PROSPECTIVE CLINICAL STUDY REPORT 2020-2021

ACCORDING TO THE STANDARDS
ICH E3 HARMONIZED GUIDELINE AND DIN EN ISO 14155

Prospective Randomized Group Controlled Trial (RCT)
of Clinical Efficacy and Safety of the
CE marked Vivira App "In-Home Therapeutic Training Program"

Type of investigational product
App for smart phones/tablets
CE marked Medical Device

Name of investigational product
Vivira In-Home Therapeutic Training Program App

Indication studied
Back Pain

Study type
Confirmative prospective randomized group controlled two arms trial (RCT)
of clinical efficacy and safety with 213 patients and active safety surveillance

Name of the sponsor
Vivira Health Lab GmbH

Protocol identification
Vivira-RCT-CIPConcept-APP-2-20-F

Development phase of study
Postmarket

Study initiation date (first patient enrolled)
Nov 16, 2020

Study completion date (first patient last visit)
March 31, 2021

Date of early study termination
n/a

Investigators
Professor K. Weise, M.D., Dr. med.
Certified Medical Specialist in Orthopedic Surgery & Traumatology
Professor of Orthopedic Surgery & Traumatology, Tübingen

Dr. Dr. H. Weise, M.D., M.DD.
Certified Medical Specialist in Maxillofacial Surgery, Tübingen

Professor H.P. Zenner, M.D., Dr. med.
Certified Medical Specialist in ORL, Head&Neck Surgery
Professor of Head&Neck Surgery, Tübingen

Name of company/sponsor signatory
Dr. Philip Heimann, CEO
Vivira Health Lab GmbH
Kurfürstendamm 54/55
10707 Berlin, Germany
Tel.: +49-151-40732078
Email: philip@vivira.com

Persons responsible for the study report
Professor K. Weise, M.D., Dr. med.
Dr. Dr. H. Weise, M.D., M.DD.
Professor H.P. Zenner, M.D., Dr. med.

Clinical Research Organization
Prof. Dr. H.P. Zenner GmbH
Bismarckstr. 66
72072 Tübingen, Germany
Phone: +49 7071 49-645
office@hpzenner.de

GCP
The study was performed in compliance with Good Clinical Practices (GCP), including the archiving of essential documents.

Date of the report
Oct. 8th, 2021

Earlier clinical reports

Febr. 13, 2020: Clinical Evaluation Report by Dr. Fritz Ley

June 2, 2020: Medical Report of Vivira User Data by Dr. Fritz Ley

July 2, 2020: Vivira-RStudy-ICHReport-APP-1-20-F

1. SYNOPSIS

Pain of the musculoskeletal system, especially as non-specific pain in the back is common[1]. Among them is low back pain (LBP), the leading cause of disability years. There is a 1-month prevalence of about 30% of the world population[1]. Non-specific LBP is more frequent than back pain with a specific cause that can be treated with a specific treatment[2]. The clinical standard of care for non-specific pain of the back takes into account physical activity and activation [3]. The German National Guideline for Non-Specific Low Back Pain suggests multidisciplinary programs[4], which explicitly may include exercise therapy. However, these multidisciplinary programs are limited to specialized centers, which limits their wide dissemination.

In the recent past, health apps have been developed to address these and other allocation problems of a sufficient spread of therapy methods. In 2017, the number of health apps released from iTunes and Google Play exceeded 300 000, with nearly 25% dealing with disease self-management. One-third of adults in the US with smartphones or tablets use health apps to achieve health behavior goals or help with medical decision making. International standards exist with regard to software engineering, privacy, security and usability of mobile apps in general (e.g. International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) standards). For health apps an important standard was set by the German DiGAV 2020 for provision within the German social security system GKV.

The present study report presents the results of a prospective group controlled confirmative study on the effectiveness and safety of the Vivira In-Home Therapeutic Training Program App (Vivira app). In both groups ITT-patients suffering from back pain including unspecific back pain were treated. The control group received TAU (therapy as usual), which was physiotherapy according to the German GKV rules. Physiotherapy prescription included physical therapy provided by physiotherapists and additional home exercises as recommended by the same physiotherapists. The experimental group received the Vivira app. The app comprises a comprehensive multidisciplinary feedback controlled treatment concept, including patient education and video-supported In-Home Therapeutic Training Program. The content of the app is in accordance with current German guidelines including the National Care Guideline "Non-Specific Back Pain" for the management of LBP [5]. Using a verbal numerical rating scale (VNRS) as primary efficacy variable the study compared pretherapeutic pain levels with the relevant data at the end of a period of up to 3 months after intention to treat. Efficacy was controlled by control a group receiving conventional physiotherapy.

[1] Ulrich, B., ; J. Weitz, Jürgen; V. Penter: Apps und Mobile Health: Viele Potenziale noch nicht ausgeschöpft. Dtsch Arztebl 2018; 115(3): A-62 / B-57 / C-57.
[2] Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017 Feb 18;389(10070):736-747.
[3] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen. 2017
Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 04;166(7):514-530.
[4] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen
[5] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen

Further, efficacy variables included functional scores on quality of life, mobility, stability and coordination as well as transient medication consumption.

N=213 ITT-patients were analyzed of which N=108 were randomly allocated to the experimental group and N=105 to the control group. The majority of cases suffered from lumbar spine pain in both groups followed by thoracic spine pain and its variants in the ITT population. Using a German language validated verbal numerical rating scale (VNRS) as primary efficacy variable for pain both groups of the ITT population produced a mean pain decrease from baseline after 2, 6 and 12 weeks. In the Vivira group pain decreased by -38.01 % (week 2), -45.73 % (week 6) and -53.11 % (week 12). In the physiotherapy group this was no more than - 2.45 %, 7.14 %, 14.62 %). Pain score changes in both groups the Vivira group and the physiotherapy group were statistically significant. T-tests at T1, T2 and T3 all revealed $p < 0.001$ for the intragroup changes in the Vivira group. For the physiotherapy intragroup changes produced significant p-values at T2 and T3 of 0.015 and 0.0003.

Confirmative group comparison demonstrated that in the Vivira app group the average pain decrease was significantly more pronounced than in the physiotherapy group. At all measurement times group differences were highly significant (t-tests: $p<0.001$ (2 weeks)/ $p<0.001$ (6 weeks)/ $p<0.001$ 812 weeks)) indicating superiority of the Vivira app. Cohen's values reveal a statistically significant large effect size difference between the Vivira group and the physiotherapy group confirming superiority of Vivira over TAU.

There were significantly more non-responders (e.g. after 12 weeks: n=42 vs. n=9) and exacerbations in the physiotherapy group than in the Vivira group. While among the Vivira patients only N=3 patients (2.8%) reported an exacerbation of pain (after 6 weeks), in the TAU group the number of patients with pain exacerbations was N=24 (22.9%, after 2 weeks), increasing to N=26 (24.8%) after 6 weeks and then decreasing until 12 weeks, but still amounting to N=18 (17.1%). Furthermore, shift analysis demonstrates that after 12 weeks N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower (N=63, 60.0%).

Using the SF-36 for the determination of QoL this improved in both groups. A statistically significant overall QoL improvement, however, was observed only in the Vivira group (T2: p=0.0252, T3: p=0.0180) suggesting superiority of Vivira over the control. When the mental score (MCS) was determined the significant treatment difference (T2: p=0.0186) between both groups at T2 supported the superiority of Vivira. When investigating the physical score (PCS) significance of improvement was given only in the Vivira group (T2: p=0.0170; T3: p=0.0184) also suggesting superiority of Vivira over the control.

Movement analyses revealed that mobility, stability (strength) and the total functional score improved significantly in both groups (T3: $p<0.0001$). Treatment differences between both groups (Vivira vs. physiotherapy) were not statistically significant. Coordination Score in both study groups showed only a statistically insignificant increase from baseline to weeks 6 and 12. Group comparison movement analysis revealed similar results for both study groups i.e. data shows a statistically significant efficacy but no superiority of Vivira.

Using the consented Medication Qualification Scale (MQS) there was a statistically significant non-linear decrease in analgesic consumption in the Vivira group by – 1.79/-67% (SD=2.24,

p=<.0001) between weeks 1 and 12. In the control group analgesic consumption is also reduced, which is -0.57/-27% (SD: 2.07, p=0.1268) but statistically not significant. The group difference is significant (p=0.0210). Thus, Vivira is not only effective in reducing analgesic consumption, but superior to control physiotherapy.

Safety data resulting from active surveillance clearly indicate with high evidence strength that the inherent rate and severity of Vivira complications is very low. Vivira produced burden each of which is fully acceptable in relation to the benefits. Furthermore, burden produced by Vivira is lower than that by physiotherapy. Moreover, complication rate and AR quality equal those of conventional physiotherapy. Finally, complication rate and AR quality are significantly lower than those of surgery. Specifically, in contrast to surgery no SARs have been observed. Thus, based on the findings it can be inferred that the probability of a patient experiencing a substantial benefit when using the Vivira App significantly outweighs the probability of suffering harm due to a risk of the device.

Conclusion. As a conclusion pain score analyses showed

- A statistically significant reduction of pain over time through the use of the App.
- A statistically significant superiority in pain reduction of the app use in comparison to TAU.
- According to Cohen Vivira achieved a large effect size superior to that of TAU
- The effect size represents a high clinical effect strength superior to that of TAU

Furthermore, statistically significant QoL improvements were observed suggesting superiority of Vivira over the control. Group comparison movement analyses revealed similar results for both study groups i.e. it shows a statistically significant efficacy but no superiority of Vivira. Moreover, Vivira was effective in reducing transient analgesic consumption being superior to the control physiotherapy. Safety findings reveal that the probability of a patient experiencing a substantial benefit when using the Vivira App significantly outweighs the probability of suffering harm due to a risk of the device.

2. TABLE OF CONTENTS

3. LIST OF ABBREVIATIONS

CE-marked *Certification Mark (European Conformity)*
CIP *Clinical Investigation Plan*
DFT *Dorsiflexion Lunge Test*
EEC *European Economic Community*
EN *European Norm*
FMS *functional movement screen*
GCP *Good Clinical Practices*, *Good Clinical Practices*
GDNÄ Gesellschaft Deutscher Naturforscher und Ärzte
GHTF *Global Harmonization Task Force*, *since 2011:IMDRF* XE "IMDRF" \t "*International Medical Devices Regulators Forum*" *Ma*
i.e. *id est - that is*
IEC *International Electrotechnical Commission*
IMDRF *International Medical Devices Regulators Forum*
ISO *International Organization for Standardization*
ITT *Intention-to-Treat*
LBP *low back pain*, *low back pain*
LOCF *last-observation-carried-forward*
mITT *Modified ITT population*
MQS *Medication Qualification Scale*
ORL Oto-Rhino-Laryngologie
PMCF *Post-Market Clinical Follow-up*
PP *Per protocol population*
RCT *Randomized controlled trial*
SAF *Safety population*
SARs *Serious adverse reactions*
SFMA *selective functional movement assessment*
SG *Study Group*
TAU *Treatment as usual*
TDD *Test driven development*
USA United States of America
VAS *Visual Analogue Scale*
VNRS *verbal numerical rating scale*,

4. DEFINITION OF TERMS

Adverse event: Side effects that are harmful; any untoward medical occurrence, unintended disease or injury, or any untoward clinical signs (including an abnormal laboratory finding) in subjects, users or other persons whether or not related to the investigational medical device. [EN ISO 14155:2011]

Adverse reaction: Events for which a *causality* link to the tested intervention is well established and strong enough (sensitive and specific)

Android: an open-source operating system used for smartphones and tablet computers.

Android device: an electronic device that runs the operating system Android.

Application: a program or piece of software designed to fulfil a particular purpose.

Backend: the data access layer of the app.

Bias: bias is a systematic deviation of an outcome measure from its true value, leading to either an overestimation or underestimation of a treatment's effect. It can originate from, for example, the way patients are allocated to treatment, the way treatment outcomes are measured and interpreted, and the way data are recorded and reported. [Adapted from GHTF SG5/N2R8:2007]

Broadband internet connection: an internet connection with a high data transfer rate.

Clinical data: the safety and/or performance information that is generated from the clinical use of a device. Clinical data are sourced from:
- clinical investigation(s) of the device concerned; or
- clinical investigation(s) or other studies reported in the scientific literature, of a similar device for which equivalence to the device in question can be demonstrated; or
- published and/or unpublished reports on other clinical experience of either the device in question or a similar device for which equivalence to the device in question can be demonstrated.

Clinical evaluation: a methodologically sound ongoing procedure to collect, appraise and analyze clinical data pertaining to a medical device and to evaluate whether there is sufficient clinical evidence to confirm compliance with relevant essential requirements for safety and performance when using the device according to the manufacturer's Instructions for Use.

Note: In exceptional cases where an instruction for use is not required, the collection, analysis and assessment are conducted taking into account generally recognised modalities of use.

Clinical evidence: the clinical data and the clinical evaluation report pertaining to a medical device. [GHTF SG5/N2R8:2007]

Clinical investigation: systematic investigation in one or more human subjects, undertaken to assess the safety or performance of a medical device.
Note: 'clinical trial' or ' clinical study' are synonymous with ' clinical investigation'.
[EN ISO 14155:2011]

Clinical investigation plan: document that states the rationale, objectives, design and proposed analysis, methodology, monitoring, conduct and record-keeping of the clinical investigation. [EN ISO 14155:2011]

Clinical performance: behavior of a medical device or response of the subject(s) to that medical device in relation to its intended use, when correctly applied to appropriate subject(s). [EN ISO 14155:2011]

Clinical safety: freedom from unacceptable clinical risks, when using the device according to the manufacturer's Instructions for Use. [MEDDEV 2.7/2 revision 2] Not simply absence of evidence of harm.

Clinical use: use of a medical device in or on living human subjects. Includes use of a medical device that does not have direct patient contact.

Confluence: is a collaboration software program developed and published by the software company Atlassian.

Coordination: the ability to use different parts of the body together smoothly and efficiently. It is one of the three functional dimensions in the Vivira movement test.

Coordination Score: an efficacy variable used in this document. A score between 0 and 100 is calculated based on the user's self-assessment in the Movement Test specifically relating to a user's abilities in the dimension of coordination.

Device registry: an organized system that uses observational study methods to collect defined clinical data under normal conditions of use relating to one or more devices to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure and that serves predetermined scientific, clinical or policy purpose(s). The term "device registry" is different from the concept of device registration and listing. [MEDDEV 2.12/2 rev2]

Educational content: short content articles provided in the Vivira app giving users information in several categories, e.g. pain, relaxation, etc.

Elixir: a functional, concurrent, general-purpose programming language that runs on the Erlang virtual machine used for the Vivira app.

Equivalent device: a device for which equivalence to the device in question can be demonstrated.

Exercise day: a day where a user has completed at least one exercise. An exercise is completed if the user confirmed that they have performed the exercise and given feedback regarding pain and their completion of reps or holding times.

Exercise difficulty: the integration of two dimensions regarding the exercises assigned to a user. The first dimension determining exercise difficulty is the intensity at which an exercise is to be performed, e.g. the number of repetitions to perform or the length of the holding time. The second dimension of exercise difficulty is defined by the exercise itself. Within each exercise group exercises can be ordered from least to most challenging. A user will progress in exercise difficulty first along all exercise intensities of the easiest exercise of a group before progressing to the next and more challenging exercise.

Exercise history: an overview in the app that depicts on which dates users exercised, which exercises they performed and which feedback they gave.

Exercise program: a specific set and ordering of exercises assigned to each user based on the input they provide at the start of and throughout their use of the Vivira app.

Exercise video: a video displayed in the Vivira app as part of the training program. In the video the correct exercise execution is demonstrated, and additional information is displayed.

Exercise progression: a term defining the Vivira exercise system, which bases exercise difficulty on user feedback. For each user the daily exercise program is compiled based on the exercises that were performed in the past and the feedback the user gave for each exercise. Only if users are able to perform one exercise at a given exercise intensity do they move on (progress) to a higher intensity for the same exercise or a more challenging exercise.

Feasibility study: a clinical investigation that is commonly used to capture preliminary information on a medical device (at an early stage of product design) to adequately plan further steps of device development, including needs for design modifications or parameters for a pivotal study. [MEDDEV 2.7/2 revision 2]

Frontend: the presentation layer of the app.

Google Play Store: a digital distribution platform, developed and maintained by Google Inc., for mobile apps on the Android operating system. The store allows users to browse and download apps developed with the Google software development kit.

GraphQL: an open-source data query and manipulation language for APIs, and a runtime for fulfilling queries with existing data.

Harmonised standards: standards whose references have been published in the Official Journal of the European Communities.

Harms: The totality of possible adverse consequences of an intervention or therapy; direct opposite of benefits, against which they must be compared.

Hazard: potential source of harm. [EN ISO 14971:2012]

Hazard due to substances and technologies: for the purpose of this document, a hazard that is seen with products that share specific characteristics. This includes products that contain the same materials and substances, material combinations, use the same technologies, produce similar abrasion, are used with the same type of surgical approach,

share the same manufacturing procedures or impurities, or share other characteristics.

Household Score: an efficacy variable used in this document. A value between 0 and 11 self-reported by the user as part of the weekly Wellbeing Journal. The score describes to what extent the user felt that they were limited in their home activities.

In-app purchase: refers to the buying of goods and services from inside an application on a mobile device, such as a smartphone or tablet.

Incident: any malfunction or deterioration in the characteristics and/or performance of a device, as well as any inadequacy in the labelling or the instructions for use which, directly or indirectly, might lead to or might have led to the death of a patient, or user or of other persons or to a serious deterioration in their state of health. [MEDDEV 2.12/rev 8]

Information materials supplied by the manufacturer: for the purpose of this document, this refers to the labelling, instructions for use and the manufacturer's promotional materials for the device under evaluation.

Intended purpose: the use for which the device is intended according to the data supplied by the manufacturer on the labelling, in the instructions and/or in promotional materials.

Investigator: individual member of the investigation site team designated and supervised by the principal investigator at an investigation site to perform critical clinical-investigation-related procedures or to make important clinical investigation-related decisions. [EN ISO 14155:2011]

iOS: an operating system used for mobile devices manufactured by Apple Inc.

iOS App Store: a digital distribution platform, developed and maintained by Apple Inc., for mobile apps on its iOS operating system. The store allows users to browse and download apps developed with Apple's iOS software development kit.

iOS device: an electronic device that runs the operating system iOS.

Java: a general-purpose programming language used for the Android operating system.

Jira: a proprietary issue tracking product developed by the software company Atlassian that allows bug tracking and agile project management, among other things.

Kotlin: a cross-platform, statically typed, general-purpose programming language with type inference used for the Android operating system.

Leisure Score: an efficacy variable used in this document. A value between 0 and 11 self-reported by the user as part of the weekly Wellbeing Journal. The score describes to what extent the user felt that they were limited in their leisure activities.

Limitation in quality of life: an efficacy variable used in this document. The limitation in quality of life is represented by a number the user selects on a numerical verbal rating scale. 0 is the lowest and 10 the highest limitation of quality of life.

