrenia).
Randomization and Blinding

Health insurance networks and word-of-mouth recommendations served as conduits to acquaint GPs within the Bavarian region of Southern Germany with the Rise-uP project. The process of randomization was facilitated through an electronic algorithm, details of which are elaborately discussed in Priebe et al. A randomization ratio of 2:1 was intentionally selected to anticipate a patient ratio of 3:1 favoring the Rise-uP arm over the control, reflecting a greater interest in the intervention arm, a trend consistent with historical recruitment patterns. Moreover, this would allow more specific analysis of subgroups of patients in the Rise-uP arm.

The responsibility of training GPs in the utilization of digital tools for this study was guaranteed by the administrative staff of the Center for Interdisciplinary Pain Medicine, Klinikum rechts der Isar (MRI), Technical University of Munich. A comprehensive software suite, provided by StatConsult—a consortium partner specializing in IT—was installed on computers within the practitioners’ offices. This suite included an eCRF, a treatment algorithm and interfaces for both the upload of patient data from tablets during baseline assessments and the facilitation of communication between the Rise-uP head office and collaborating specialists. Teleconsultations were conducted using a Cisco DX-80 videoconferencing system, with Netconnex GmbH supplying the necessary connection services.

Physicians in the CG received no specific information about the Rise-uP protocol. While neither the physicians nor the patients were blinded to the group assignments, they remained uninformed about the specific details pertaining to the counterpart group, maintaining a level of impartiality in the study’s execution. Recruitment of patients started in August 2017 and was closed March 2019. Last follow-up was assessed March 2020.

Intervention

The German National Guideline for the treatment of NLBP was taken as the basis of the Rise-uP approach. Yet, we deviated from guidelines in one relevant issue: The guideline demands determination of risk for developing chronic pain after 4 weeks of failed treatment via STarT-Back questionnaire. In contrast, since Rise-uP focused on early intervention for preventing chronic pain, this risk was determined immediately at the beginning of the study inclusion. In case of a high risk, the GP was offered a teleconsultation with a pain specialist from the medical staff at the Rise-uP head office or the Algesiologikum pain clinic. In the teleconsultation, the medical record of the patient as well as the next appropriate treatment steps were discussed.

As the key intervention, patients in the Rise-uP group were granted access to the Kaia back pain app with the instruction to use the app as frequent as possible. The Kaia back pain app has been described elsewhere in detail. Furthermore, see Priebe et al for a description of the overall communication and data flow of the Rise-uP trial. The basic points of the Rise-uP approach were:

- Clinical investigations including red and yellow flags (STarT Back score) at baseline.
- Re-visitations depending on the risk of development of chronic back pain and clinical progress or improvement.
- Patients were provided access to the Kaia back pain app.
- Electronic questionnaires for follow-up measurements via email.

At the first appointment with a GP, NLBP patients received information about the Rise-uP project and a screening for inclusion and exclusion criteria. After signing the informed consent, a set of questionnaires comprising the determination of the risk of developing chronic pain as well as the outcomes was completed on a tablet and transferred to the physician’s computer and to the central database at StatConsult.

Standard of care considering the adherence to the national guideline was provided to CG patients.

In order to compensate patients for participations, vouchers totaling 10€ were provided to patients for answering a set of questionnaires at baseline (T0) and at the follow-ups after 3 months (T1), 6 months (T2) and 12 months (T3). The trial was conducted according to the Declaration of Helsinki. The study protocol was approved by the Ethics Committee of the University Hospital Klinikum rechts der Isar, Technical University of Munich (TUM) (272/17 S) with date from 08/03/2017. First patient was included 09/08/2017. The trial was registered at DRKS (German Clinical Trials Register; WHO Primary Register) with
registration number DRKS0001548 with date from 12/17/2018. Last patient was recruited into the study at 06/18/2019. Last data collection from last patient in was after 12-month observation period and was performed with date from 06/18/2020.

Outcomes
Pain intensity (primary outcome), functional ability, psychopathological and wellbeing parameters as well as pain graduation (secondary outcomes) were assessed as clinical outcome parameters (PROMs) at baseline (T0) as well as at 3-month (T1), 6-month (T2) and 12-month follow-up (T3) via self-report questionnaires. Furthermore, the participating health insurances AOK, DAK and BARMER provided individual healthcare cost data. As a PREM, patients’ treatment satisfaction was assessed.