Mobility Score: an efficacy variable used in this document. A score between 0 and 100 is calculated based on the user's self-assessment in the Movement Test specifically relating to a user's abilities in the dimension of mobility.

Movement Test: a self-assessment in the Vivira app, which is performed every four weeks. Based on the user's ability to perform a set of defined movements a score is compiled for each dimension (Stability, Mobility, Coordination) and a total score is calculated.

MVP: a derivation of the model–view–controller (MVC) architectural pattern which mostly used for building user interfaces for the Android operating system.

Onboarding: the configuration phase of the app, where users answer questions that determine the assignment of a therapy program. Onboarding is completed after users have answered all questions and created a Vivira account.

Pain score: the number a user selects on a numerical rating-scale that represents the intensity of their perceived pain.

Paywall: an arrangement whereby access is restricted to users who have paid to use the app.

Phoenix: a web development framework written in the functional programming language Elixir.

PMCF plan: the documented, proactive, organized methods and procedures set up by the manufacturer to collect clinical data based on the use of a CE-marked device corresponding to a particular design dossier or on the use of a group of medical devices belonging to the same subcategory or generic device group as defined in Directive 93/42/EEC. The objective is to confirm clinical performance and safety throughout the expected lifetime of the medical device, the acceptability of identified risks and to detect emerging risks on the basis of factual evidence. [MEDDEV 2.12/2 rev.2]

PMCF study: a study carried out following the CE marking of a device and intended to answer specific questions relating to clinical safety or performance (i.e. residual risks) of a device when used in accordance with its approved labelling. [MEDDEV 2.12/2 rev.2]

Premium user: a person who uses or operates Vivira and is charged based on the usage duration they selected.

Progression system: the system defining how user feedback after each exercise determines the next assigned set of exercises. The progression system ensures that users are assigned exercises that are in line with their individual abilities.

Progress visualization: a section of the app that provides graphic representations of the self-reported input that users have made in the Wellbeing Journal and the Movement Test.

Risk: combination of the probability of occurrence of harm and the severity of that harm. [EN ISO 14971:2012]

Risk management: systematic application of management policies, procedures and practices to the tasks of analyzing, evaluating, controlling and monitoring risk. [EN ISO 14971:2012]

RxJava: a programming library used for the Android operating system.

RxKotlin: a programming library used for the Android operating system.

Safety: Substantive evidence of an absence of harm. Not simply absence of evidence of harm.

Serious adverse event:
Serious adverse events6 during clinical investigations: in their most severe forms, threaten life or function. If suspected to be product-related (adverse reactions), might be significant enough to lead to important changes in the way the medicinal product is developed (e.g., change in dose, population, needed monitoring consent forms).

Adverse event that
a) led to death,
b) led to serious deterioration in the health of the subject, that either resulted in
1) a life-threatening illness or injury, or
2) a permanent impairment of a body structure or a body function, or
3) in-patient or prolonged hospitalization, or
4) medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function,
c) led to foetal distress, foetal death or a congenital abnormality or birth defect.
NOTE: Planned hospitalization for a pre-existing condition, or a procedure required by the CIP [Clinical Investigation Plan], without serious deterioration in health, is not considered a serious adverse event. [EN ISO 14155:2011]

[6] As defined by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, document E2A (available at www.ich.org)

Side effects: Unintended effects, not necessarily of harmful nature.

Smartphone: a mobile phone that performs many of the functions of a computer, typically having a touchscreen interface, internet access, and an operating system capable of running downloaded apps.

Stability: Stability is the resistance of a muscle or group of muscles to control joint position and balance. It is one of the three functional dimensions in the Vivira movement test.

Stability Score: an efficacy variable used in this document. A score between 0 and 100 is calculated based on the user's self-assessment in the Movement Test specifically relating to a user's abilities in the dimension of stability.

Sufficient clinical evidence: an amount and quality of clinical evidence to guarantee the scientific validity of the conclusions.

Surveillance of harms, passive: participants spontaneously report

Surveillance of harms, active: participants are asked (in structured questionnaires or interviews or predefined laboratory or other diagnostic tests at prespecified time intervals)

Swift: a general-purpose, multi-paradigm, compiled programming language developed by Apple Inc. for iOS.

Tablet PC: a mobile device, typically with a mobile operating system and touchscreen display processing circuitry, and a rechargeable battery in a single, thin and flat package.

Total Functional Score: an efficacy variable used in this document. The total score (between 0 and 100) is calculated based on the user's self-assessment in the Movement Test.

Toxicity: Describes product-related harms. The term may be most appropriate for laboratory-determined measurements, although it is also used in relation to clinical events. Abnormal laboratory values may be described as laboratory-determined toxicity. The disadvantage of the term "toxicity" is that it implies causality. If authors cannot prove causality, the terms "abnormal laboratory measurements" or "laboratory abnormalities" are more appropriate to use.

Test-driven development (TDD): a process of modifying the code in order to pass a test designed previously. In Software Engineering, It is sometimes known as "Test First Development."

Test user: a person who uses or operates Vivira free of charge for a limited time.

User: a person who uses or operates an application.

User interface: the means by which the user and the Vivira app interact.

Visualization: Human perception and cognition of an image.

Vivira account: Users of the Vivira app create an account, which enables them to access the app using their email address and a password. Each Vivira user and Vivira account is assigned a unique user identification (User ID).

Wellbeing Journal: a set of questions users answer on a weekly basis. Questions cover pain intensity in the last week, limitations in quality of life, limitations at home, limitations at work and limitations in leisure activities.

Work Score: an efficacy variable used in this document. A value between 0 and 11 self-reported by the user as part of the weekly Wellbeing Journal. The score describes to what extent the user felt that they were limited in their work.

5. ETHICS

Independent Ethics Committee

The study concept was advised by the *Ethik-Kommission der Landesärztekammer Baden-Württemberg* under the appraisal number F-2020-122.

Clinical Trial Registry

The study is registered under DRKS00022781 in the German Clinical Trials Registry DRKS.

Ethical Conduct of the Study

The Ethics Committee appraisal was obtained. After oral clarification by a doctor patients signed the written declaration of informed consent. Access to personal data was restricted to clinical investigators and patients. The data access of the manufacturer is restricted to anonymous result data. The study was conducted in accordance with the ethical principles that have their origins in the Declaration of Helsinki and in compliance with Good Clinical Practices (GCP), including the archiving of essential documents. Data protection follows Bundesdatenschutzgesetz and EU DSchGV.

Patient Information and Consent

After oral clarification by a doctor patients signed the written declaration of informed consent to participate in the study and agreeing with their data usage. In addition, they received a written patient information including study description and purpose, possible AEs, name and address of the insurer and explicit data protection information.

Continuing communication with the EC

The following information was provided to the EC:
a) serious adverse events;
If applicable, the following additional information is or was intended to be provided to the EC:
b) requests for deviations, and reports of deviations, if the deviation affects subject's rights, safety and wellbeing, or the scientific integrity of the clinical investigation.

Under emergency circumstances, deviations from the CIP to protect the rights, safety and well-being of human subjects may proceed without prior approval of the sponsor and the EC. Such deviations shall be documented and reported to the EC as soon as possible.
c) progress reports, including safety summary and deviations;
d) amendments to any documents already approved by the EC;
e) if applicable, notifications of suspensions or premature termination;
f) if applicable, justification and request for resuming the clinical investigation after a suspension;
g) clinical investigation report or its summary.

Compensation and additional health care

Compensating subjects for costs resulting from participation in the clinical investigation (e.g. transportation) was appropriate and allowed by national regulations. It includes transportation costs and incentives up to 90 Euros. This amount did not unduly encourage the subjects to participate. Arrangements were made by the study director for additional health care for subjects who suffer from an adverse event as a result of participating in the clinical investigation.

6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE

Sponsor
Vivira Health Lab GmbH, Berlin (Dr. Philip Heimann, CEO)

CRO
Clinical Research Organization
Prof. Dr. H.P. Zenner Clinical GmbH, Tübingen

Investigators
Prof. Dr. med. Kuno Weise, M.D.
Dr. Dr. Hannes Weise. M.D., M. DD.
Prof. Dr. med. Hans-Peter Zenner, M.D.

Steering, data collection and monitoring
Clinical investigators and monitors of the CRO Prof. Dr. H.P. Zenner GmbH

Statistics and statistical report
Dipl.-Stat. Michael Bulitta

Study report
Prof. Dr. med. Kuno Weise, M.D.
Dr. Dr. Hannes Weise. M.D., M. DD.
Prof. Dr. med. Hans-Peter Zenner, M.D.

7. INTRODUCTION

Pain of the musculoskeletal system, especially as non-specific pain in the back are common. Among them is low back pain (LBP), the leading cause of disability years[7]. There is a 1-month prevalence of about 30% of the world population. Non-specific LBP is more frequent than back pain with a specific cause that can be treated with a specific treatment[8] . The clinical standard of care for non-specific pain of the back takes into account physical activity and activation[9]. The German National Guideline for Non-Specific Low Back Pain suggests multidisciplinary programs[10], which explicitly may include exercise therapy. However, these multidisciplinary programs are limited to specialized centers, which limits their wide dissemination.

In the recent past, health apps have been developed to address these and other allocation problems of a sufficient spread of therapy methods. In 2017, the number of health apps released from iTunes and Google Play exceeded 300 000, with nearly 25% dealing with disease self-management. One-third of adults in the US with smartphones or tablets use health apps to achieve health behavior goals or help with medical decision making. International standards exist with regard to software engineering, privacy, security and usability of mobile apps in general (e.g. International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) standards). For health apps an important standard was set by the German DiGAV 2020 for provision within the German social security system GKV.

The present study report presents the results of a prospective group controlled confirmative study on the effectiveness and safety of the Vivira In-Home Therapeutic Training Program App (Vivira app). The app was used for the treatment of ITT-patients suffering from pain of the back. The app comprises a comprehensive multidisciplinary feedback controlled treatment concept, including patient education and video-supported physiotherapy. The content of the app is in accordance with current German guidelines including the National Care Guideline "Non-Specific Back Pain" for the management of LBP11. Using a verbal numerical rating scale (VNRS) as primary efficacy variable the study compared pretherapeutic pain levels with the relevant data at the end of a period of up to 3 months after intention to treat. Efficacy was controlled by a control group receiving conventional physiotherapy. Further, efficacy variables included functional scores on quality of life, mobility, stability and coordination as well as medication reduction.

[7] Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017 Feb 18;389(10070):736-747.
[8] Ibid
[9] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen. 2017
Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 04;166(7):514-530.
[10] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen
[11] AWMF 2017. Nationale Versorgungsleitlinie Kreuzschmerzen

8. STUDY OBJECTIVES

Study objective is the prospective group controlled confirmative determination of postmarket clinical efficacy and safety of the Vivira App “In-Home Therapeutic Training Program”. Safety was addressed by an active surveillance of AEs.

9. STUDY PRODUCT

Product design

According to the software safety classification as per EN 62304, Vivira is a classified as a Class B software.

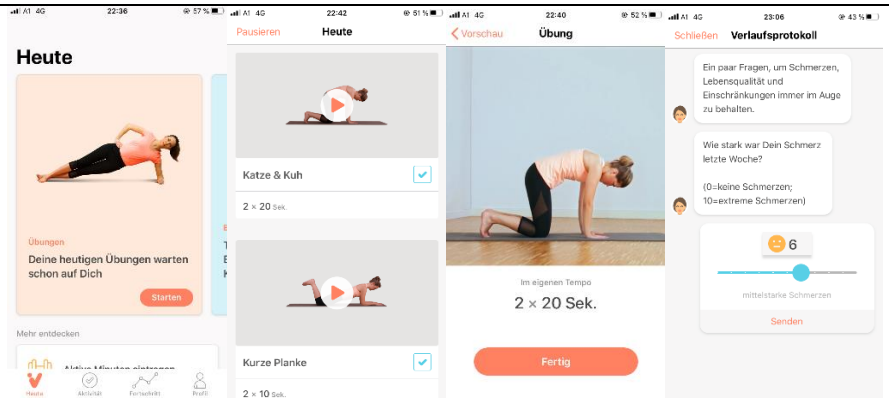
Product variants

None

Product properties

Vivira is a digital application that is used on smartphones and tablets. Users tap on the device’s screen to control the application. They watch the video exercise demonstrations and read instructions and information. They input information regarding their wellbeing and configuration needs. The Vivira App is designed to allow handling according to today's established design standards. A more detailed description may be found in the Description of Product.

Use requires broadband internet connection and a smartphone or tablet PC.

Category	Product
Product Name	Vivira
Screen shots of user interface	
Project Name	n/a
Product Models	app for iOS devices, app for Android devices
Device group, subgroup	Software app
CE Status	CE-marked 2017
Market	The product is on the market in Germany.
Classification	Class I according to MDD appendix VII
Accessories	None
Manufacturer(s)	Vivira Health Lab GmbH, Kurfürstendamm 54/55, 10707 Berlin
User group	Patients, lay persons
Medical indication	• Gonarthrosis • Nonspecific knee pain • Coxarthrosis • Nonspecific hip pain • Nonspecific back pain • Spine osteochondrosis
Age group	18 yrs. and older
Gender	m/f
Type and severity of the medical condition	Acute and chronic moderate to severe
Range of time	Transient application Number of repeat exposures: daily for 3 months up to unlimited period of time
Intended use	Vivira helps users to reduce physiotherapeutically treatable pain of the musculoskeletal system

Nature and duration of contact to the human body	Vivira has no contact with the human body
Material in contact to the human body	None
Organs in contact with material	None
Tissue in contact with material	None
Fluids in contact with material	None

Table 1: Vivira App Product

Use

Patients go through an

- **introduction** to configure their training program. The patient confirms that from a medical point of view, he is able to train on his own. If he does not confirm this, he is requested to consult a doctor before using Vivira. Afterwards they will receive
- **daily specific exercises** tailored to their complaint region, their indication and their progression ability. Exercise instructions are accompanied by tutorial videos that demonstrate the correct execution of the respective exercise. The exercises of one day form a training session, and for each exercise patients give
- **feedback** on pain and their ability to perform the exercise. This feedback determines the
- **progression of the exercises** in further course. Throughout the course of the treatment patients record their daily physical activity and weekly their progress in pain, quality of life and limitations. They perform a monthly movement test for stability, mobility and neuromuscular control of their locomotor system. The training and activity level, the training progress, as well as the course curves are visualized in
- **progression curves.**
- **Key data** about the program and the progress achieved can be exported at any time and shared with a doctor/therapist. Patients can manage their Vivira account, their training programs and their data directly in the app at any time.

Interventions

The Vivira training system was developed by the Vivira Medical Board in cooperation with the Vivira physiotherapy team. Chairman of the Vivira Medical Board is Dr. med. Markus Klingenberg, who is an Ärztekammer certified specialist for Orthopedics and Orthopedic Surgery with Beta Klinik, Bonn, Germany. The following interventions are part of the program:

Independent exercising. Independent of place and time, patients are prompted to complete four exercises per day. Their correct execution is demonstrated through video and text, respectively.

Personalized progression. Patients start after stating their area of complaint, their indication (provided available) and their limitations, a correspondingly specific training program. The personalization of the exercise contents is patient feedback controlled from the first exercise performed. Via the feedback after each exercise, exercises are adapted to the patient's abilities on a daily basis. The progression algorithm allows patients to constantly practice in their range of abilities, so that they are neither under- nor overchallenged. At the beginning of use, the algorithm approaches the capability range "from below", so that no excessive demands are made. This minimizes the risk of AEs.

Functional training. The effectiveness of the exercises is provided by the functional exercise approach. Musculoskeletal functionality as a whole is trained according to the functional approach and may contribute to pain reduction. Focus is on mobility (agility), stability (strength) and neuromuscular control (coordination).

Continuous cycle (exercises, progress recording, tests). Patients experience a continuous cycle of exercises, progress recording and movement tests. The cycle continues over the entire application time, over levels of difficulty, creates awareness for one's own change process, motivates and thus contributes to the adherence to therapy.

Education. Within the framework of education, the patient learns an adequate disease model. Thus, he is thought to be motivated to perform the upcoming exercises correctly and with a high compliance and adherence.

10. INVESTIGATIONAL PLAN

Overall Study Design and Plan-Description

The study is a confirmative prospective randomized group controlled two arms trial (RCT) of clinical efficacy and safety with active AE surveillance. The experimental group used Vivira. The control treatment (comparator) consisted of the GKV-standard physiotherapy (TAU, treatment as usual).

Discussion of Study Design, including the Choice of Control

In the study, patients received the app free of charge and were randomly assigned to one of the two treatment groups. In the first group, patients used the app for 12 weeks from the beginning. In the second group patients received physical therapy/physiotherapy free of charge and were offered the app after 12 weeks. This control treatment (comparator) consisted of the GKV-standard physiotherapy as a typical setting with a physiotherapist. According to the GKV regulation (Heilmittelkatalog), a regular prescription of physiotherapy consisted usually of six individual sessions. If the treating physician did not prescribe the physiotherapy this was prescribed by an investigating doctor. No influence was exerted on timing, possible waiting time or additional physiotherapy sessions, so that these also corresponded to the standard GKV procedure. Thus, the effectiveness and safety of app and physiotherapy could be compared scientifically. Moreover, regardless of which of the two groups the patients were in: as an ethical result an individual benefit could be expected in both groups.

At visit T0, following the informed consent, all participants received a clinical investigation by the investigating doctor. Furthermore, doctors performed a structured interview including case history and an active assessment of inclusion and of all exclusion criteria. At the time points T1-T3 doctors reperformed telemetrically the structured interview including interim case history and an active assessment of AEs. At the time points T0-T3 all study patients made their statements on rating score and questionnaires and performed and documented the movement tests.

Study objective. Study objective was the prospective controlled determination of postmarket clinical efficacy and safety of the Vivira App “In-Home Therapeutic Training Program”.

Hypotheses. Major hypothesis was that the RCT can distinguish Vivira app efficacy from standard GKV physiotherapy efficacy.

Method of assigning patients to treatment groups. To minimize stratification bias a stratified block randomization was chosen.

Homogeneity. From the scientific and ethical point of view, the RCT was characterized by homogeneous groups to meet scientific criteria for significance and power and for limiting the number of patients to meet ethical criteria of the Declaration of Helsinki. For these scientific and ethical reasons study indication was limited to back pain allowing to generate homogeneous groups. A preceding retrospective study (2018-2019; Ref. Vivira-RStudy-ICHReport-APP-1-20-F) revealed that back pain is the key symptom for pain of the musculoskeletal system when treated using the Vivira app. For this reason, the group of patients with back pain was taken as homogeneous group to be allocated to both the experimental group and the control group.

For the treatment, however, of the remaining pain symptoms of the musculoskeletal system including hip and knee pain evidence of Vivira efficacy has also to be provided. A transfer of evidence between the primary area of complaint (unspecific back pain), as assessed in this study, and the secondary areas of complaint is to be substantiated through the cohort study mentioned above (Ref. Vivira-RStudy-ICHReport-APP-1-20-F) as well as an additional cohort study (Ref. Vivira-RStudyReport-2020-21-1-21-F) that was performed simultaneously to the study presented here. In addition, a metaanalysis is provided (Ref. Vivira-Meta-1-21-F).