Questionnaires
An 11-point numeric rating scale (NRS; 0 = no pain, 10 = unbearable pain) was used to assess current pain, as well as maximum and average pain over the last 4 weeks period. For the calculation of the pain index, the current, maximum and average pain intensity were averaged.25,26 For measuring psychopathological symptoms, the Depression-Anxiety-Stress-Scale (DASS)27,28 was applied. Assessing functional ability was conducted using the Hannover Functional Ability Questionnaire (HFAQ, FFbH-R “Funktionsfragebogen Hannover Rücken”).29 The Veterans RAND 12-Item Health Survey (VR-12)30 was used to assess mental and physical wellbeing. By application of the Graded Chronic Pain Status, patient pain severity was graded into four classes (grade 1: low disability-low intensity; grade 2: high pain intensity and low pain-related disability; grade 3: moderate pain-related disability; grade 4: severely limiting).31,32

The German version of the STarT Back (STarT-G)23,24 and the PainDetect pain-questionnaire (PD-Q)33 were used to determine the risk of developing chronic pain and to estimate the neuropathic pain component. Patient’s satisfaction was assessed in the end of the trial by the ZUF-8 questionnaire.34,35

Rise-uP patients completed the baseline questionnaires on the day of study inclusion via tablet, while paper/pencil was used for the CG. The follow-up questionnaires were sent to patients via Email and completed via the internet in both groups.

Healthcare Cost Data
In order to compare treatment costs between groups, the statutory health insurances AOK, DAK and BARMER provided health cost data for each patient for the observational period and a period of one year before study inclusion (pre-observational period). This allowed a comparison of differences between the costs in the pre-observational and the observational period (difference-in-differences) between groups. Healthcare cost data included overall costs as well as specific costs for inpatient hospital and outpatient treatment, inability to work, as well as information about medication and remedies.

Sample Size
The power analysis was performed on the primary endpoint “pain intensity” at 12 months after study inclusion. Subgroup-analyses had to be considered in the power analysis. With the subgroups risk of developing chronic pain (low, medium, high, CG) and pain duration (acute, subacute, recurrent, CG) a 4 × 4 study plan with 16 subgroups resulted. Therefore, power analysis was based on a two-ways split-plot analysis of variance (ANOVA) with the within-factor time (T0 vs T1 vs T2 vs T3) and the between-factor subgroup (medium effect size, α = 5% and β = 20%). Power analysis revealed 640 patients in total. We conservatively considered a follow-up drop-out rate of 50% for T3 which resulted in a sample size of 1280 patients. Since subgroup analysis is mandatory only in the Rise-uP group, a 3:1 ratio (Rise-uP vs CG) was taken as a basis.

Statistical Analysis
For comparing the pain outcome after 12 months between groups, a split-plot ANOVA with the between-factor “group” (Rise-uP vs CG) and the within-factor “measure point” (T0 vs T1 vs T2 vs T3) was run over NRS pain ratings at T0, T1, T2, and T3. A multiple regression model was used to control for baseline values. A Jump-To-Reference approach with 2000 bootstrap samples was applied for imputing missing values. Cluster effects were tested by calculating the intra-cluster-coefficient (ICC). For this purpose, the primary endpoint data were subjected to a mixed linear model with a random factor “cluster”.

https://doi.org/10.2147/JPR.S473250
DovePress
Journal of Pain Research 2024:17
2242

Priebe et al Dove Press
Powered by TCPDF (www.tcpdf.org)
Furthermore, we compared the percentage of pain reduction between groups. A difference-score was calculated by subtracting the pain value of T0 from the values of T1 or T2 or T3 (Δpain) for each patient. Then, for each measure point, the Δpain scores were divided by the baseline value to determine the percentage of pain reduction (Δ pain %) which controls for baseline values. Then, Δ pain % scores were subjected to split-plot ANOVA with the between-factor “group” (Rise-uP vs CG) and the within-factor “measure point” (T1 vs T2 vs T3) to compare differences in Δpain % scores between both groups at each measure point.

For comparing symptom improvement in the secondary outcomes at the end of the observational period (anxiety, depression, stress, functional ability, mental and physical wellbeing, and pain severity grades), Δ scores for all secondary outcomes were calculated by subtracting T0 values from the values from T3 for each patient. Then, Δ scores of all secondary outcomes were subjected to a multivariate analysis of variance (MANOVA) with the between-factor “group” (Rise-uP vs CG). A potential significant MANOVA main effect was post-hoc tested by univariate analyses of variance (SPSS).

**Behavioral Tracking Analysis**

In order to investigate the relationship between frequency of app usage and pain improvement, user data were extracted from the log-files of the Kaia database. The further analysis code employed a data-driven, machine learning and statistical bottom-up approach to identify complex usage clusters among participants of the Rise-uP study. The process began with data preprocessing, where adherence signals were smoothed and interpolated to handle missing data. This cleaned set of data was used to extract various adherence metrics such as the number of training periods and mean adherence during these periods. Additionally, pain data were processed to provide a continuous pain profile for each user. These metrics formed the basis for training a decision tree classifier, a supervised machine learning model, which categorized users into predefined adherence and pain clusters. In the next step, linear and non-linear dimensionality reduction techniques like Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP) was applied. These techniques reduced the complexity of the data, making it easier to visualize and identify patterns. After dimensionality reduction, K-means clustering, an unsupervised learning algorithm, was applied. This algorithm grouped users into clusters based on their adherence and pain behavior, uncovering more intricate usage patterns that might not have been evident through simple analysis. The analysis code also used statistical methods to detect the relationship between adherence and pain improvement. Additionally, kernel density estimation was used to visualize the distribution of pain changes within clusters.

**Cost Analysis**

A difference-in-differences (DID) approach was calculated to compare pre-post estimates of outcomes and costs for Rise-uP and CG. The DID approach is a method for adjusting for any unobservable determinants that are time-invariant and account for variation in outcomes and costs. The economic evaluation was performed as a cost-effectiveness analysis considering current pain as effect measures and total costs. As a point estimator of cost-effectiveness, we calculated the incremental cost-effectiveness ratio (ICER), that is, the ratio of the differences in mean costs $\overline{C}$ and mean health effects $\overline{E}$ between Rise-uP and standard care during the 12-month follow-up period. Simultaneous bootstrapping of incremental cost and incremental effect estimates addressed estimation uncertainty. Those replications were plotted on a cost-effectiveness plane (CEP).

ZUF-8 scores (PREMs) were compared between groups using a one-tailed t-test for independent samples. Significance level was set $\alpha = 5\%$ for all analyses. Bonferroni-correction was applied. All analyses were run using SPSS, IBM, versions 27 and 28 or the statistical analysis software R®.

**Role of Funding Source**

The study was funded by the German Innovationsfonds (G-BA), grant number: 01NVF16014. The funding institution of the study had no role in study design, data collection, analysis or interpretation and was not involved in the writing process.
Results
Sample Characteristics and Baseline Values
After finalizing physician recruitment, 85 medical practices with 117 affiliated GPs in different rural and urban areas of Bavaria were pooled. Six GPs in 4 medical practices withdrew before randomization for unknown reasons. Then, 81 medical practices with 111 affiliated GPs were randomized either to the Rise-uP or the control group in an approximately 2:1 ratio. Seventy-three Rise-uP GPs (48 medical practices) were randomized to the Rise-uP group, and 38 control GPs (33 practices) were randomized to the CG. Fifteen GPs in 14 medical practices in the Rise-uP group and 10 GPs in 10 medical practices in the CG did not enroll patients for unknown reasons. Thus, 58 GPs (34 medical practices) in the Rise-uP group and 28 GPs (23 medical practices) in the CG contributed to patient inclusion.

Overall, 930 patients were included into the Rise-uP arm (mean inclusion per GP $M = 13$), and 307 patients were included into the CG (mean inclusion per GP $M = 8$). Patient drop-outs are depicted in Figure 1 in detail. In total 524 Rise-uP patients and 226 control patients completed all follow-ups. Yet, 15 of those patients (13 in the Rise-uP group and 2 in the CG) had to be excluded since they showed a T0 pain index = 0. Finally, data of 511 patients in the Rise-uP group and 224 patients in the CG were available. Table 1 provides an overview of the patients’ characteristics of all included patients.

There were no prominent differences regarding demographic and medical parameters between the Rise-uP and the control group. The significantly different parameters showed only small effect size (age: $d = 0.28$; risk to development of chronic pain and pain duration: $V < 0.25$). Additionally, baseline measures of outcome parameters are provided in Table 2.

At baseline, there was neither a between-group difference detected in the primary outcome (pain intensity) nor in anxiety, depression and stress. Small but significant differences occurred in functional ability, mental wellbeing, and physical wellbeing with stronger symptoms in the Rise-uP group compared to the control group. However, in wellbeing parameters, scores of both groups were not in the pathological ranges (both groups within the range of 1 SD of the standardized mean = 50).