Selection of timing for each patient. Experimental patients were free to decide when to exercise. Thus, experimental patients choose their individually preferred application time. Control patients made appointments with a physiotherapists on their own. No influence was exerted on timing and possible waiting time, so that these correspond to the usual GKV procedure.

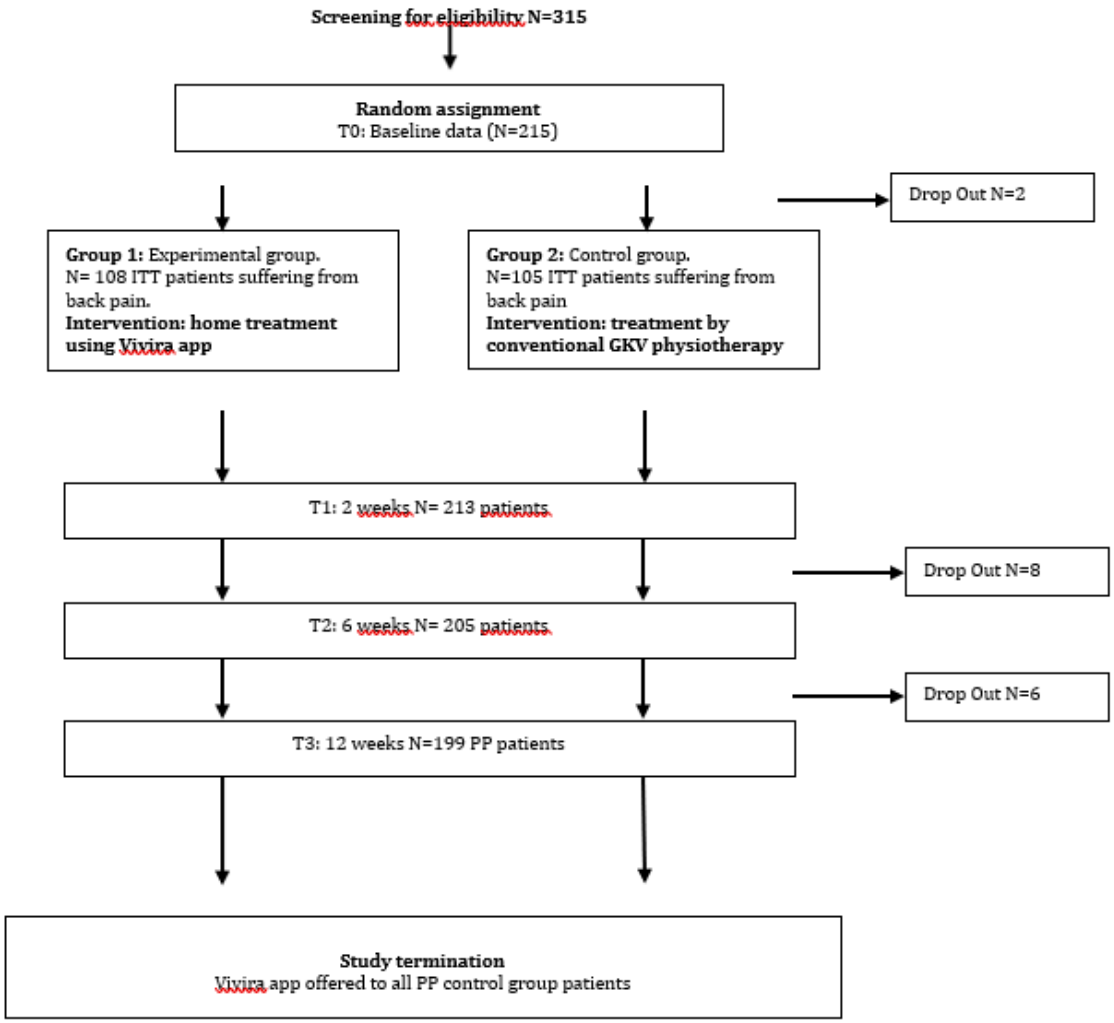
Blinding. Blinding was not feasible.

Prior and concomitant therapy. Other accompanying therapy was not influenced by the therapy with the app.

Overview

The subsequent scheme shows overall design and procedures of the study. Major steps were enrollment, allocation, follow-up and analysis.

After screening, the decision on inclusion and/or exclusion was made. The next step was the random assignment and thus the allocation to the experimental group or control group. Within the follow up data acquisitions are performed. After 12 weeks the study was completed for each individual patient. After last patient out, a complete per protocol analysis was performed.



Selection of Study Population

Screening

N= 315 individuals were screened. Screening was based on in- and exclusion criteria.

Inclusion criteria

Adult (18 years and older) patients

- with back (spine) pain (compare subsequent ICD-list)
- with baseline pain score of 4 or more

- with ability to provide the informed consent
- possessing a smart phone or tablet

Patients fulfilling all criteria and who provided the informed consent were taken as ITT-patients. The following list displays all backpain pathologies treatable by Vivira that possess an ICD code according to ICD 10 (German version):

ICD 10	Bezeichnung
M53.2	Instabilität der Wirbelsäule
M53.8	Sonstige näher bezeichnete Krankheiten der Wirbelsäule und des Rückens
M53.9	Krankheit der Wirbelsäule und des Rückens, nicht näher bezeichnet
M54.4	Lumboischialgie
M54.5	Kreuzschmerz - Inkl.: Lendenschmerz; Lumbago o.n.A.; Überlastung in der Kreuzbeingegend - Exkl.: Flankenschmerz-Hämaturie-Syndrom (N39.81); Lumbago durch Bandscheibenverlagerung (M51.2); Lumboischialgie (M54.4)
M54.6	Schmerzen im Bereich der Brustwirbelsäule - Exkl.: Schmerzen durch Bandscheibenschaden (M51.-)
M54.8	Sonstige Rückenschmerzen
M54.9	Rückenschmerzen, nicht näher bezeichnet - Inkl.: Rückenschmerzen o.n.A.
M99.02	Segmentale und somatische Funktionsstörungen: Thorakalbereich [thorakolumbal]
M99.03	Segmentale und somatische Funktionsstörungen: Lumbalbereich [lumbosakral]
M99.04	Segmentale und somatische Funktionsstörungen: Sakralbereich [sakrokokzygeal, sakroiliakal]
M99.82	Sonstige biomechanische Funktionsstörungen: Thorakalbereich [thorakolumbal]
M99.83	Sonstige biomechanische Funktionsstörungen: Lumbalbereich [lumbosakral]
M99.84	Sonstige biomechanische Funktionsstörungen: Sakralbereich [sakrokokzygeal, sakroiliakal]
M99.92	Biomechanische Funktionsstörung, nicht näher bezeichnet: Thorakalbereich [thorakolumbal
M99.93	Biomechanische Funktionsstörung, nicht näher bezeichnet: Lumbalbereich [lumbosakral]
M99.94	Biomechanische Funktionsstörung, nicht näher bezeichnet: Sakralbereich [sakrokokzygeal, sakroiliakal]
M42.0-	Juvenile Osteochondrose der Wirbelsäule - Inkl.: Scheuermann-Krankheit; Vertebra plana [Calvé-Krankheit] - Exkl.: Kyphose als Haltungsstörung
M42.1-	Osteochondrose der Wirbelsäule beim Erwachsenen
M42.9-	Osteochondrose der Wirbelsäule, nicht näher bezeichnet

Table 2: Spine pain pathologies treatable by Vivira and that possess an ICD code according to ICD 10. German version.

Exclusion criteria

General

- No pain, pain score 3 or less
- Previous movement therapy with a DiGA (Vivira, Kaia, etc.)

- Use of analgetics prior inclusion
- Pregnancy
- Limited legal or insufficient language capacity
- Patients who are not able to follow the exercise protocol
    - e.g. significantly impaired vision or blindness

Serious internal disorders

- Severe organ failure, e.g.
    - Condition after heart attack
    - Need for dialysis
    - Cardiovascular decompensation
    - Pulmonary insufficiency
- Inflammation
    - Past or present rheumatological disease
    - Acute inflammatory diseases
    - Feverish condition
- Coagulopathy
    - Thrombosis
    - Blood coagulation disorders including anticoagulant therapy

Musculoskeletal

- Any bone disease
- Injuries or surgery
    - Fresh bone or joint fractures
    - Injury to spinal column, knee or hip joint
    - Condition after
        - Spine, hip or joint surgery
        - Osteotomy (an operation to correct the axis of the leg)
        - Arthrodesis (joint stiffening) in one of the two knee or hip joints

- Inflammatory disease
    - Spinal column or joint inflammatory disease
    - Situation after spinal column or joint inflammatory disease
- Spinal tumor
- Osteochondrosis dissecans
- Bone necrosis
- Hip dysplasia
- Acute instability of the knee or hip joint
- Free joint bodies
- Disc pathology
    - Slipped disc
    - Acute herniated disc or other disorder with radiation to the legs (radiculopathy, sensorimotor failure)
    - Herniated disc in the past
- Clinically relevant bone marrow edema
- Osteoporosis

Neuropsychiatric

- Serious neurological disorders e.g.
    - Stroke
    - Paralysis
    - Multiple sclerosis
    - Convulsions
- Posture insecurity
    - Neurological motor disorders
    - Sensomotoric disorders
    - Vertigo
- Skin sensitivity disorder
- Psychoses
- Dementia
- Drug or alcohol abuse

Oncological

- Metastases of malignant tumors
- Acute malignant disease

Treatments

Treatments administered

The experimental group used Vivira. The level of exercises provided by Vivira depended on the feedback of the patient. The feedback thus leads to an extensive individualization of the therapy within the framework of the available interventions. The control treatment (comparator) consisted of the GKV-standard physiotherapy (TAU, treatment as usual).

Identity of investigational product
Vivira App

Method of assigning patients to treatment groups
To minimize stratification bias a stratified block randomization was chosen.

Selection of level in the study
The app's software offers a feedback mechanism, which allows the level of the exercises to be adjusted individually to the patient.

Selection of timing for each patient
The app's software allows an individual timing by the patient.

Blinding
Patients were naturally not blinded by the product.

Prior and concomitant therapy
Other accompanying therapy was not influenced by the therapy with the app.

Treatment compliance
Compliance with treatment motion sequences was asked for using the structured interview.

Treatment adherence
Adherence was asked for using the structured interview**.**

11. Efficacy Variables

11.1 Overview

Efficacy
Specific primary efficacy variable assessed

- Pain

Specific secondary efficacy variables assessed

- Quality of Life
- Functional score
- Mobility
- Stability
- Coordination

Further investigations incorporated

- Adherence
- Compliance
- Transient drug use

No laboratory tests were conducted. Subsequently all variables are tabulated:

11.2 Primary efficacy endpoint

The study used the German validated categorical spread VNRS according to the working group of Prof. Diener, Neurologische Universitätsklinik Essen (Aicher et al 2011[12]). In addition, the proposed German-language category summary was integrated within the high pain domain (Weber et al 2005)[13]. The VNRS has 5 spread German language categories.

12 Aicher B, H Peil, B Peil2 and H-C Diener (2011): Pain measurement: Visual Analogue Scale (VAS) and Verbal Rating Scale (VRS) in clinical trials with OTC analgesics in headache. Cephalalgia 32(3) 185–197.
13 M. Weber, J. Schüz, J. Kuball, H. Gamm, J. Jage: Schmerzerfassung bei invasiven diagnostischen Prozeduren. Schmerz 2005, 19:513–520.

Scale range	German pain category
0	Keine Schmerzen
1	Leichte Schmerzen
2-4	Mäßiger Schmerz
5-7	Starker Schmerz
8-10	Sehr starker Schmerz

Table 3: German pain category descriptions and correspondent scale ranges of the pain VNRS

11.3 Secondary outcomes

Mobility, Stability, Coordination and Total Functional Score

Mobility, stability, and coordination scores were calculated based on a patient´s reported ability to perform the following list of movements without pain and to meet the defined criteria. The scores for a patient feedback "yes" (confirming to meet the defined criteria without pain) are defined in the following table. Patient feedback "no" is scored with 0. Total score and individual scores for mobility, stability, and coordination were calculated. Score values range from 0 to 100. The movements within this assessment were selected and adapted (where required) from established movement screens:

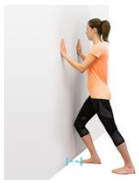
- SFMA – selective functional movement assessment (Cook et al., 2006a, b[14])
- FMS – functional movement screen (Cook 201015)
- DFT - Dorsiflexion Lunge Test (Bennell et al., 199816)

- Other tests (Tong et al., 201417; Magee, 201418; Schellenberg et al., 200719).

[14] Cook, G., Burton, L., & Hogenboom, B. (2006a) 'The use of fundamental movements as an assessment of function - Part 1', NAJSPT, 1(2), pp- 62-72
Cook, G., Burton, L., & Hogenboom, B. (2006b). 'Pre-participation screening: the use of fundamental movements as an assessment of function - Part 2', NAJSPT, 1(3), pp. 132-139.
[15] Cook, G. Movement: Functional Movement Systems: Screening, Assessment and Corrective Strategies. Aptos, CA: Target Publications; 2010.
[16] Bennell, K.L., Talbot, R.C., Wajswelner, H., Techovanich, W., Kelly, D.H., Hall, A.J. Intra-rater and inter-rater reliability of a weight-bearing lunge measure of ankle dorsiflexion. Aust J Physiother. 1998;44(3):175-180.

17 Tong, T.K., Wu, S., Nie, J. (2014) 'Sport-specific endurance plank test for evaluation of global core muscle function.' Phys Ther Sport, 15(1), pp. 58-63.
18 Magee, D.J. (2014) Orthopedic Physical Assessment. St Louis, Mo: Saunders Elsevier.
19 Schellenberg, K.L., Lang, J.M., Chan, K.M., Burnham, R.S. (2007) 'A clinical tool for office assessment of lumbar spine stabilization endurance: prone and supine bridge maneuvers.' Am J Phys Med Rehabil, 86(5), pp. 380-386.

Nr	Movement	Image	Criteria	Total Score	Mobility Score	Stability Score	Coordination Score
1	Cervical flexion		Chin touches sternum	6,6	10	0	0
2a	Cervical rotation - left		Chin touches left collar bone	3,3	5	0	0
2b	Cervical rotation - right	See 2a	Chin touches right collar bone	3,3	5	0	0
3	Multisegmental Flexion		Fingertips touch ground with extended knees	6,6	10	0	0
4a	Multisegmental Rotation - left		Heterolateral shoulder is visible in mirror with mutisegmental rotation, no lateralflexion for compensation	3,3	5	0	0
4b	Multisegmental Rotation - right	See 4a	Heterolateral shoulder is visible in mirror with mutisegmental rotation, no lateralflexion for compensation	3,3	5	0	0
5	Multisegmental Extension (wall)		Heels, pelvis, shoulder blades and back of the head are in contact with wall simultaneously	6,6	10	0	0
6	Extension + full shoulder anteversion		Full abduction of shoulder is possible while heels, buttocks, shoulder blades, back of the head or hands stay in contact with wall	6,6	10	0	20

7a	Knee to Wall - left		Knee touches the wall with approx. 10 cm distance between toes and wall	3,3	5	0	0
7b	Knee to Wall - right	See 7a	Knee touches the wall with approx. 10 cm distance between toes and wall	3,3	5	0	0
8a	In-Line Lunge - left		Backward lunge with knee touching the ground is possible, no evasive movements	3,3	5	0	10
8b	In-Line Lunge - right	See 8a	Backward lunge with knee touching the ground is possible, no evasive movements	3,3	5	0	10
9	Deep Squat		Buttocks touch heels, knees stay together	6,6	10	0	20
10	Heelsit		Buttocks touch heels, knees stay together	6,6	10	0	0
11a	Bird Dog - left		Position held at least 10 seconds	3,3	0	10	10
11b	Bird Dog - right	See 11a	Position held at least 10 seconds	3,3	0	10	10
12	Forearm Plank		Position held at least 60 seconds	6,6	0	20	0
13a	Side Plank – left		Position held at least 30 seconds	3,3	0	10	0
13b	Side Plank - right	See 13a	Position held at least 30 seconds	3,3	0	10	0

14	Shoulder Bridge		Position held at least 60 seconds	6,6	0	20	0
15a	One Legged Stand - left		One legged stand held for at least 10 seconds, no evasive movements of upper body	3,3	0	10	10
15b	One Legged Stand - right	See 15a	One legged stand held for at least 10 seconds, no evasive movements of upper body	3,3	0	10	10
Maximum Score				**100**	**100**	**100**	**100**

Table 4: Overview of mobility, stability, coordination and total functional score retrieved by the app.

Score evaluation. Resulting scores of total functional ability, mobility, stability and coordination were categorized as follows:

Score range	Ability within category
0-40	Low
41-70	Medium
71-90	High
91-100	Very high

Table 5: Score evaluation. Resulting scores of total functional ability, mobility, stability and coordination were categorized as indicated.

Quality of Life (QoL)

For QoL determination patients complete the German version of the generic SF-36 questionnaire (36-Item Short Form Health Survey [SF-36]) prior and after treatment at T0, T2 and T3. The SF-36 Health Survey is a widely used classical and validated inventory. The use of scales and scoring follows the SF-36 Health Survey Manual and Interpretation Guide.

Validated main outcome variables are the changes in PCS and MCS (PCS: physical component score; MCS: mental component score; Ellert et al. 2005 [20]). These two subscales can be summarized via the Z-score sum according to the proposal of O`Brien (1984[21]) for multiple endpoints and used as an overall endpoint.

11.4 Application schedule

Patients were asked to exercise at least three days per week, using the training plan provided in the Vivira App. Adherence to the training program was considered as long as there was at least one training performed per week.

The time of day for exercising was the patient's decision, no relation to meals had to be observed.

11.5 Persons responsible for the measurements

At the time points T0-T3 all study patients made their statements on rating score and questionnaires and performed and documented the movement tests. Structured interviews that included active recording of AEs and ARs were performed by the investigators.

11.6 Frequency and timing of measurements

The following table displays how frequency and timing of measurements were performed:

[20] Ellert U, Lampert T, Ravens-Sieberer U.: Measuring health-related quality of life with the SF-8. Normal sample of the German population. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2005 Dec;48(12):1330-7. doi: 10.1007/s00103-005-1168-5.

[21] O`Brien, P: Procedures for comparing samples with multiple endpoints. Biometrics (1984) 40, 1079-1087

Timing	Investigations
Screening	Inclusion criteria Exclusion criteria
T0 (baseline)	Clinical investigation Case history Pain intensity VNRS Mobility score Strength score Total functional score Coordination score QoL SF-36
T1	Structured medical interview Active AE/AR recording Pain intensity VNRS Mobility score Strength score Total functional score Coordination score QoL SF-36
T2	Structured medical interview Active AE/AR recording Pain intensity VNRS Mobility score Strength score Total functional score Coordination score QoL SF-36
T3	Structured medical interview Active AE/AR recording Pain intensity VNRS Mobility score Strength score Total functional score Coordination score QoL SF-36

Table 6: Frequency and timing of measurement

11.7 Persons other than the investigators responsible for evaluation of clinical outcomes

Dipl.-Stat. Michael Bulitta
CRM Biometrics, Rheinbach

11.8 Procedures (blinding, anonymization)

Blinding was not feasible for the patients. The data were collected by the investigators, pseudonymized, and made available for statistical analysis. After completion of the statistical analysis, all resulting metadata was completely anonymized and can be disclosed to manufacturer and authorities. The same anonymized data can be used for scientific publications

11.9 Appropriateness of measurements

The study used the German validated categorical spread VNRS according to the working group of Prof. Diener, Neurologische Universitätsklinik Essen (Aicher et al 2011[22]). In addition, the proposed German-language category summary was integrated within the high pain domain (Weber et al 2005)23. The VNRS has 5 spread German language categories.

An early German-language VNRS with 5 categories was already published in 1988 by Gablenz et al. [24] and compared with the VAS. Already at that time there was a r=0.84 for VAS/VNRS in German. Weber et al. 2005[25] also found a very good correlation between numerical rank and German verbal categorical scales in N=263 (Spearman correlation coefficient 0.86). They noticed that the numerical digit 3 as transition between mild and moderate pain as well as the digits 5 and 6 as transition between moderate and severe pain led to overlaps between the categories of the verbal scale in German. This was taken up by the Diener working group[26]. Due to the small number of cases in the high pain range, the categories "very severe pain" and "strongest possible pain" were combined by Weber et. al (2005)[27]. When compared, Weber et al. showed 73.6% agreement. The weighted κ-measure accordingly means a good agreement with 0.72 (95% confidence interval 0.66-0.79).

Using the VAS the working group of Diener in Essen also validated the German-language VNRS (Aicher et al 2011[28]). Data for this analysis was collected in a randomized, placebo-controlled, double-blind, multicenter, parallel-group study with six treatment arms. Patients included in the study were trained in the use of the VAS by naming categories of headache in German. These data were used to assess the degree of order consistency between the VAS and VNRS, to derive cutoff points for rescaling the continuous VAS to a discrete ordinal scale using receiver operating characteristic method, and to determine interrater and test-retest reliability. Test-retest evaluations were possible over several weeks or sometimes even months. Test-retest assessments for all categories of VNRS were comparable in magnitude to the observed interindividual reliability. This resulted in a shift to higher or lower categories of the VAS in no more than 20.6% of patients using the cut-off points. Shifts occurred mainly between adjacent categories. Depending on the VNRS category, complete agreement between test and retest (re-test) was demonstrated in 61.0-91.4% of patients.

[22] Aicher B, H Peil, B Peil2 and H-C Diener (2011): Pain measurement: Visual Analogue Scale (VAS) and Verbal Rating Scale (VRS) in clinical trials with OTC analgesics in headache. Cephalalgia 32(3) 185–197.
[23] M. Weber, J. Schüz, J. Kuball, H. Gamm, J. Jage: Schmerzerfassung bei invasiven diagnostischen Prozeduren. Schmerz 2005, 19:513–520
[24] Gablenz, B. Heinen, D. Kirsch und E. Lanz: Schmerzerfassung, Beschreibung einer neuen Methode. Der Schmerz (1988) 2:144-150
[25] M. Weber, J. Schüz, J. Kuball, H. Gamm, J. Jage: Schmerzerfassung bei invasiven diagnostischen Prozeduren. Schmerz 2005, 19:513–520.
[26] Aicher B, H Peil, B Peil2 and H-C Diener (2011): Pain measurement: Visual Analogue Scale (VAS) and Verbal Rating Scale (VRS) in clinical trials with OTC analgesics in headache. Cephalalgia 32(3) 185–197
[27] M. Weber, J. Schüz, J. Kuball, H. Gamm, J. Jage: Schmerzerfassung bei invasiven diagnostischen Prozeduren. Schmerz 2005, 19:513–520
[28] Aicher B, H Peil, B Peil2 and H-C Diener (2011): Pain measurement: Visual Analogue Scale (VAS) and Verbal Rating Scale (VRS) in clinical trials with OTC analgesics in headache. Cephalalgia 32(3) 185–197.

Approximately 75% of patients registered pain intensity on the VAS in the same order as indicated on the VNRS. In 12.6% of patients, designations were confused regarding pain intensity. Using cutoff point analysis, the cutoff points were determined. The study showed that the VNRS categories cannot be equally spaced on the VAS and that, contrary to previous assumptions, the pain intensity descriptors are less clear. Thus, using the German language a language specific spreading is necessary. Thus, according to Diener and coworkers (Aicher et al 2011[29]) specific German wording spreading was used in the present investigation (see "Scale Range" in the table of chapter “Primary efficacy endpoint”).

11.10 Data Quality Assurance

Data quality was provided by the monitors.

12. STUDY PATIENTS

Disposition of Patients

- Screening N= 315
- Random assignment N= 215
- Allocation to the
  - experimental group N= 108
  - control group N= 105
- After 2 weeks N= 213
- After 6 weeks N= 205
- After 12 weeks N= 199

Intention-to-Treat (ITT)-Patients

All patients who were randomized and showed values for the primary variable at baseline were included in the ITT analysis. If at further points in time observations were missing, they were imputed using the last-observation-carried-forward-method. The ITT analysis is the primary one in this study.

29 Ibid

Modified ITT population (mITT)

All patients who were randomized and showed values for the primary variable at baseline and after 2 weeks were included in the mITT analysis. Missing values after 6 weeks and after 12 weeks were imputed by means of the last-observation-carried-forward method (LOCF) for the primary variable. Missing values for other variables were not replaced. This analysis was only carried out for the primary variable.

Per protocol population (PP)

In the PP analysis all patients who were randomized and showed no missing data for the primary variable were assessed.

Safety population (SAF)

All patients who worked with the Vivira App or underwent physiotherapy were actively assessed for safety. The SAF population was identical to the ITT population.

Protocol Deviations

Patients who entered the study even though they did not satisfy the entry criteria at P0

- None

Patients who developed withdrawal criteria during the study and were withdrawn

- N=7

All protocol deviation patients are included in the drop-outs.

Group	Periode of Time (T0-T3)	Gender	Age	Details
Vivira App	0-1	m	74	Body weight increased too much for performing excercises
	0-1	w	57	Mental disease
	1-2	w	49	Mental disease
	1-2	m	62	Stroke
	0-1	w	63	Telephone problems
Physiotherapy	0-1	m	61	Bone fracture
	1-2	m	70	Cancer

Table 7: Protocol Deviations

Drop-Outs

Group	Periode of Time (T0-T3)	Gender	Age	Details	AR
Vivira App	0-1	m	74	Body weight increased too much for performing excercises	No
	0-1	w	57	Mental disease	No
	1-2	w	49	Mental disease	No
	0-1	m	68	Nausea during app excercises	Yes
	2-3	m	79	Patient has discontinued study	n/a
	1-2	m	30	Patient has discontinued study	n/a
	0-1	m	78	Patient has discontinued study	n/a
	1-2	m	36	Patient has discontinued study	n/a
	1-2	m	62	Stroke	No
	0-1	w	63	Telephone problems	n/a
Physiotherapy	1-2	m	70	Cancer	No
	1-2	w	51	Patient has discontinued study	n/a
	1-2	w	56	Patient has discontinued study	n/a
	0-1	m	63	Patient has discontinued study	n/a
	0-1	m	72	Patient has discontinued study	n/a
	0-1	m	61	Bone fracture	No

Table 8: Drop Outs

13. EFFICACY EVALUATION

Data Sets Analyzed

ITT	N=213
mITT	N=207
PP	N=199
Male	N=100
Female	N=113
Anatomic localization	N=213
Age	N=213

Baseline Characteristics

ITT population
N=213 from the original 215 ITT-patients were analyzed. N=2 withdrew their consent prior to the baseline assessment and intervention, respectively. On average, both study groups were very well balanced regarding the variable age. Furthermore, the distribution of the age categories was similar between both study groups. 100 males and 113 females were analysed. There were some slight differences between the two study groups regarding the gender variable. As a result of the block randomization slightly more males and lesser females were enrolled in the Vivira App group compared to the Physiotherapy group (see table).

			Vivira App (n=108)	**Physiotherapy (n=105)**
Age (years)	Mean		57.37	57.26
	SD		13.80	13.50
	Median		60.50	59.00
Age categories (years)	18-35	n (%)	11 (10.19)	9 (8.57)
	36-45		12 (11.11)	10 (9.52)
	46-55		16 (14.81)	18 (17.14)
	56-65		33 (30.56)	38 (36.19)
	66-75		29 (26.85)	25 (23.81)
	76-85		7 (6.48)	5 (4.76)
Gender	male	n (%)	57 (52.78)	43 (40.95)
	female	n (%)	51 (47.22)	62 (59.05)

Table 9: Demography distribution (ITT). n = sample size, SD = standard deviation.

PP population

From the 213 ITT patients N=199 remained as PP population. This corresponds to the prestudy case number calculation. The prestudy case number calculation had revealed N=198 participants to be required for the study. Demography results for the PP population did not differ from those of the ITT population (see subsequent Table).

			Vivira App (n=98)	**Physiotherapy (n=101)**
Age (years)	Mean		57.14	57.16
	SD		13.53	13.68
	Median		60.00	59.00
Age categories (years)	18-35	n (%)	10 (10.20)	9 (8.91)
	36-45		11 (11.22)	10 (9.90)
	46-55		15 (15.31)	17 (16.83)
	56-65		30 (30.61)	36 (35.64)
	66-75		27 (27.55)	24 (23.76)
	76-85		5 (5.10)	5 (4.95)
Gender	male	n (%)	50 (51.02)	41 (40.59)
	female	n (%)	48 (48.98)	60 (59.41)

Table 10: Demography results (PP). n = sample size, SD = standard deviation.

Clinical findings

Both study groups were relatively well balanced regarding the localization of pain in the population excluding the drop-outs. The majority of cases suffered from lumbar spine pain 7 lower back pain in both groups (Vivira App: Lumbago/Lumboichalgia 40.74 %/40.74 %, Physiotherapy: 42.86 %/ 31.43 %), followed by thoracic spine pain / upper back region and its variants in the ITT population (Vivira App: 18.53 %, Physiotherapy: 25.71 %).

Pain localization		Vivira App (n=108) n (%)	Physiotherapy (n=105) n (%)
Thoracic spine ICD M54.9		13 (12.04)	13 (12.38)
Thoracic spine and brachialgia ICD M54.9		6 (5.56)	5 (4.76)
Thoracic and cervical spine ICD M54.9		1 (0.93)	9 (8.57)
Lumbar spine	Lumbago ICD M54.5	44 (40.74)	45 (42.86)
	Lumbagoichalgia ICD M54.4	44 (40.74)	33 (31.43)

Table 11: Localization of pain (ITT). n = sample size.

The corresponding results for the PP Population were similar to those of the ITT analysis (see next table).

Localization		Vivira App (n=98) n (%)	Physiotherapy (n=101) n (%)
Thoracic spine ICD M54.9		11 (11.22)	11 (10.89)
Thoracic spine and brachialgia ICD M54.9		5 (5.10)	5 (4.95)
Thoracic and cervical spine ICD M54.9		1 (1.02)	8 (7.92)
Lumbar spine	Lumbago ICD M54.5	40 (40.82)	44 (43.56)
	Lumbagoichalgia ICD M54.4	41 (41.84)	33 (32.67)

Table 12: Localization of symptoms (PP). n = sample size.

Regarding further clinical findings both study groups were also relatively well balanced in the ITT population (see next table). The results of the PP analysis regarding medical history and clinical findings did not differ from those of the ITT analysis (not shown).

Most frequent pathological clinical findings were lordosis (15 (13.89)/20 (19.05)), massive kyphosis and spine instability (14 (12.96)/10 (9.52)) in the thoracic and thoracolumbal region and lordosis (8 (7.41)/12 (11.43) as well as massive kyphosis, spine instability and pelvic obliquity (6 (5.56/-)) in the lumbar, lumbosacral, sacral, sacrococcygeal and sacroiliacal region. Atrophic musculature contributing to spine instability was found in 28 (25.93)/27 (25.71) of the cases.

Localization	Spine findings		Vivira App (n=108) n (%)	Physiotherapy (n=105) n (%)
Thoracic Thoracolumbal	Physiologic	n (%)	79 (73.15)	75 (71.43)
	Lordosis		15 (13.89)	20 (19.05)
	Massive kyphosis Spine instability ICD M99.02; M99.82; M99.92; M53.2; M53.8; M53.9; M54.8		14 (12.96)	10 (9.52)
Lumbar Lumbosacral Sacral Sacrococcygeal Sacroiliacal	Physiologic		94 (87.04)	93 (88.57)
	Lordosis		8 (7.41)	12 (11.43)
	Massive kyphosis Spine instability Pelvic obliquity ICD M 99.83; 99.84; 99;03; 99.04; M53.2; M53.8; M53.9; M54.8		6 (5.56)	-
Musculature	Atropic Spine instability M53.2; M53.8; M53.9; M54.8		28 (25.93)	27 (25.71)
	Physiologic		80 (74.07)	78 (74.29)

Table 13: Other clinical findings (ITT). n = sample size.

14. MEASUREMENTS OF TREATMENT ADHERENCE AND COMPLIANCE

Adherence. As shown in the subsequent figure Vivira PP patients (N=98) used the app on 1 to 7 days per week. Most frequent use was at 6 days. Mean was 5.77 days which corresponds to an adherence of 82,4% (5.77/7 days per week). In the control group all PP patients (N=101) received physiotherapy between 5 and 24 sessions (mean: 6.94 sessions), the majority of them 6 sessions. Adherence to the physiotherapy group was defined as the percentage of the used number of 695 sessions vs. the planned number of 714 sessions resulting in an adherence of 97.33%.

Compliance. 10/98 Vivira PP patients performed the exercises different from how the app prompted them to execute the exercises. Consequently, in the Vivira group compliance was calculated as the number of patients who performed the exercises as proposed by the app (N=89) vs. the total number of patients (N=98). Thus, compliance was 90.81%. In the physiotherapy group the number of PP patients who have been assigned additional home exercises by the therapist (N=82) was determined vs. the number of those who really performed the exercises as proposed by the therapist (N=68). This results in a compliance of 82.93%.

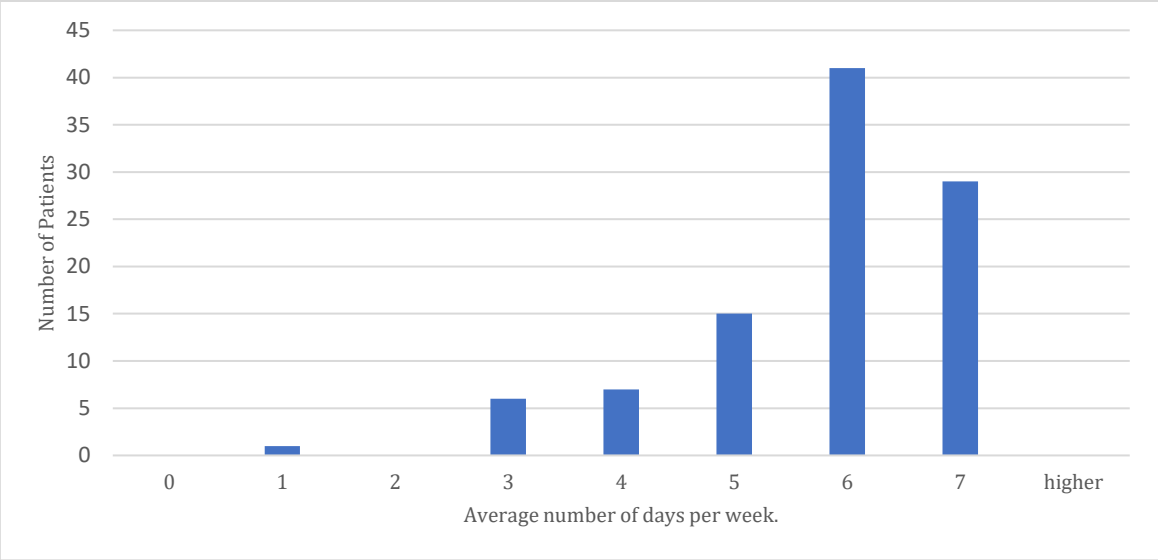


Figure 1: Frequency of weekly day use of the app. Frequency is indicated as the number of users (absolute number of Vivira PP patients).

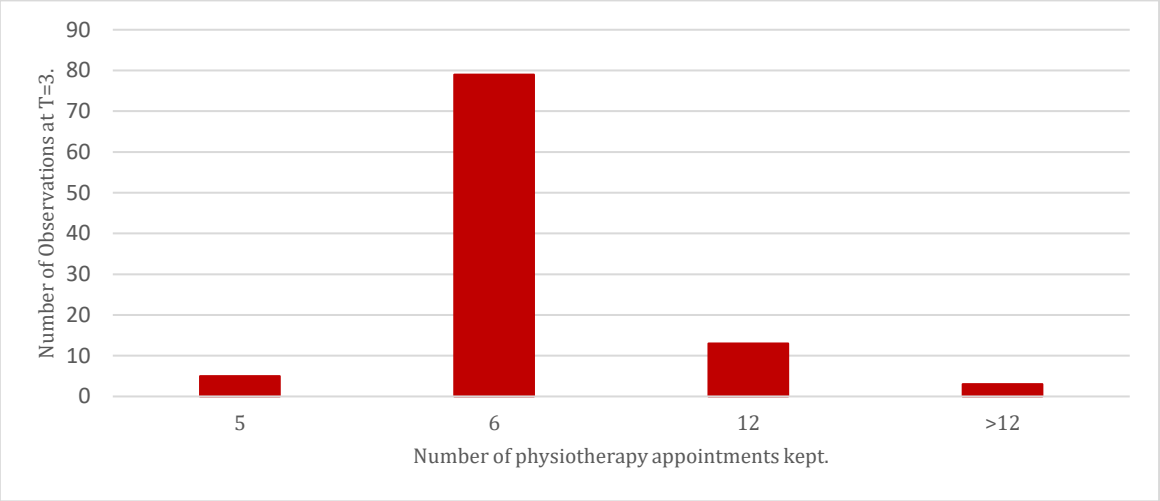


Figure 2: Histogram of absolute number of physiotherapy sessions completed within 12 weeks. PP patients

ViviRa PP Group N=98	**Physiotherapy PP Group** N=101
Adherence	**Adherence**
5.77 days/7 days per week (82,4%)	n= 695/714 sessions (97.33%)
Compliance	**Compliance**
N=89/98 patients (90,81%)	N=68/82 patients (82.93%)

Table 14: Adherence and compliance of both PP groups. Adherence: For the Vivira PP group the mean value/week was divided by the maximally possible 7 days/week. Adherence to the physiotherapy group is defined as the percentage of used number of sessions vs. the planned number of sessions. Compliance: In the Vivira group the number of patients who performed exercises vs. the number of patients who performed them as proposed by the app. In the physiotherapy group compliance is the number of patients who have been assigned additional home exercises by the therapist vs. the number of those who really performed the exercises as proposed by the therapist.