Outcomes
Primary Outcome
Pain decreased in both groups over time ($F(3;2178) = 208.75; p < 0.001; \eta = 0.223$). Yet, after 12 months, Rise-uP patients reported less pain compared to the CG ($p=0.005$ [$<0.001$ to $0.035$]) (Figure 2A). This effect remained stable when controlling for baseline characteristics (see Table 3).

Moreover, the Rise-uP patients showed a stronger decrease in pain intensity than the CG (baseline – T1: Rise-uP: −33% vs CG: −14%; baseline – T2: Rise-uP: −39% vs CG: −21%; $p<0.001$; baseline – T3: Rise-uP: −46% vs CG: −24%; $p's <0.001$). It is worth mentioning that the pain reduction in the Rise-uP group exceeded the threshold of −33% which is considered to be clinically relevant$^{26}$, by far, while the control group (−24%) did not pass the 33% threshold. This indicates clinical superiority of the Rise-uP approach (Figure 2B).

In order to test clinical superiority of the Rise-uP approach more directly, an ex-post logistic regression controlling for baseline parameters, with the predictor “group” (Rise-uP vs CG) and the binary criterion “pain reduction” (<33% vs >33%) was computed. Descriptively, 58% of the Rise-uP patients and 47% of the control patients showed a pain reduction >33%. Logistic regression revealed an odds ratio OR = 1.62 ($p < 0.001$) in favor of the Rise-uP group. This reflects a significantly higher proportion of patients with a pain decrease >33% in the Rise-uP group compared to the CG.

Secondary outcomes
The multivariate analyses of variance (MANOVA) over the $\Delta$s of the secondary outcomes showed significant between-group differences for all secondary outcomes after 12 months ($F(6;697) = 10.528; p < 0.001; \eta = 0.083$). Furthermore, there was not only significant higher symptom relief in the Rise-uP group compared to the control group. The control group even deteriorates in anxiety, depression, stress and mental wellbeing (Figure 3).

In summary, analyses of the primary and secondary outcomes underline the superiority of the Rise-uP treatment approach compared to the CG.
Figure 1 Overview of the patients’ flow. From 930 patients in the Rise-uP group and 307 in the CG 511 patients in the Rise-uP group and 224 patients in the CG completed all follow-ups (T1, T2, T3).
Table 1 Sample Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Rise-uP (n=930)</th>
<th>CG (n=307)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>65%</td>
<td>64%</td>
<td>n.s.</td>
</tr>
<tr>
<td>Male</td>
<td>35%</td>
<td>36%</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42.0 (12.4)</td>
<td>37.0 (12.6)</td>
<td>sig.</td>
</tr>
<tr>
<td><strong>Height (cm)</strong></td>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>171 (11)</td>
<td>172 (10)</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Weight (kg)</strong></td>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77.5 (19.7)</td>
<td>77.8 (18.5)</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td>29%</td>
<td>27%</td>
<td>n.s.</td>
</tr>
<tr>
<td>Non-academic</td>
<td>71%</td>
<td>73%</td>
<td></td>
</tr>
<tr>
<td><strong>Employment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>87%</td>
<td>87%</td>
<td>n.s.</td>
</tr>
<tr>
<td>No</td>
<td>13%</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td><strong>Taking pain killer</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>36%</td>
<td>37%</td>
<td>n.s.</td>
</tr>
<tr>
<td>No</td>
<td>64%</td>
<td>63%</td>
<td></td>
</tr>
<tr>
<td><strong>Pain Duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute</td>
<td>52%</td>
<td>31%</td>
<td>sig.</td>
</tr>
<tr>
<td>Subacute</td>
<td>24%</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>Recurrent</td>
<td>24%</td>
<td>46%</td>
<td></td>
</tr>
<tr>
<td><strong>Chronification Risk</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>47%</td>
<td>62%</td>
<td>sig.</td>
</tr>
<tr>
<td>Medium</td>
<td>36%</td>
<td>28%</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>17%</td>
<td>10%</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All patients who were included to the study are displayed except those who withdrew consent resulting in N = 1237 (Rise-uP: 930 patients vs CG: 307 patients). Bold numbers indicate Bonferroni-corrected (α' = 0.01) significance (two-tailed).