	Vivira Group **Days per week/12 weeks**	**Physiotherapy Group** **Number of sessions/12 weeks**
Min	1.00	1.00
Mean	5.77	6.94
Median	6.00	6.00
Max	7.00	24.00
StD	1.21	2.94

Table 15: Number of applications. For the ViviRa PP group the average weekly usage of the app in days is indicated. For physiotherapy PP group the number of physiotherapy appointments kept during 12 weeks is displayed.

Therapy start. All Vivira patients (N=98/100%) had started the exercises at time T1 (2 weeks). No earlier than after time point T2 (6 weeks) all physiotherapy patients had also started their therapy and completed it by time point T3 (12 weeks) with the above mentioned adherence and compliance.

15. EFFICACY RESULTS

15.1 Primary efficacy variable pain score (VNRS)

Primary efficacy variable was pain which was determined using the validated verbal numerical rating scale (VNRS). In both groups results of the statistical analyses (displayed in the subsequent tables and figures including mean (± 95 % CI) profiles and box-and-whisker plots) reveal a mean pain decrease from baseline after 2, 6 and 12 weeks in the ITT population. In the Vivira group pain decreased by -2.47 score points (-38.01 %, week 2), -2.92 score points (-45.73 %, week 6) and -3.35 score points (-53.11 %, week 12). In the physiotherapy group this was -0.33 score points (- 2.45 %, week 2), - 0.58 score points (7.14 %, week 6) and -0.91 score points (14.62 %, week 12). In both groups pain score differences between T0 and T2 (6 weeks) and T3 (12 weeks) were statistically significant. Statistical significance may be derived directly from the confidence interval[30] . T-tests at T1, T2 and T3 all revealed $p < 0.001$ for the intragroup changes in the Vivira group. For the physiotherapy intragroup changes produced significant p-values at T2 and T3 of 0.015 and 0.0003.

Cohen's d was used for a quantitative and metric-free estimation of the effect size. According to Cohen[31] 0.20 is a small, 0.50 a medium and 0.80 a large effect size. Thus, with Cohen´s d values from 1.44-1.76 the present investigation reveals a large intragroup effect size in the Vivira group and with 0.22-0.52 small to medium effects sizes in the physiotherapy group. In both groups the statistically significant clinical effect sizes are sufficiently high to produce a sufficient clinical effect strength[32].

Confirmatory group comparison reveals that in the Vivira App group the average pain decrease was significantly more pronounced than in the physiotherapy group (week 2: -2.47 score points (-38.01 %, Vivira App) versus -0.33 score points (- 2.45 %, physiotherapy), week 6: -2.92 score points (-45.73 %, Vivira App) versus 0.58 score points (7.14 %, physiotherapy), week 12: -3.35 score points (-53.11 %, Vivira App) versus -0.91 score points (14.62 %, physiotherapy), respectively). Group differences are highly significant (t-Test: $p<0.0001$ from 2-12 weeks therapy period) indicating superiority of the Vivira app. This was confirmed by the sensitivity analysis (Mann-Whitney test: $p< 0.0001$ for all timepoints). Using Cohen's delta group comparison of change of VNRS from baseline to weeks 2, 6 and 12 revealed intergroup

[30] Statistical significance with a 5%-level is given if CI values do not include the value 0 (zero).
[31] C.f.: Cohen J.: A power primer. In: Psychological Bulletin. Band 112, 1992, S. 155–159.
[32] Prof. Dr. Weise

values from 0.98-1.08, suggesting a large effect size of difference between both groups. Thus, this supports superiority of Vivira over conventional physiotherapy.

The later demonstrated shift analysis differentiating responders from non-responders revealed that the higher number of exacerbations in the physiotherapy group contributes significantly to the superiority of Vivira.

Furthermore, after 2 weeks T0/T1 score difference for Vivira was already significant whereas for physiotherapy difference was not yet significant. Contributing to this significant group difference at time T1 is the fact that at this time physiotherapy had generally not yet started, whereas Vivira patients were able to start their exercises without any time delay.

Similar results were obtained for the mITT and the PP populations.

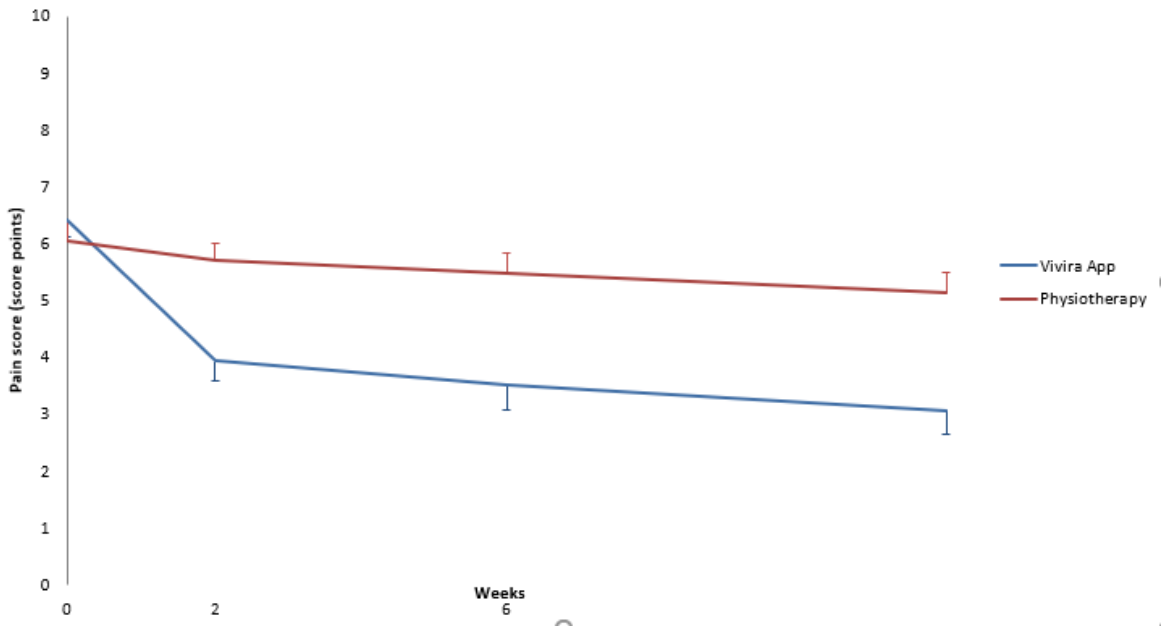
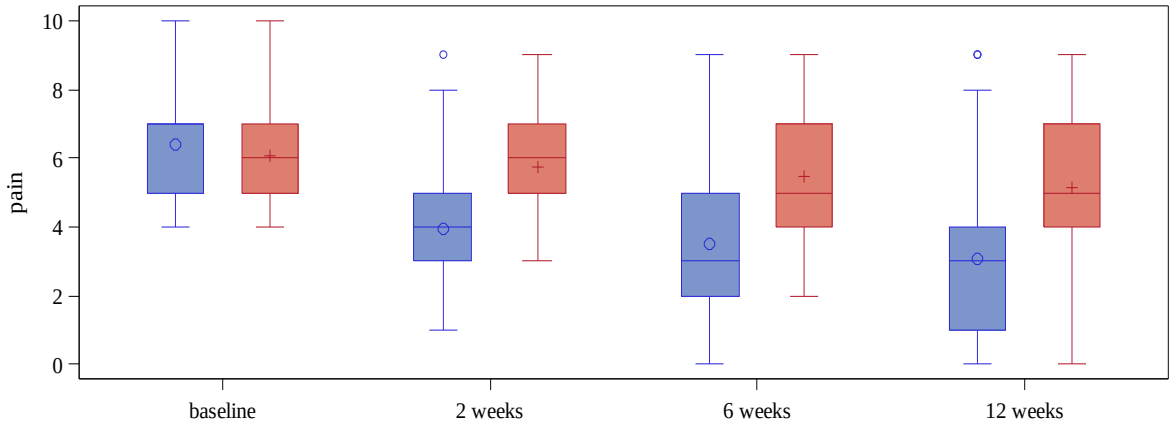


Figure 3: Pain score (VNRS) over time (ITT).

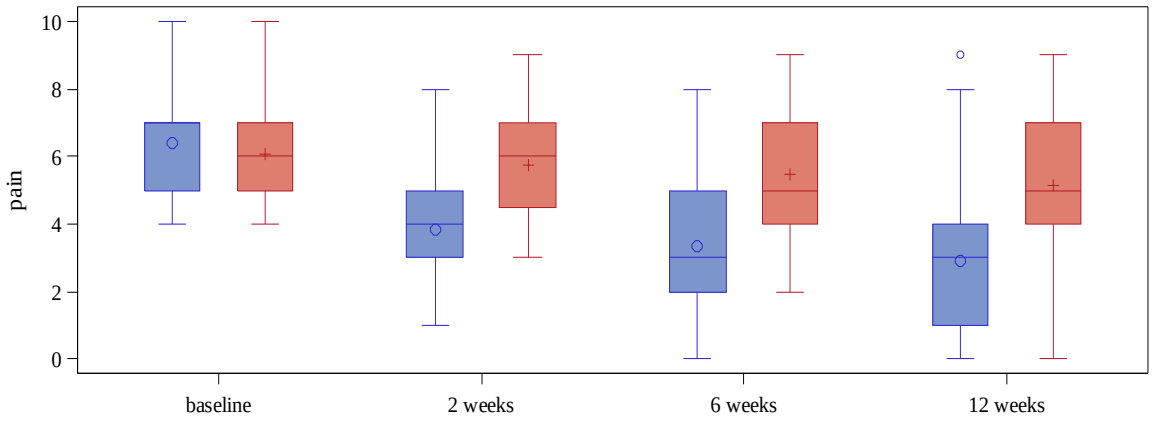
a.

Figure: Pain score values (ITT)



b.

Figure: Pain score values (mITT)



c.

Figure: Pain score values (PP)

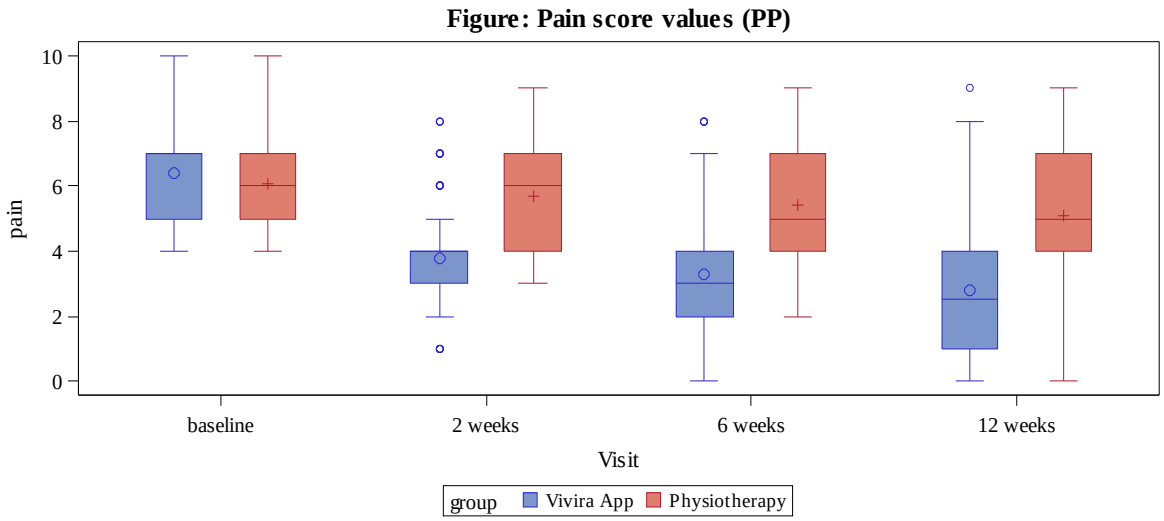


Figure 4: Pain score (VNRS) over time – boxplot (a: ITT; b: mITT; c: PP). In both study groups a mean decrease from baseline in pain score (VNRS) could be observed after 2, 6 and 12 weeks in the ITT population, respectively. However, the average decrease was more in the Vivira App group than in the Physiotherapy group.

Pain score (VNRS) (score points)	Vivira App (n=108)				Physiotherapy (n=105)			
	Baseline	Week 2	Week 6	Week 12	Baseline	Week 2	Week 6	Week 12
Mean	6.42	3.94	3.50	3.06	6.05	5.71	5.47	5.13
SD	1.65	1.79	2.21	2.18	1.64	1.48	1.80	1.91
Median	7.00	4.00	3.00	3.00	6.00	6.00	5.00	5.00
Cohen's delta: change from baseline	-	1.44	1.52	1.76	-	0.22	0.34	0.52
p-value	-	< 0.001	< 0.001	< 0.001	-	0.123	0.015	0.0003
Cohen's delta: treatment effect	-	1.08	0.98	1.01				
p-value	-	< 0.001	< 0.001	< 0.001				

Table 16: Pain score VNRS (ITT). VNRS = Verbal numerical rating scale, n = sample size, SD = Standard deviation.

a.

		Vivira App (n=108)			Physiotherapy (n=105)		
		Week 2	Week 6	Week 12	Week 2	Week 6	Week 12
Absolute pain score (VNRS) change (score points)	Mean	-2.47	-2.92	-3.35	-0.33	-0.58	-0.91
	SD	1.74	2.07	2.05	1.42	1.65	1.50
	Median	-2.00	-3.00	-3.00	0	0	-1.00
	95 % CI	-2.80; -2.14	-3.31; -2.52	-3.74; -2.96	-0.61; -0.06	-0.90; -0.26	-1.20; -0.62
	Cohen's delta treatment effect	1.35	1.26	1.37			
	p-value	< 0.001	< 0.001	< 0.001			
Relative pain score (VNRS) change (%)	Mean	-38.01	-45.73	-53.11	-2.45	-7.14	-14.62
	SD	22.86	30.63	29.49	24.21	28.26	25.28
	Median	-41.43	-50.00	-57.14	0	0	-16.67
	95 % CI	-42.37; -33.65	-51.57; -39.89	-58.74; -47.49	-7.14; +2.23	-12.61; -1.67	-19.51; -9.73
	Cohen's delta treatment effect	1.51	1.31	1.40			
	p-value	< 0.001	< 0.001	< 0.001			

b.

		Vivira App (n=103)			Physiotherapy (n=104)		
		Week 2	Week 6	Week 12	Week 2	Week 6	Week 12
Absolute pain score (VNRS) change (score points)	Mean	-2.59	-3.06	-3.51	-0.34	-0.59	-0.92
	SD	1.69	2.02	1.95	1.43	1.65	1.51
	Median	-2.00	-3.00	-3.00	0	0	-1.00
	95 % CI	-2.92; -2.26	-3.45; -2.66	-3.90; -3.13	-0.61; -0.06	-0.91; -0.27	-1.22; -0.63
	Cohen's delta treatment effect	1.44	1.35	1.50			
	p-value	< 0.001	< 0.001	< 0.001			
Relative pain score (VNRS) change (%)	Mean	-39.85	-47.95	-55.69	-2.48	-7.21	-14.76
	SD	21.77	29.61	27.70	24.33	28.39	25.37
	Median	-42.86	-50.00	-57.14	0	0	-16.67
	95 % CI	-44.11; -35.60	-53.74; -42.16	-61.11; -50.28	-7.21; 2.26	-12.73; -1.69	-19.69; -9.83
	Cohen's delta treatment effect	1.62	1.40	1.54			
	p-value	< 0.001	< 0.001	< 0.001			

c.

		Vivira App (n=98)			Physiotherapy (n=101)		
		Week 2	Week 6	Week 12	Week 2	Week 6	Week 12
Absolute pain score (VNRS) change (score points)	Mean	-2.61	-3.09	-3.57	-0.37	-0.62	-0.97
	SD	1.71	2.04	1.96	1.43	1.66	1.50
	Median	-2.00	-3.00	-3.50	0	0	-1.00
	95 % CI	-2.96; -2.27	-3.50; -2.68	-3.97; -3.18	-0.65; -0.08	-0.95; -0.30	-1.27; -0.67
	Cohen's delta treatment effect	1.43	1.34	1.50			
	p-value	< 0.001	< 0.001	< 0.001			

Relative pain score (VNRS) change (%)	**Mean**	-40.32	-48.69	-56.82	-2.89	-7.76	-15.54
	SD	22.04	29.95	27.69	24.53	28.58	25.29
	Median	-42.86	-50.00	-62.50	0	0	-16.67
	95 % CI	-44.74; -35.90	-54.69; -42.68	-62.38; -51.27	-7.73; +1.95	-13.41; -2.12	-20.53; -10.55
	Cohen's delta treatment effect	1.61	1.40	1.56			
	p-value	< 0.001	< 0.001	< 0.001			

Table 17: Absolute and relative pain score (VNRS) changes from baseline (a: ITT; b: mITT; c: PP). In both study groups a mean decrease from baseline in pain score (VNRS) could be observed after 2, 6 and 12 weeks in the ITT population, respectively. However, the average decrease was more in the Vivira App group than in the Physiotherapy. VNRS = Verbal numerical rating scale, n = sample size, SD = Standard deviation, CI = confidence interval.

a.

Treatment effect		**Week 2**	**Week 6**	**Week 12**
Absolute pain score changes (score points)	**Mean**	-2.14	-2.34	-2.44
	95 % CI	-2.57; -1.71	-2.84; -1.83	-2.92; -1.95
	p	<0.0001	<0.0001	<0.0001
	Standardized |Δ|	1.34	1.25	1.35
Relative pain score changes (%)	**Mean**	-35.56	-38.59	-38.49
	95 % CI	-41.92; -29.19	-46.56; -30.63	-45.92; -31.06
	p	<0.0001	<0.0001	<0.0001

b.

Treatment effect		**Week 2**	**Week 6**	**Week 12**
Absolute pain score changes (score points)	**Mean**	-2.26	-2.47	-2.59
	95 % CI	-2.68; -1.82	-2.98; -1.97	-3.07; -2.11
	p	<0.0001	<0.0001	<0.0001
	Standardized |Δ|	1.44	1.34	1.48
Relative pain score changes (%)	**Mean**	-37.37	-40.74	-40.93
	95 % CI	-43.71; -31.05	-48.69; -32.79	-48.21; -33.64
	p	<0.0001	<0.0001	<0.0001

c.

Treatment effect		Week 2	Week 6	Week 12
Absolute pain score changes (score points)	**Mean**	-2.24	-2.46	-2.60
	95 % CI	-2.68; -1.80	-2.98; -1.94	-3.08; -2.11
	p	<0.0001	<0.0001	<0.0001
	Standardized |Δ|	1.41	1.33	1.49
Relative pain score changes (%)	**Mean**	-37.43	-40.92	-41.28
	95 % CI	-43.95; -30.91	-49.11; -32.74	-48.69; -33.87
	p	<0.0001	<0.0001	<0.0001

Table 18: Treatment effects for pain score (VNRS) changes from baseline (a: ITT; b; mITT; c: PP).
VNRS = Verbal numerical rating scale, CI = Confidence interval, p = p-value.

Shift tables for pain scores

This chapter presents the shift tables for the pain score change from baseline to weeks 2, 6 and 12 in the ITT population.

2 Weeks. In the Vivira App group 97 patients improved by at least one score point (89.8 %) and none of the patients deteriorated in this group. In the physiotherapy group only 42 patients (40.0 %) improved by at least 1 score point. Moreover, in this group deteriorations were observed in 24 patients (22.9 %).

a.

Vivira App (n=108)									
Pain at baseline (score-points)	**Pain after Week 2 (Score Points)**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4	2	8	6	2	-	-	-	-	-
5	-	3	5	9	1	-	-	-	-
6	-	1	5	4	3	2	-	-	-
7	4	3	6	9	1	4	4	-	-
8	-	-	1	3	1	2	4	1	-
9	-	1	2	3	2	1	1	1	1
10	-	-	1	-	-	-	-	1	-

b.

Physiotherapy (n=105)									
Pain at baseline (score-points)	**Pain after Week 2 (Score Points)**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4			3	12	2	5	1	-	-
5			1	4	11	5	3	1	-
6			-	-	6	4	2	-	-
7			1	4	3	5	7	4	-
8			-	1	-	3	3	5	1
9			-	-	1	-	2	3	0
10			-	-	-	-	1	1	-

Table 19: Shift table for pain scores after 2 weeks (ITT). (a: Vivira; b: physiotherapy). Numbers indicate the number of patients with improvements (pain reductions) of at least 1 scorepoint.

6 weeks. After 6 weeks, the pain score improved by 1 score point or more in 95 cases (88,2 %) in the Vivira App group (ITT). In rare cases (n=3, 2.8 %) deteriorations were observed in this group. A pain score of 0 or 1 score points was observed in 23 cases after 6 weeks. In the physiotherapy group only 50 (47.6 %) cases showed an improvement of 1 or more score points after 6 weeks (ITT). Deteriorations were documented in 26 cases (24.8 %).

a.

Vivira App (n=108)										
Pain at baseline (score-points)	**Pain after Week 6 (Score Points)**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**1**
4	-	7	7	2	1	1	-	-	-	-
5	1	1	3	4	6	2	-	-	1	-
6	1	4	3	2	4	-	1	-	-	-
7	3	5	3	7	5	2	2	4	-	-
8	-	-	-	1	2	3	-	4	2	-
9	-	1	2	1	1	2	-	4	-	1
10	-	-	-	-	1	-	-	-	1	-

b.

Physiotherapy (n=105)										
Pain at baseline (score-points)	**Pain after Week 6 (score points)**									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4	-	-	1	4	9	6	1	1	1	-
5	-	-	-	4	5	10	4	1	1	-
6	-	-	-	2	3	3	0	2	2	-
7	-	-	1	2	2	7	3	3	6	-
8	-	-	-	1	-	2	1	2	6	1
9	-	-	-	-	-	1	-	2	2	1
10	-	-	-	-	-	-	-	-	1	1

Table 20: Shift table for pain scores after 6 weeks (ITT), (a: vivira; b: physiotherapy). Numbers indicate the number of patients with improvements (pain reductions) of at least 1 scorepoint.

12 weeks. After 12 weeks, the pain score improved by 1 score point or more in 99 cases (91.7 %) in the Vivira App group (ITT). Deteriorations were not observed. However, in the physiotherapy group only 63 (60.0 %) cases showed an improvement of 1 or more score points after 6 weeks (ITT). Deteriorations were documented in 18 cases (17.1 %).

a.

Vivira App (n=108)										
Pain at baseline (score-points)	**Pain after Week 12 (Score Points)**									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4	3	9	4	0	2	-	-	-	-	-
5	1	3	1	6	5	2	-	-	-	-
6	-	3	6	3	2	-	1	-	-	-
7	4	7	3	6	4	1	4	2	-	-
8	-	-	-	4	1	-	3	3	1	-
9	-	2	3	2	0	2	1	0	1	1
10	-	-	-	-	-	-	1	-	-	1

b.

Physiotherapy (n=105)										
Pain at baseline (score-points)	**Pain after Week 12 (Score Points)**									
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4	-	-	3	11	7	1	1	-	-	-
5	1	1	2	2	5	7	5	2	-	-
6	-	-	-	-	2	6	1	2	1	-
7	-	-	-	1	4	6	3	6	4	-
8	-	-	-	1	0	2	3	3	2	-
9	-	-	-	-	1	-	-	4	-	1
10	-	-	-	-	-	-	-	-	1	1

Table 21: Shift table for pain scores after 12 weeks (ITT). (a: Vivira; b: physiotherapy). Numbers indicate the number of patients with improvements (pain reductions) of at least 1 scorepoint.

Summary. The next figure and the next table summarize the shift tables of the pain scores in the ITT population. The corresponding results of the mITT and PP analyses did not differ from those of the ITT population.

After 12 weeks, after both groups had been treated with similar adherence and compliance, N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower, at N=63 (60.0%). At the earlier time points T1 (2 weeks) and T2 (6 weeks), the differences were even greater: here, structural and procedural influences may play a significant role in the differences in results between the groups: in the control group, the start of therapy was significantly delayed by waiting times until the first therapy session with the physiotherapist. Only after time point T2 (6 weeks) all physiotherapy patients had also started their therapy and completed it by time point T3 (12 weeks) with the above mentioned adherence and compliance.

		Vivira App (n=108)			Physiotherapy (n=105)		
		Improved	**No change**	**Exacerbation**	**Improved**	**No change**	**Exacerbation**
Week 2	**n (%)**	97 (89.8)	11 (10.2)	-	42 (40.0)	39 (37.1)	24 (22.9)
Week 6	**n (%)**	95 (87.9)	10 (9.3)	3 (2.8)	50 (47.6)	29 (27.6)	26 (24.8)
Week 12	**n (%)**	99 (91.7)	9 (8.3)	-	63 (60.0)	24 (22.9)	18 (17.1)

Table 22: Summary of shift table results for pain scores (ITT).

Furthermore, the significantly higher number of exacerbations in the physiotherapy group is striking. While among the Vivira patients only N=3 patients (2.8%) reported an exacerbation of pain at T2 (6 weeks), which had disappeared after 12 weeks, the exacerbation situation in the physiotherapy group was significantly different. After 2 weeks, the number of patients with pain exacerbations was N=24 (22.9%), increasing to N=26 (24.8%) after 6 weeks (see figure below). Only after 12 weeks the number of patients with pain increase decreased, but still amounted to N=18 (17.1%).

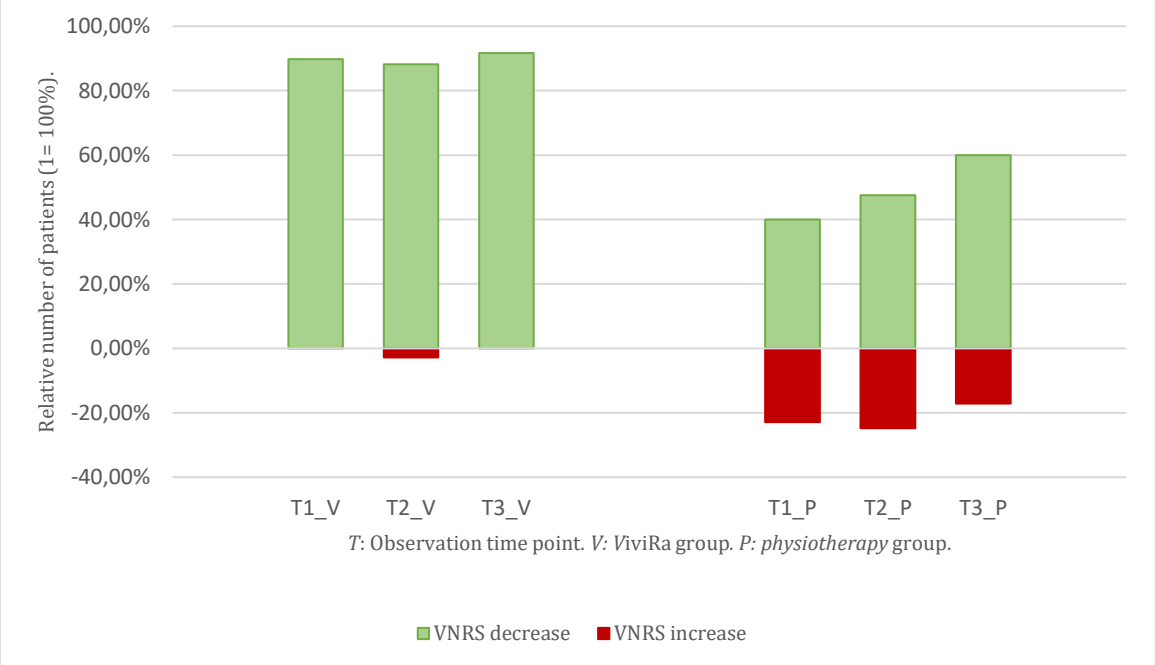


Figure 5: Relative number of ITT patients with improvements or exacerbations of pain by at least one point on VNRS scale (1=100%). Green: decrease, Red: increase.

15.2 Efficacy subgroup analysis

15.2.1 Pain score by anatomic localization of disease

The majority of cases the patients belonged to the subgroup "lumbar spine pain" in both study groups of the ITT population. On average, the relative pain score reductions were of the same magnitude in this subgroup as in the overall group (week 2: -40.76 % (Vivira App) versus -2.00 % (physiotherapy); week 6: -47.98 % (Vivira App) versus -6.94 % (Physiotherapy), week 12: - 54.78 % (Vivira App) versus -16.36 % (physiotherapy). Intragroup significance was <0.001 over 2, 6 and 12 weeks (Vivira) vs. 0.2357, 0.0396 and 0.0004 in the control. Cohen´s effect sizes were 1.57, 1.59, 1.81 (Vivira) vs. 0.19, 0.33, 0.57 (control). Group comparison revealed p= <0.001, <0.001, <0.001 (Vivira vs. control) and Cohen´s d of 1.20, 1.03, 0.99 (Vivira vs. control) confirming superiority of Vivira.

Also for the subgroup "thoracic spine pain" much higher relative mean pain score reductions were observed in the Vivira App group than in the physiotherapy group (week 2: -30.47 % (Vivira App) versus -6.39 % (physiotherapy); week 6: -44.42 % (Vivira App) versus -7.12 % (physiotherapy), week 12: - 53.61 % (Vivira App) versus -13.04 % (physiotherapy). The subgroup cervical spine pain was too small for a valid conclusion. mITT and PP analyses showed similar results as the ITT analysis (data not shown).

a.

Pain score (VRNS) (score points)		**Vivira App (n=105)**				**Physiotherapy (n=105)**			
		Baseline	**Week 2**	**Week 6**	**Week 12**	**Baseline**	**Week 2**	**Week 6**	**Week 12**
Thoracic spine	**n**	13	13	13	13	13	13	13	13
	Mean	6.15	4.23	3.46	2.92	5.38	4.92	5.31	4.69
	SD	1.86	1.96	2.40	2.40	1.26	1.55	1.84	2.02
	Cohen's 1	-	1.01	1.26	1.52	-	0.33	0.05	0.42
	p-value	-	0.0172	0.0041	0.0009	-	0.4139	0.9023	0.3063
	Cohen's 2	-	0.39	0.87	0.80				
	p-value	-	0.3293	0.0385	0.0532				
Cervical spine	**n**	1	1	1	1	9	9	9	9
	Mean	8.00	6.00	4.00	3.00	6.33	6.44	5.78	6.22
	SD	0.00	0.00	0.00	0.00	1.58	1.42	2.11	1.39
	Cohen's 1	-	-	-	-	-	0.07	0.30	0.07
	p-value	-	-	-	-	-	0.8775	0.5367	0.8764
	Cohen's 2	-	0.62	1.69	4.63				
	p-value	-	-	-	-				
Lumbar spine	**n**	88	88	88	88	78	77	79	78
	Mean	6.47	3.81	3.40	3.00	6.05	5.75	5.48	5.03
	SD	1.63	1.75	2.23	2.20	1.64	1.47	1.80	1.91
	Cohen's 1	-	1.57	1.59	1.81	-	0.19	0.33	0.57

	p-value	-	<0.001	<0.001	<0.001	-	0.2357	0.0396	0.0004
	Cohen's 2	-	1.20	1.03	0.99				
	p-value	-	<0.001	<0.001	<0.001				
Missing	**n**	6	6	6	6	5	5	5	5
	Mean	6.00	5.00	5.00	4.33	7.20	5.60	5.40	6.00
	SD	1.55	1.67	1.26	1.37	2.17	0.89	1.67	2.00

b.

Pain score (VNRS) Change (score points)		**Vivira App (n=108)**			**Physiotherapy (n=105)**		
		Week 2	**Week 6**	**Week 12**	**Week 2**	**Week 6**	**Week 12**
Thoracic spine	**n**	13	13	13	13	13	13
	Mean	-1.92	-2.69	-3.23	-0.46	+0.08	-0.69
	SD	1.66	2.02	2.20	1.56	1.93	1.65
Cervical spine	**n**	1	1	1	9	9	9
	Mean	-2.00	-4.00	-5.00	+0.11	-0.56	-0.11
	SD	0.00	0.00	0.00	0.60	2.01	1.05
Lumbar spine	**n**	88	88	88	77	79	78
	Mean	-2.66	-3.07	-3.47	-0.31	-0.56	-1.03
	SD	1.75	2.09	2.03	1.40	1.57	1.54
Missing	**n**	6	6	6	5	5	5
	Mean	-1.00	-1.00	-1.67	-1.60	-1.80	-1.20
	SD	1.10	1.10	1.37	1.82	1.48	0.84

c.

Pain score (VNRS) Change (%)		**Vivira App (n=108)**			**Physiotherapy (n=105)**		
		Week 2	**Week 6**	**Week 12**	**Week 2**	**Week 6**	**Week 12**
Thoracic Spine	**n**	13	13	13	13	13	13
	Mean	-30.47	-44.42	-53.61	-6.39	-7.12	-13.04
	SD	20.28	30.31	30.45	28.21	36.89	32.23
Cervical spine	**n**	1	1	1	9	9	9
	Mean	-25.00	-50.00	-62.50	+2.57	-7.23	-0.42
	SD	0.00	0.00	0.00	9.14	28.568	15.537
Lumbar Spine	**n**	88	88	88	77	79	78
	Mean	-40.46	-47.98	-54.78	-2.00	-6.94	-16.36
	SD	22.82	30.69	29.40	24.30	27.79	25.37
Missing	**n**	6	6	6	5	5	5
	Mean	-16.19	-14.94	-26.07	-16.67	-22.78	-17.22
	SD	15.92	15.92	20.08	25.53	17.61	10.09

Table 23: Pain score (VNRS) from baseline to week 12 by anatomic localization (ITT). a.: absolute values; b.: absolute changes; c.: percent changes. VRNS = Verbal numerical rating scale, n = sample size, SD = Standard deviation.

15.2.2 Pain score by age

In all age groups up to 75 yrs., Vivira app reduces the pain to a higher extent than the physiotherapy. Age groups from 18-75 all produced a highly significant treatment superiority of Vivira over control. The highest age group was to small for a convincing subgroup analysis. Details may be seen in the subsequent table. The results of the mITT and PP analyses did not differ from those of the ITT population (data not shown).

a.

ain score (VNRS) (score points)		Vivira App (n=105)				Physiotherapy (n=105)			
		Base line	Week 2	Week 6	Week 12	Baseline	Week 2	Week 6	Week 12
18-35 yrs	n	11	11	11	11	9	9	9	9
	Mean	5.73	3.45	2.91	2.27	5.78	5.78	5.11	4.56
	SD	1.35	1.13	1.45	1.62	1.48	1.48	1.76	1.33
	Cohen's 1	-	1.83	2.01	2.33	-	0	0.41	0.87
	p-value	-	0.0004	0.0001	<0.0001	-	1.000	0.3985	0.0846
	Cohen's 2	-	1.79	1.37	1.55				
	p-value	-	<0.001	<0.001	<0.001				
36-45 yrs	n	12	12	12	12	10	9	9	9
	Mean	6.25	4.00	3.33	2.75	6.40	5.78	6.00	6.11
	SD	1.76	1.48	2.42	2.36	1.71	1.20	1.80	1.76
	Cohen's 1	-	1.39	1.40	1.70	-	0.43	0.23	0.17
	p-value	-	0.0027	0.0030	0.0004	-	0.3693	0.6275	0.7223
	Cohen's 2	-	1.33	1.27	1.63				
	p-value	-	<0.001	<0.001	<0.001				
46-55 yrs	n	16	17	16	16	18	19	19	20
	Mean	6.31	3.59	2.88	2.56	5.56	5.58	5.37	4.15
	SD	1.40	1.70	1.93	1.59	1.82	1.64	1.92	2.01
	Cohen's 1	-	1.75	2.06	2.51	-	0.01	0.10	0.74
	p-value	-	<0.0001	<0.0001	<0.0001	-	0.0298	0.7628	0.9676
	Cohen's 2	-	1.19	1.29	0.88				
	p-value	-	<0.001	<0.001	0.012				
56-65 yrs	n	33	32	33	33	38	38	38	38
	Mean	6.48	4.16	3.94	3.27	6.29	5.79	5.24	5.26
	SD	1.80	2.02	2.49	2.48	1.66	1.45	1.72	1.80
	Cohen's 1	-	1.21	1.18	1.5	-	0.32	0.62	0.60
	p-value	-	<0.0001	<0.0001	<0.0001	-	0.0116	0.0081	0.1667
	Cohen's 2	-	0.94	0.62	0.93				
	p-value	-	<0.001	0.015	<0.001				
66-75 yrs	n	29	30	3029	30	25	25	25	24
	Mean	6.83	4.07	3.47	3.37	6.12	5.56	5.56	5.46
	SD	1.63	2.00	2.19	2.31	1.54	1.61	1.94	2.11
	Cohen's 1	-	1.52	1.76	1.76	-	0.36	0.32	0.36

	p-value	-	<0.0001	<0.0001	<0.0001	-	0.2173	0.2635	0.2143
	Cohen's 2	-	0.83	1.01	0.95				
	p-value	-	<0.001	<0.001	<0.001				
76-85 yrs	**n**	7	6	6	6	5	5	5	5
	Mean	6.00	4.00	4.33	3.83	5.40	6.20	6.80	5.80
	SD	1.73	1.41	2.16	1.94	1.52	1.30	1.30	1.30
	Cohen's 1	-	1.27	0.86	1.18	-	0.57	0.99	0.28
	p-value	-	0.0427	0.1614	0.0607	-	0.6668	0.1570	0.3978
	Cohen's 2	-	1.62	1.43	1.22				
	p-value	-	0.026	0.047	0.078				

b.