Table 2 Symptoms at Baseline

<table>
<thead>
<tr>
<th></th>
<th>Rise-uP vs CG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rise-uP (N=930)</td>
</tr>
<tr>
<td><strong>Pain Intensity (0–10)</strong></td>
<td>5.36</td>
</tr>
<tr>
<td><strong>Anxiety (0–21)</strong></td>
<td>6.29</td>
</tr>
<tr>
<td><strong>Depression (0–21)</strong></td>
<td>8.52</td>
</tr>
<tr>
<td><strong>Stress (0–21)</strong></td>
<td>12.56</td>
</tr>
<tr>
<td><strong>Functional ability (0% - 100%)</strong></td>
<td>71.0%</td>
</tr>
<tr>
<td><strong>Mental wellbeing (M = 50, SD = 10)</strong></td>
<td>44.94</td>
</tr>
<tr>
<td><strong>Physical wellbeing (M = 50, SD = 10)</strong></td>
<td>40.30</td>
</tr>
<tr>
<td>GCPS Grade 1</td>
<td>28%</td>
</tr>
<tr>
<td>GCPS Grade 2</td>
<td>21%</td>
</tr>
<tr>
<td>GCPS Grade 3</td>
<td>47%</td>
</tr>
<tr>
<td>GCPS Grade 4</td>
<td>4%</td>
</tr>
</tbody>
</table>

Notes: Means and SDs for the outcome parameters at T0 (baseline) for the Rise-uP group and the control group (all included patients). Bold numbers indicate Bonferroni-corrected (α' = 0.01) significance (two-tailed).
Analysis of the Relationship of Symptom Improvement and App Usage

Obviously, overall usage of the app decreased after the first 3 months and the physical exercise pillar was preferred compared to mindfulness training and education especially after 3 months (Figure 4).

Table 3 Regression Model Controlling for Baseline Characteristics

<table>
<thead>
<tr>
<th>CG vs Rise-uP</th>
<th>Beta (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>0.72 (0.30 to 1.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age</td>
<td>0.03 (0.02 to 0.05)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Pain Duration</td>
<td></td>
<td>0.083</td>
</tr>
<tr>
<td>Acute</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Subacute</td>
<td>0.12 (−0.32 to 0.56)</td>
<td>0.580</td>
</tr>
<tr>
<td>Recurrent</td>
<td>0.47 (0.05 to 0.88)</td>
<td>0.028</td>
</tr>
<tr>
<td>Chronification Risk</td>
<td>0.051</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>Medium</td>
<td>−0.40 (−0.82 to 0.03)</td>
<td>0.068</td>
</tr>
<tr>
<td>High</td>
<td>−0.80 (−1.5 to −0.10)</td>
<td>0.025</td>
</tr>
<tr>
<td>Wellbeing</td>
<td>0.02 (−0.01 to 0.05)</td>
<td>0.120</td>
</tr>
<tr>
<td>Depression</td>
<td>0.02 (−0.01 to 0.06)</td>
<td>0.220</td>
</tr>
<tr>
<td>Anxiety</td>
<td>0.06 (0.02 to 0.10)</td>
<td>0.004</td>
</tr>
<tr>
<td>Stress</td>
<td>−0.01 (−0.04 to 0.03)</td>
<td>0.700</td>
</tr>
<tr>
<td>Functional Ability</td>
<td>0.01 (0.00 to 0.02)</td>
<td>0.230</td>
</tr>
</tbody>
</table>

Notes: The regression analysis with group (Rise-uP vs CG) as predictor and baseline characteristics as covariates shows still a significant main effect of group. The bold line indicates that the superiority of the Rise-uP approach remains stable when controlling for baseline characteristics. Abbreviation: CI, confidence interval.
Correlation analyses confirmed the previous findings\textsuperscript{10} of no relationship between frequency of Kaia back pain app usage and pain reduction ($\Delta$ pain $\%$) occurring over the whole observational period ($r = -0.042$; $p = 0.355$). Therefore, we used an AI-based bottom-up approach including machine learning with supervised learning (decision tree classifiers) and unsupervised learning (K-means clustering) and various statistical methods for data processing and interpretation to identify more complex usage behaviors.

Figure 3 Illustration of the changes in secondary outcomes both for the Rise-uP group and the CG at T3. Rise-uP patients report significantly higher improvement in all secondary outcomes compared to the CG. Data of 511 Rise-uP patients and 224 CG patients were available for this plot. Statistical testing was conducted as ITT (intention to treat).

Notes: * $p < 0.05$ significancy. $\Delta$FFbH = difference in “Funktions-Fragebogen-Hannover” from baseline to 12-months; $\Delta$VR-12 = difference in wellbeing from baseline to 12-months; $\Delta$DASS = difference in anxiety, depression and stress from baseline to 12-months.