		Vivira App (n=108)			**Physiotherapy (n=105)**		
		Week 2	**Week 6**	**Week 12**	**Week 2**	**Week 6**	**Week 12**
18-35 yrs	**n**	11	11	11	9	9	9
	Mean	-2.27	-2.82	-3.45	0.00	-0.67	-1.22
	SD	1.42	1.94	1.81	2.12	2.35	1.79
36-45 yrs	**n**	12	12	12	9	9	9
	Mean	-2.25	-2.92	-3.50	-0.78	-0.56	-0.44
	SD	1.48	1.98	1.88	0.83	1.33	1.13
46-55 yrs	**n**	17	16	16	19	19	20
	Mean	-2.65	-3.44	-3.75	+0.05	-0.16	-1.30
	SD	1.93	2.39	2.24	1.22	1.57	1.59
56-65 yrs	**n**	32	33	33	38	38	38
	Mean	-2.38	-2.55	-3.21	-0.50	-1.05	-1.03
	SD	1.66	1.86	2.06	1.31	1.45	1.38
66-75 yrs	**n**	30	30	30	25	25	24
	Mean	-2.73	-3.33	-3.43	-0.56	-0.56	-0.75
	SD	1.98	2.22	2.19	1.53	1.53	1.62
76-85 yrs	**n**	6	6	6	5	5	5
	Mean	-2.00	-1.67	-2.17	+0.80	1.40	+0.40
	SD	1.79	1.86	1.72	1.30	1.52	0.89

c.

		Vivira App (n=108)			**Physiotherapy (n=105)**		
		Week 2	**Week 6**	**Week 12**	**Week 2**	**Week 6**	**Week 12**
18-35 yrs	**n**	11	11	11	9	9	9
	Mean	-38.38	-47.06	-60.48	+6.31	-5.67	-17.98
	SD	19.22	27.45	29.87	37.09	41.66	28.80
36-45 yrs	**n**	12	12	12	9	9	9
	Mean	-34.78	-48.79	-58.47	-10.29	-7.98	-6.30
	SD	19.92	29.92	30.07	10.19	20.30	18.15
46-55 yrs	**n**	17	16	16	19	19	20
	Mean	-41.60	-52.55	-57.72	+3.99	-0.14	-23.52
	SD	25.99	36.71	27.80	23.38	31.20	28.15
56-65 yrs	**n**	32	33	33	38	38	38
	Mean	-36.54	-41.56	-51.81	-5.56	-15.22	-15.67
	SD	22.29	29.28	29.75	21.36	21.22	20.42

66-75 yrs	**n** **Mean** **SD**	30 -39.96 25.31	30 -48.69 29.26	30 -50.69 30.09	25 -7.24 23.20	25 -8.20 23.20	24 -12.54 28.93
76-85 yrs	**n** **Mean** **SD**	6 -31.68 121.47	6 -27.12 36.63	6 -35.91 30.86	5 +19.00 1.32.48	5 +31.86 39.23	5 +10.00 22.36

Table 24: *P*ain score (VNRS) from baseline to week 12 by age subgroups (ITT). a.: absolute values; b.: absolute changes; c.: relative changes (%). VNRS = Verbal numerical rating scale, n = sample size, SD = Standard deviation, yrs =years.

15.2.3 Pain score by gender

The mean pain score reductions were much more marked in the Vivira app group compared to physiotherapy both for male and female patients. A significant gender difference between the experimental group and the comparator was not observed. The corresponding PP analysis led to similar results than the ITT analysis (data not shown).

a.

		Vivira App (n=108)				**Physiotherapy (n=105)**			
		Baseline	**Week 2**	**Week 6**	**Week 12**	**Baseline**	**Week 2**	**Week 6**	**Week 12**
male	**n** **Mean** **SD**	57 6.33 1.65	57 3.72 1.73	57 3.28 2.19	57 2.79 2.08	43 5.91 1.54	43 5.51 1.62	43 5.40 1.71	43 4.77 2.17
	Cohen's 1	-	1.54	1.59	1.90	-	0.25	0.31	0.61
	p-value	-	<0.001	<0.001	<0.001	-	0.1802	0.1758	0.0063
	Cohent's 2	-	1.07	1.09	0.93				
	p-value	-	<0.001	<0.001	<0.001				
female	**n** **Mean** **SD**	51 6.51 1.65	51 4.20 1.83	51 3.75 2.23	51 3.37 2.28	62 6.15 1.71	62 5.85 1.37	62 5.52 1.88	62 5.39 1.68
	Cohen's 1	-	1.33	1.42	1.60	-	0.19	0.35	0.45
	p-value	-	<0.001	<0.001	<0.001	-			
	Cohent's 2	-	1.03	0.86	1.02				
	p-value	-	<0.001	<0.001	<0.001				

b.

Change (score points)		**Vivira App (n=108)**			**Physiotherapy (n=105)**		
		Week 2	**Week 6**	**Week 12**	**Week 2**	**Week 6**	**Week 12**
male	**n** **Mean** **SD**	57 -2.61 1.78	57 -3.05 2.19	57 -3.54 2.16	43 -0.40 1.45	43 -0.51 1.39	43 -1.14 1.71
female	**n** **Mean** **SD**	50 -2.31 1.70	50 -2.76 1.95	51 -3.14 1.92	62 -0.29 1.41	62 -0.63 1.81	62 -0.76 1.33

c.

Change (%)		Vivira App (n=108)			Physiotherapy (n=105)		
		Week 2	Week 6	Week 12	Week 2	Week 6	Week 12
male	n Mean SD	57 -40.65 23.51	57 -48.04 31.24	57 -56.17 29.97	43 -4.48 25.55	43 -7.20 22.63	43 -19.52 30.61
female	n Mean SD	51 -35.05 21.96	51 -43.15 30.03	51 -49.70 28.87	62 -1.05 23.35	62 -7.10 31.76	62 -11.22 20.40

Table 25: Pain score by sex (ITT).
a.: absolute values; b.: absolute pain score changes; c.: relative pain score changes (%).
n = sample size, SD = Standard deviation.

Conclusion of pain score analyses

In the Vivira group pain decreased by -2.47 score points (-38.01 %, week 2), -2.92 score points (-45.73 %, week 6) and -3.35 score points (-53.11 %, week 12). In the physiotherapy group this was no more than -0.33 score points (- 2.45 %, week 2), - 0.58 score points (7.14 %, week 6) and -0.91 score points (14.62 %, week 12). Between T0 and T2 (6 weeks) and T0 and T3 (12 weeks) the pain score changes in the Vivira group and in the physiotherapy group were statistically significant (t tests: $p<0.001$/ $p<0.0019$), thus demonstrating efficacy of both Vivira and physiotherapy.

Confirmative group comparison demonstrated that in the Vivira app group the average pain decrease was significantly more pronounced than in the physiotherapy group (week 2: -2.47 score points (-38.01 %, Vivira App) versus -0.33 score points (- 2.45 %, physiotherapy), week 6: -2.92 score points (-45.73 %, Vivira App) versus 0.58 score points (7.14 %, physiotherapy), week 12: -3.35 score points (-53.11 %, Vivira App) versus -0.91 score points (14.62 %, physiotherapy), respectively). At all measurement times group differences were highly significant (t-tests: $p<0.001$/ $p<0.001$/ $p<0.001$) indicating superiority of the Vivira app. This was supported by the sensitivity analysis (Mann-Whitney test: $p<0.0001$/$p<0.0001$/ $p<0.0001$).

With Cohen's d absolute values of 1.26 (relative value=1.31) between T0 and T2 and 1.37 (relative value=1.40) between T0 and T3 the present investigation reveals a large effect size difference between the Vivira group and the physiotherapy group. Confirming superiority of Vivira over TAU these statistically significant effect sizes represent a high clinical effect strength[33].

There are significantly more non-responders in the physiotherapy group than in the Vivira group (n=41 vs. n=9). Furthermore, shift analysis revealed a higher number of pain exacerbations in the physiotherapy group. While among the Vivira patients only N=3 patients

33 Prof. Dr. Weise

(2.8%) reported an exacerbation of pain (after 6 weeks), in the TAU group the number of patients with pain exacerbations was N=24 (22.9%, after 2 weeks), increasing to N=26 (24.8%) after 6 weeks and then decreasing until 12 weeks, but still amounting to N=18 (17.1%). Moreover, shift analysis demonstrates that after 12 weeks N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower (N=63, 60.0%). At the earlier time points T1 (2 weeks) and T2 (6 weeks), the differences were even more pronounced. This notwithstanding after 2 weeks T0/T1 VNRS difference for Vivira was already significant whereas for physiotherapy pain score difference was not yet significant.

Pain score efficacy subgroup analysis by anatomic localization of disease, by age or by gender did not reveal significant subgroup differences.

15.3 Secondary outcomes

15.3.1 Quality-of-life/SF-36

The quality-of-life was measured with the generic SF-36 questionnaire prior and after treatment, i.e. at baseline and after 2, 6 and 12 weeks. Validated outcome variables are the changes in PCS and MCS (PCS: physical component score; MCS: mental component score; Ellert et al. 2005[34]). These are described in the subsequent table and figure. In addition, these two subscales were combined via the Z-score sum according to the proposal of O`Brien (1984)[35] for multiple endpoints and chosen as an overall endpoint.

QoL improved in both groups. A statistically significant overall QoL improvement, however, was observed only in the Vivira group (T2: p=0.0252, T3: p=0.0180) suggesting superiority of Vivira over TAU. Furthermore, the significant treatment difference (p=0.0332) between both groups at T2 supports this notion. When the mental score (MCS) was determined the significant treatment difference (p=0.0186, p=0.0415) between both groups at T2 and T3 supports the superiority of Vivira, although the intraindividual improvements in both groups remained insignificant. When investigating the physical score (PCS) significance of improvement was given only in the Vivira group (T2: p=0.0170; T3: p=0.0184) also suggesting superiority of Vivira over TAU. Furthermore, the trend in the treatment difference between both groups at T2 may support this notion.

[34] Ellert U, Lampert T, Ravens-Sieberer U.: Measuring health-related quality of life with the SF-8. Normal sample of the German population. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2005 Dec;48(12):1330-7. doi: 10.1007/s00103-005-1168-5.

[35] O`Brien, P: Procedures for comparing samples with multiple endpoints. Biometrics (1984) 40, 1079-1087

		T0 Vivira	T1	T2	T3	T0 Physio-therapy	T1	T2	T3
Overall Quality of Life	n	108	103	99	97	105	104	101	101
	Mean	64.27	65.65	69.21	69.72	65.37	63.20	64.50	67.43
	SD	16.05	16.08	15.45	16.58	14.99	15.87	15.61	15.63
	T-Dif	-3.1	-6.8	-9.0	-6.8				
p-value treatment difference		0.6086	0.2723	0.0332	0.3189				
p-value change from baseline		-	0.5357	0.0252	0.0180	-	0.3122	0.6857	0.3348
Physical Score (PCS)	n	108	103	99	97	105	104	101	101
	Mean	61.91	63.57	67.64	68.06	63.71	61.93	63.77	67.59
	SD	17.22	17.59	17.00	19.61	17.73	18.36	17.83	17.66
	T-Dif	1.8	-1.6	-3.9	-0.47				
p-value treatment difference		0.4527	0.5120	0.1180	0.8591				
p-value change from baseline		-	0.4877	0.0170	0.0184	-	0.4775	0.9803	0.1171
Mental Score (MCS)	n	108	103	99	97	105	104	101	101
	Mean	67.97	68.74	71.56	72.19	67.63	65.11	65.58	67.14
	SD	18.49	18.36	18.12	17.35	15.19	16.00	17.46	17.23
	T-Dif	-0.34	-3.6	-6.0	-5.0				
p-value treatment difference		0.8819	0.1310	0.0186	0.0415				
p-value change from baseline		-	0.7634	0.1604	0.0938	-	0.2439	0.3716	0.8312

Table 26: Quality-of-life/SF-36. n = sample size, SD= Standard deviation, T-Dif: LS Means treatment difference

15.3.2 Mobility, Stability, Coordination and Total Functional Scores

Mobility score. At baseline, the average mobility score was approximately similar in both study groups in the ITT population. In both groups mobility score improved. The development was virtually similar in both study groups. In both study groups a statistically significant (p=0.0004/0.0031) mean increase of a similar magnitude was observed after 12 weeks, respectively. After 6 weeks, however, the Vivira group had responded earlier (p=0.0064 (Vivira) vs. p=0.0440 (TAU)) suggesting Vivira´s efficacy. Nevertheless, significant group differences were not observed.

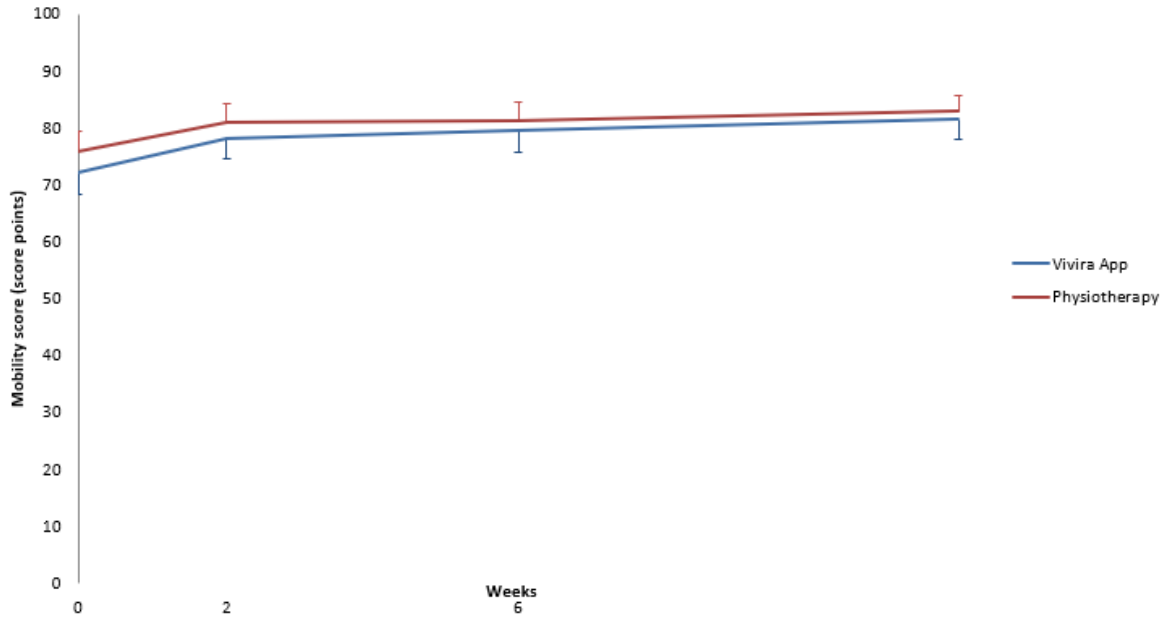
Stability score. The Stability Score (strength score) showed also comparable mean baseline values of a medium category (41-70) between the Vivira app and the physiotherapy group (ITT). In both groups stability (strength) score improved (p<0.0001/p<0.0001, week 12). The development was virtually similar in both study groups. In the Vivira group, the mean stabilty/strength score increased significantly by 22.17 score points from baseline to follow-up; T0-T3). This corresponds to a 35.84 % increase (p<0.0001) suggesting Vivira´s efficacy. Interestingly, this strength increase corresponds in both groups to the frequently observed exercise-induced muscle soreness ("*Muskelkater*") that has been reported by patients for the beginning of therapy. In both groups a similar increase was observed. After the intervention both groups were in a high ability category on average (71-90). Significant group differences were not observed.

Coordination score. The mean baseline values were well balanced regarding the Coordination Score, too. In both study groups only a slight increase from baseline to weeks 6 and 12 was observed, respectively. Data does not demonstrate Vivira´s efficacy. Significant group differences were not observed.

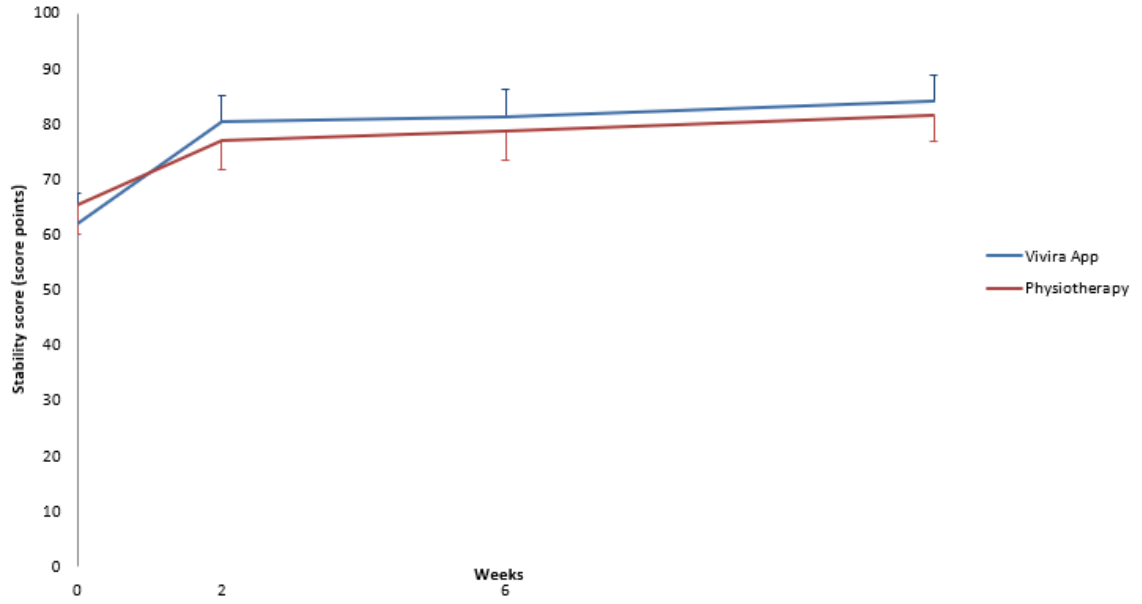
Total functional score. The Total Functional Score showed also comparable mean baseline values. In both groups the score improved statistically significant ($p<0.0001$ Vivira after 12 weeks/$p<0.0001$ TAU after 12 weeks). In the Vivira group the mean total functional score increased significantly from baseline to week 12 by 13.49 (19.18 %) suggesting Vivira´s efficacy. In the Vivira App group and in the Physiotherapy group similar increases were noted on average. Thus, significant group differences were not observed.

Summary. Corresponding displays of the results are presented in the subsequent table and figures. The corresponding data of the PP analysis produced similar results (data not shown). In summary, mobility, stability (strength) and total functional scores show a statistically significant efficacy of Vivira. No improvement of the coordination score was found. The development of all movement scores was virtually similar in both study groups. Thus, in the movement scores superiority of Vivira vs. TAU is not given.