Figure 4 Frequency of active usage days as well as the particular pillars of the Kaia back pain app program over the observation period. Overall usage of the app decreased after the first 3 months and the physical exercise pillar was preferred compared to mindfulness training and education especially after 3 months.

https://doi.org/10.2147/JPR.S473250
DovePress
Journal of Pain Research 2024:17
Two issues are noteworthy. Firstly, the clusters differ regarding the maintenance of adherence over time with only short and low usage in the beginning (Cluster 1) up to adherence over the whole observation time (cluster 4). Secondly, all usage clusters lead to a similar improvement in pain intensity (Figure 5).

Cost-Effectiveness Analysis
The DiD cost analysis of the individual cost components showed a saving of 99.04 € in inpatient costs in the Rise-uP group compared to the CG with NLPB as the main diagnosis ($p = 0.022$). Furthermore, outpatient costs were also

Figure 5: Illustration of the 4 app usage clusters which were identified. Left column: Figures with “day” on the X-axis and “adherence index” on the Y-axis. This figure displays the adherence index (smoothed measure of user engagement) over time (in days) for different users per found cluster. This figure shows patterns of user engagement over time. The variation in adherence patterns can help us differentiate user clusters, such as those with continuous engagement versus those with intermittent or declining engagement. Right column: Figure with “n logs” on the X-axis and “change in Pain Index” on the Y-axis (density plots). This figure presents the relationship between the number of app logs (n logs) and the change in pain values (difference between the pain index at the end of the study and the start). This figure illustrates how user engagement correlates with changes in their pain index. This relationship helps identify which user behaviors are associated with better pain management outcomes.
significantly reduced on average by 20.46 € in the Rise-uP group than in the CG ($p = 0.030$). However, the DiD cost analysis found an average increase of 6.17 € in medication costs for Rise-uP ($p = 0.801$). Instead, the costs for remedies decreased by 12.69 € in favour of Rise-uP ($p = 0.151$). A cost reduction of 120.80 € in favour of Rise-uP was also demonstrated in the costs of incapacity for work ($p = 0.181$). Overall, the analysis of the total costs revealed that the total costs of the Rise-uP group decreased from 411.65 € to 372.45 € per patient within twelve months, while the costs of CG increased by 207.75 €. As a result, the DiD cost analysis yielded a decrease in total costs of 246.95 € (81%) in favour of Rise-uP ($p = 0.021$). A cost-effectiveness analysis showed cost savings of €416.21 per point reduction on the NRS pain scale in favour of the Rise-uP group.

For Rise-uP (98% of all estimation parameters), most of the cost-effect pairs were located in the lower left quadrant of the CEP (Figure 6). In the case of the pain indices, where a lower value indicates a more effective intervention, this quadrant suggests better effects and fewer costs.

**Analysis of Patients’ Satisfaction**

In order to assess patients’ satisfaction with treatment the ZUF-8 questionnaire$^{34,35}$ has been sent out via email. Two hundred and forty-one patients of the Rise-uP group and 106 control patients completed the questionnaire. Patients’ satisfaction was significantly higher in the Rise-uP group ($M=25.38; SD=5.65$) compared to the control group ($M=21.56; SD=6.18$), $p < 0.001$.

**Discussion**

To the best of our knowledge, the Rise-uP trial stands as a pioneering study in the exploration of digital strategies for managing NLBP, incorporating a comprehensive digital framework. This includes a shared eCRF with clinical decision support, teleconsultations between GPs and pain specialists for patients at high risk for chronic pain, and the multimodal...
Kaia back pain app, offering physical exercises, mindfulness training, and educational content. This study pursued two main objectives: (1) to conduct a thorough long-term evaluation of the effects of the Rise-uP approach over the time of 12 month, focusing on primary and secondary outcomes (PROMs and PREMs) and (2) to evaluate the cost-effectiveness of the Rise-uP approach in comparison to standard care practices. Additionally, the study delved into analysis of behavioral tracking of the Kaia back pain app. Firstly, the key findings revealed that the Rise-uP approach demonstrated better clinical outcomes for all measured parameters. Secondly, the Rise-uP approach proved to be significantly more cost-effective than conventional care methods. Lastly, patients treated under the Rise-uP protocol reported greater satisfaction with their treatment than those in the control group.

The findings presented here demonstrate the sustained effectiveness of digital treatment with a medical app over 12 months. Although there was also improvement in the CG over time, this did not reach the clinically significant pain reduction threshold of −33% at any point in time, with a −24% peak in pain reduction after 12 months. In contrast, Rise-uP patients achieved this crucial level of pain reduction as early as the first measurement and reported an average pain reduction of −46% after 12 months. Given the large sample size of the Rise-uP trial, reaching statistical significance was straightforward, underscoring the importance of the clinical relevance of these effects to establish superiority.

The Rise-uP trial also assessed cost parameters and underlined that, from a payer’s perspective, the effectiveness of medical interventions should not only be measured in terms of symptom reduction but also in terms of their cost-effectiveness. The cost analysis identified slight cost savings in the Rise-uP group during the observation period compared to increased costs in the control group during the same period. This resulted in higher cost-effectiveness of the Rise-uP approach compared to standard care, with savings of €416.21 per point reduction on the NRS pain scale. Given the significant financial burden of back pain on global healthcare systems, these findings of higher cost-effectiveness are particularly relevant for policy makers and health insurance providers. Our analysis indicates that the superior cost-effectiveness was primarily due to reduced costs for inpatient treatment and incapacity for work, indicating that the Rise-uP approach is effective in preventing the escalation of symptoms requiring inpatient treatment, thereby reducing costs associated with incapacity for work. Consequently, the Rise-uP approach holds the potential to significantly improve the treatment of NLBP in terms of both clinical and economic outcomes.

Statistical analysis showed that the Kaia back pain app was the central element for the positive clinical effects observed. Its efficacy in treating NLBP has been corroborated by multiple studies. Notably, the Rise-uP trial marks the first example, to our knowledge, of an app being integrated as a primary intervention within a real-world treatment framework.

The treatments received by the Rise-uP group extend beyond the use of the Kaia back pain app. The Rise-uP approach also includes an eCRF with clinical decision support and teleconsultations for individuals with increased risk of developing chronic pain. However, our prior findings indicate that the beneficial impact of teleconsultations on high-risk patients was directly linked to increased usage of the Kaia back pain app among those who had teleconsultations. While we acknowledge the eCRF’s utility in facilitating information sharing and supporting clinical decisions, we do not consider it played a major role in the outcomes. It appears to be a beneficial supplement that electronically navigates through guidelines without significantly influencing the results.

The Rise-uP approach offers an early, self-determined and location-independent treatment to the patient. This offers an extraordinary opportunity to initiate, to continue or to bridge treatment gaps in NLBP. Given the pivotal role of the Kaia back pain app within the Rise-uP framework, it suggests a move towards simplifying the telematic infrastructure, using an app as a primary interface.

Another issue of Rise-uP requires attention. We found that after 3 month the effect of the teleconsultation on pain reduction in patients with a high risk for chronic pain was fully mediated by a higher Kaia app adherence. We did not find a general relationship between app usage and symptom reduction over time in the 3-months data. The missing relationship was confirmed also in the present analysis for the whole observation period of 12 months. Hence, it becomes evident that the assertion “the more training, the better the outcome” is not universally applicable, as suggested by Priebe et al. Rather, our AI-based bottom-up analysis has unveiled four distinct patterns of usage characterized by varying levels of training intensity. Remarkably, each identified pattern appears to be effective in mitigating pain, suggesting that individuals may self-select the most suitable training “dosage” and discontinue their regimen upon achieving desired levels of pain alleviation. This insight into complex usage
dynamics is anticipated to play a pivotal role in customizing treatment protocols through medical applications for individual patients, as proposed by Priebe et al. To accurately tailor treatments to specific patient phenotypes, it is imperative to conduct extensive research to delineate different patient clusters based on their usage patterns. Thus, behavioral tracking as used in the Rise-uP study could be relevant for the future of medical apps as it could provide a deeper understanding of user behavior and adherence patterns. By identifying distinct clusters, this information could tailor interventions and content to specific user groups, improving overall engagement and effectiveness. Additionally, this analysis can inform product development, allowing for data-driven decisions to enhance user experience and outcomes. Understanding these patterns also aids in personalizing the app’s recommendations, potentially leading to better pain management and user satisfaction.

Despite demonstrating the superiority of the Rise-uP approach, integrating such innovative digital strategies into the healthcare system poses considerable challenges. Research has consistently shown that patients value digital enhancements as beneficial adjuncts to their treatment regimens and that patients are often more receptive to digital medicine than healthcare professionals. This receptiveness is further evidenced by the Rise-uP patients, who reported greater satisfaction with their digitally augmented treatment compared to those in the control group.

In order to facilitate the adoption of digital elements into clinical practice, it is essential for GPs and other healthcare providers to receive comprehensive training and support at the outset. During the Rise-uP trial, detailed guidance and account management for participating GPs was imperative for the smooth operation of the project. Integrating all digital functions of the Rise-uP approach into a single app could potentially ease technical requirements and enhance professional acceptance. This strategy aligns with the Technology Acceptance Model (TAM) which highlights perceived ease of use (PEOU) and perceived usefulness (PU) as critical factors for technology adoption. While the Rise-uP trial has convincingly demonstrated PU, PEOU remains a key consideration for the complex Rise-uP treatment strategy’s integration into healthcare systems. Thus, centralizing the treatment algorithm within an app interface could be advantageous.

Strength and Weaknesses
The main strength of our study is the real-world application of a comprehensive treatment strategy, albeit with potential limitations in internal validity due to a lack of control over multiple confounding variables. For example, treatment in the CG was not standardised, although we instructed GPs that the NVL had to be considered. Simultaneously, this enhances the external validity of the study, as it better reflects real-world conditions. Despite the efficacy of medical apps and digital interventions being established in RCTs (high internal validity), the “clean setting” of such studies often fails to capture real-life complexities (low external validity).

Another limitation refers to (1) significant dropout rates over time, with (2) a notable difference between the Rise-uP group (44%) and the CG (28%). While our dropout rates over 12 months with three follow-ups align with known standards, the disparity between groups warrants attention. However, this could suggest that integrating Rise-uP into patients’ everyday routines may demand much of their attention and daily life capacity, making continued follow-up participation challenging. Last, the two groups differed in some baseline characteristics (for example, age) which may affect results. However, when controlling for those differences by regression analysis, the superiority of the Rise-uP approach remained stable.

Conclusion
The Rise-uP trial has unequivocally shown that a digital treatment approach can significantly improve NLBP management compared to standard care treatment. Increased pain reduction, improvement in function and quality of life, increased patient satisfaction and cost savings are single and together important arguments for innovation and implementation. The Rise-uP trial underscores the potential of digital interventions to transform the treatment of NLBP, offering effective, accessible, and cost-effectiveness solutions.

Data Sharing Statement
Pseudonymized raw data referring to the present MS are available upon request from the corresponding author (TRT). Data are not publicly available for data protection reasons.
Acknowledgments
Rise-up is granted by the Innovationsfonds of the Gemeinsamer Bundesausschuss (G-BA) Germany (Grant No: 01NVF16014). We thank Margit Hermann, Florian Günther, Alexander Günter and Martin Steidler (AOK Bayern), Dr. Ursula Marschall, Dr. Claudia Wöhler, Regina Kotzbacher, Sandra Lehnerth (BARMER) and Markus Hinz and Sophie Schwab (DAK) for permanent support and collaboration in the Rise-up project. AOK is partner in the consortium, and Barmer und DAK are collaborators. We further thank Christine Schiessl, Simone Dohmen, Martina Firsching, Theodora Malamoussi and Richard Pecka (Algesiologikum Munich) for organizing and conducting the teleconsultations. We thank Pfizer Germany to generously allocate the basic version of the “Therapy Navigator” to Klinikum rechts der Isar, Technical University Munich, for unrestricted use in scientific projects in low back pain.

Disclosure
Janosch A. Priebe received a consultancy fee for data analysis of demographic data from Kaia subscribers via internet download and grants from Innovationsfond G-BA. Linda Kerkmeyer, Katharina K. Haas, Katharina Achtert, Leida F. Moreno Sanchez, Paul Stockert Maximilian Spannagl, Julia Wendlinger, Reinhard Thoma, Siegfried Ulrich Jedamzik, Jan Reichmann, Sebastian Franke, and Leonie Sundmacher. Volker E. Amelung declare that they have no conflict of interest. Thomas R. Toelle declares consultancies, travel grants and speaking fees for AOP Orphan, Almiral Hermal, Bionest Partners, Benkitt Renkiser, Grünenthal, Hexal, Indivior, Kaia Health, Lilly, Medscape Mundipharma, MSD, Novartis, Pfizer, Recordati Pharma, Sanofi-Aventis, and TAD Pharma, and grants from Tölle. The authors report no other conflicts of interest in this work.

References


29. Kohlmann T, Raspe H. [Hannover functional questionnaire in ambulatory diagnosis of functional disability caused by backache] (Der funktionsfragebo...