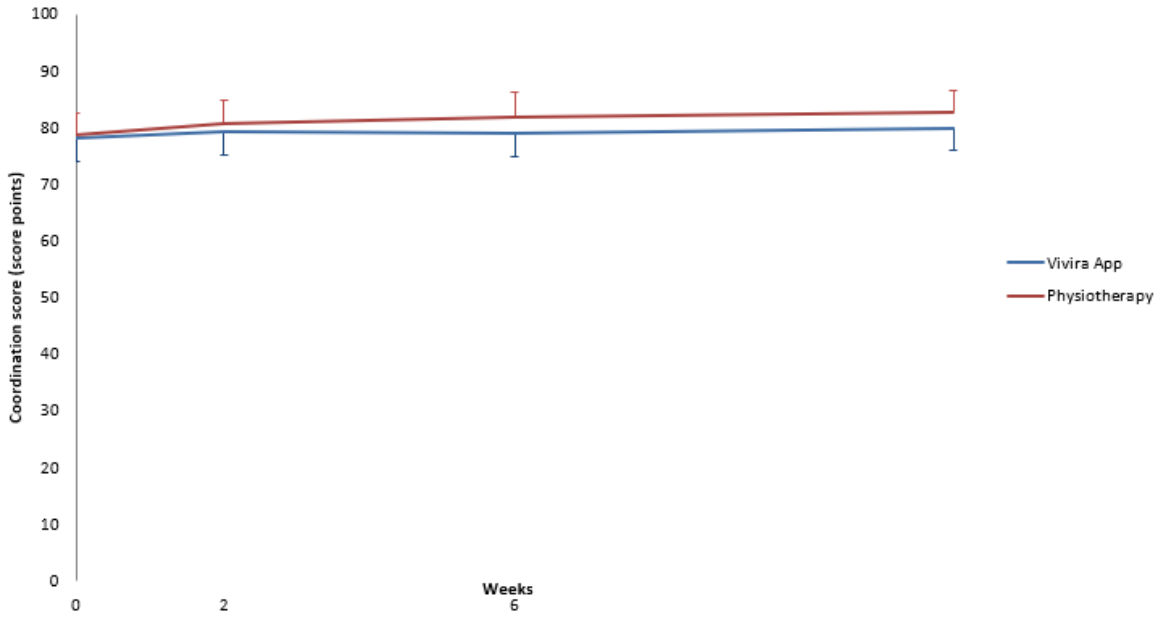
a.


100
90
80
70
60
50
40
30
20
10
0
Mobility score (score points)
0
2
6
Weeks
Vivira App
Physiotherapy

b.


100
90
80
70
60
50
40
30
20
10
0
Stability score (score points)
0
2
6
Weeks
Vivira App
Physiotherapy

c.



d.

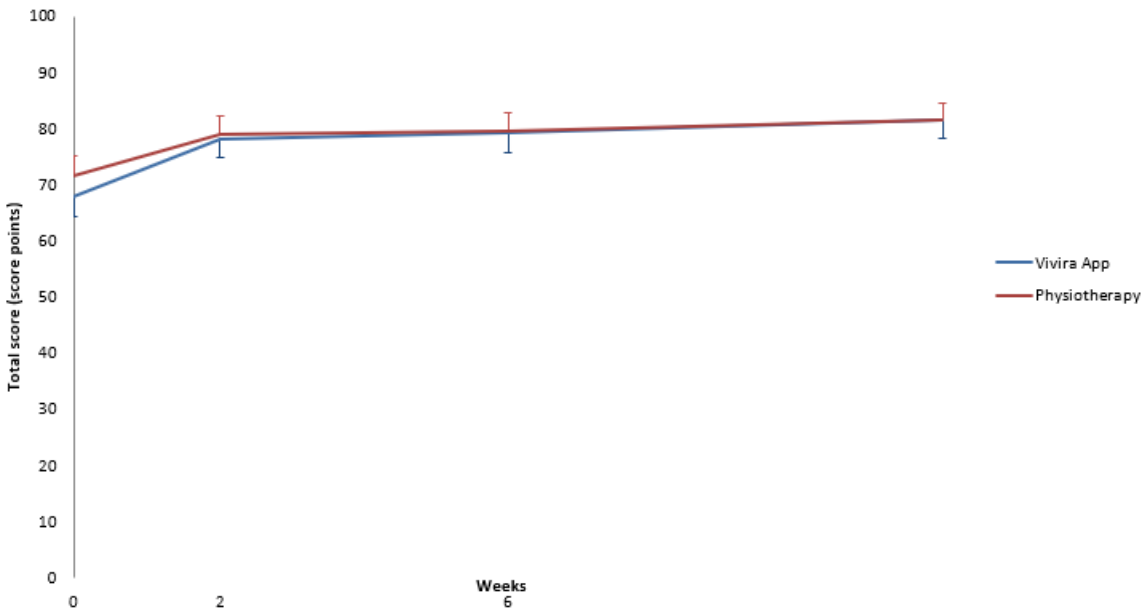


Figure 6: Movement scores over time – a: Mobility; b: Stability/Strength; c: Coordination; d: Total functional score (ITT). Mean + 95 % confidence intervals.

		Vivira App (n=108)				Physiotherapy (n=105)			
		Base-line	Week 2	Week 6	Week 12	Base-line	Week 2	Week 6	Week 12
Mobility score	**n**	108	103	99	97	105	102	101	101
	Mean	72.08	78.50	79.44	81.44	75.95	80.98	81.09	82.82
	SD	19.97	18.19	18.49	17.08	18.44	17.05	17.94	14.33
	Median	75.00	80.00	80.00	80.00	80.00	80.00	80.00	85.00
p-value treatment difference		0.143	0.2607	0.5240	0.5399				
p-value change from baseline		-	0.0208	0.0064	0.0004	-	0.0428	0.0440	0.0031
Stability score	**n**	108	103	99	97	105	102	101	101
	Mean	61.85	80.29	81.21	84.02	65.24	76.86	78.61	81.49
	SD	29.23	25.03	25.33	22.90	26.89	26.02	26.38	23.97
	Median	60.00	100.00	100.00	100.00	60.00	80.00	90.00	90.00
p-value treatment difference		0.3797	0.3376	0.4782	0.4475				
p-value change from baseline		-	<0.0001	<0.0001	<0.0001	-	0.0018	0.0004	<0.0001
Coordination score	**n**	108	103	99	97	105	102	101	101
	Mean	78.15	79.13	78.99	79.90	78.76	80.78	81.78	82.67
	SD	21.58	20.49	21.50	19.92	19.15	20.86	22.33	18.70
	Median	80.00	80.00	80.00	80.00	80.00	80.00	80.00	80.00
p-value treatment difference		0.8263	0.5665	0.3688	0.3137				
p-value change from baseline		-	0.7359	0.7791	0.5470	-	0.4686	0.2995	0.1396
Total functional score	**n**	108	103	99	97	105	102	100	101
	Mean	67.99	78.11	81.48	81.48	71.66	78.81	79.46	81.55
	SD	19.43	16.97	15.48	15.48	17.55	16.94	16.97	14.66
	Median	67.65	82.50	85.80	85.50	72.60	79.20	82.50	82.50
p-value treatment difference		0.1491	0.7675	0.9258	0.9728				
p-value change from baseline		-	<0.0001	<0.0001	<0.0001	-	0.0032	0.0014	<0.0001

Table 27: Mobility, stability (strength), coordination and total functional scores (ITT). SD= Standard deviation.

Conclusion of movement analyses

In both groups mobility, stability/strength and Total Functional Score improved statistically significantly. Coordination Score in both study groups showed only an insignificant increase from baseline to weeks 6 and 12. Group comparison analysis revealed similar results for both study groups. Thus, efficacy of Vivira can be shown, but no superiority.

15.4 Transient concomitant pain medication

At T0, analgetic drug use was an exclusion criterion. Thus, at T0 all patients started without analgetics. Should the necessity of pain medication arise during the study, however, this would be allowed. The resulting concomitant drug use was determined using the consented Medication Qualification Scale (MQS), Version III (Harden et al. 2005)[36]. The MQS was developed as a consented tool to co-quantify relevant aspects of medications prescribed for chronic nonmalignant pain. The underlying ratings of detriment weight were relatively consistent (α = .84). The next table displays the drugs used and their detriment weights:

Steroid	4,4
Benzodiazepine	3,9
Muscle relaxant	3,8
Opioid	3,7
Ibuprofen	3,4
Novalminsulfone, acetylacetic acid	2,3
Tricyclic antidepressant	2,3
Anxiolytics	2,1
Antidepressant serotonin reutake inhibitor	1,7
Ointment	1,1

Table 28: Concomitant drugs and detriment weights. Acc. to Harden et al. 2005

Relative dosage level is taken into according the subsequent dosage score table:

SCORE	
1	Subtherapeutic dose or occasional use
2	Lower 50 % of the therapeutic dose range
3	Upper 50 % of the therapeutic dose range
4	Supratherapeutic dose

Table 29: Dosage Level/Relative Dosage Scores (Harden et al. 2005).

The following figure and table show the MQS over time from T0-T3 (baseline to week 12) divided into the three time phases T0-T1 (weeks 1-2), T1-T2 (weeks 3-6), and T2-T3 (weeks 7-12). There is a statistically significant non-linear decrease in analgesic consumption in the Vivira group by – 1.79/-67% (SD=2.24), p=<.0001) between weeks 1 and 12. In the control group analgesic consumption is also reduced, which is -0.57/-27% (SD=.2.07, 0.1268) but statistically not significant. The group difference is significant (p=0.0210). Thus, Vivira is not only effective in reducing analgesic consumption, but superior to the control physiotherapy.

36 Harden, RN, Weinland SR, Remble TA, Houle TT, Colio S, Steedman S and Kee WG. Medication Quantification Version III. J. Pain 6 (2005) 364-371.

Interestingly, in the control group medication did not change between weeks 1 and 6. This may be explained by the delayed initiation of physiotherapy as a consequence of TAU, which typically started after week 6, then resulting in the observed medication reduction.

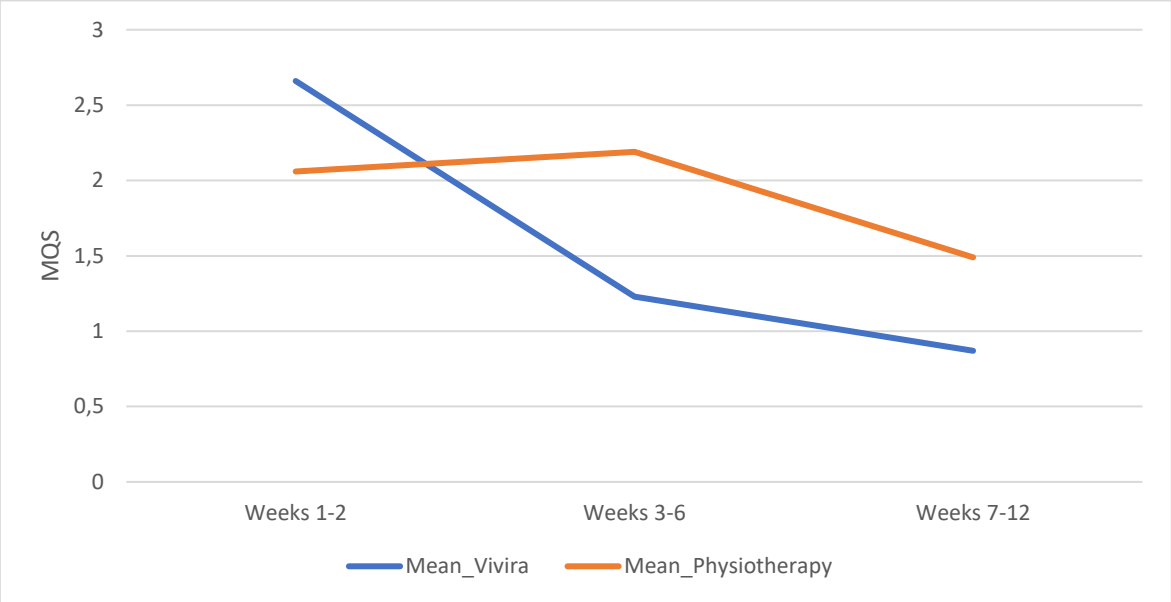


Figure 7: Drug score over time. MQS means over time from T0-T3 (baseline to week 12) divided into the 3 time phases T0-T1 (weeks 1-2), T1-T2 (weeks 3-6), and T2-T3 (weeks 7-12). SDs are indicated in the subsequent table. MQS was computed according to Harden et al. 2005.

MQS		Vivira App (n=34)			Physiotherapy (n=36)		
		Weeks 1-2	**Weeks 3-6**	**Weeks 7-12**	**Weeks 1-2**	**Weeks 3-6**	**Weeks 7-12**
Total	**Mean SD**	**2.66 1.40**	**1.23 1.63**	**0.87 1.40**	**2.06 1.54**	**2.19 1.53**	**1.49 1.59**
Changes from Periode of Weeks 0-2	**Mean SD**	-	**-1.43 2.42**	**-1.79 2.24**	-	**0.13 2.25**	**-0.57 2.07**
	p-value treatment difference		0.0070	0.0210		-	
	p-value change from T1		0.0003	<.0001		0.7302	0.1268

Table 30: Drug score over time. MQS over time T0-T3 (baseline to week 12) divided into the 3 time phases T0-T1 (weeks 1-2), T1-T2 (weeks 3-6), and T2-T3 (weeks 7-12). MQS was computed according to Harden et al. 2005.

16.STATISTICAL METHODS/ANALYTICS

16.1 Patient disposition and baseline characteristics

Demography and other baseline characteristics were summarized descriptively by treatment group. For quantitative data, means, medians, standard deviations and extremes were determined. For qualitative data, categorical variables were summarized using the cell frequencies and percentage of patients in each demographic category.

16.2 Efficacy

For the confirmatory analysis presented here, a testing strategy was followed that controls for study-wide error of the 1st kind. Continuous variables were summarized using the mean, standard deviation, median, minimum, maximum and 95 % confidence interval values. Categorical variables were summarized using the cell frequencies and percentage of patients in each category. Figures show the time course of outcomes by treatment group for key quantitative outcomes.

16.2.1 Primary efficacy variable

The primary efficacy variable pain score (VNRS) was summarized using the mean, standard deviation, median, minimum, maximum and 95 % confidence interval values. Score differences and Cohen's delta were calculated for confirmatory treatment group comparisons (Vivira vs. TAU/physiotherapy) as well as for the intragroup score changes from baseline. Confirmatory treatment group differences were tested using the t-test. Subgroup analyses were performed by localization of symptoms, age group, sex, non-responders and patients showing an exacerbation of the pain score. Non-responders are defined as patients who did not achieve any form of improvement in the pain score. Patients showing an exacerbation are defined as showing a worsening of the pain score. Additionally, shift tables for the pain score were produced. Odds ratios were omitted because they did not fit the parameter. For the 10-point scale, the non-parametric Wilcoxon rank-sum test was calculated instead.

16.2.2 Secondary outcomes

Mobility, Stability, Total Function, Quality of Life/SF-36. For confirmatory group comparisons continuous variables were summarized using the mean, standard deviation, median, minimum, maximum and 95 % confidence interval values. Categorical variables were summarized using the cell frequencies and percentage of patients in each category.

16.3 Safety

Adverse effects were listed.

16.4 Missing data

Missing data for the primary efficacy variable was replaced using the Last-observation-carried-forward (LOCF) method. Further missing data was not replaced.

Tabulation of individual response data

For data protection reasons, all data is displayed as metadata in anonymized form. In principle, an individual data presentation could be used to draw conclusions about the individual: reliable anonymization may not be possible.

Study treatment and other therapy interactions

The influence of both Vivira treatment and comparator on drug consumption was investigated. See the relevant result chapter on drug consumption.

By-patient displays

For data protection reasons, all data is displayed as metadata in anonymized form. In principle, an individual data presentation could be used to draw conclusions about the individual: reliable anonymization may not be possible.

17. EFFICACY CONCLUSIONS

Using a German language validated verbal numerical rating scale (VNRS) as primary efficacy variable for pain both groups of the ITT population produced a mean pain decrease from baseline after 2, 6 and 12 weeks. In the Vivira group pain decreased by -2.47 score points (-38.01 %, week 2), -2.92 score points (-45.73 %, week 6) and -3.35 score points (-53.11 %, week 12). In the physiotherapy group this was no more than -0.33 score points (- 2.45 %, week 2), - 0.58 score points (7.14 %, week 6) and -0.91 score points (14.62 %, week 12). Pain score changes in both groups the Vivira group and the physiotherapy group were statistically significant. T-tests at T1, T2 and T3 all revealed $p < 0.001$ for the intragroup changes in the Vivira group. For the physiotherapy intragroup changes produced significant p-values at T2 and T3 of 0.015 and 0.0003 demonstrating efficacy of both Vivira and physiotherapy.

Confirmative group comparison demonstrated that in the Vivira app group the average pain decrease was significantly more pronounced than in the physiotherapy group (week 2: -2.47 score points (-38.01 %, Vivira App) versus -0.33 score points (- 2.45 %, physiotherapy), week 6: -2.92 score points (-45.73 %, Vivira App) versus 0.58 score points (7.14 %, physiotherapy), week 12: -3.35 score points (-53.11 %, Vivira App) versus -0.91 score points (14.62 %, physiotherapy), respectively). At all measurement times group differences were highly significant (t-tests: $p<0.001$/ $p<0.001$/ $p<0.001$) indicating superiority of the Vivira app. This was supported by the sensitivity analysis (Mann-Whitney test: $p<0.0001$/$p<0.0001$/ $p<0.0001$)

For a quantitative and metric-free estimation of the effect size Cohen's d was used. With Cohen's d absolute values of 1.26 (relative value=1.31) between T0 and T2 and 1.37 (relative value=1.40) between T0 and T3 the present investigation reveals a large effect size difference between the Vivira group and the physiotherapy group. Confirming superiority of Vivira over TAU these statistically significant effect sizes represent a high clinical effect strength[37].

Similar results were obtained for the mITT and the PP populations.

There were significantly more non-responders (e.g. after 12 weeks: n=42 vs. n=9) and exacerbations in the physiotherapy group than in the Vivira group. Shift analysis revealed that the higher number of non-responders together with a higher number of pain exacerbations in the physiotherapy group may contribute to explain the statistically significant superiority of Vivira. While among the Vivira patients only N=3 patients (2.8%) reported an exacerbation of pain (after 6 weeks), in the TAU group the number of patients with pain exacerbations was N=24 (22.9%, after 2 weeks), increasing to N=26 (24.8%) after 6 weeks and then decreasing until 12 weeks, but still amounting to N=18 (17.1%).

Furthermore, shift analysis demonstrates that after 12 weeks N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower (N=63, 60.0%). At the earlier time points T1 (2 weeks) and T2 (6 weeks), the differences were even more pronounced. This notwithstanding after 2 weeks T0/T1 VNRS difference for Vivira was already significant whereas for physiotherapy pain score difference was not yet significant. Data suggest that structural and procedural influences may play a significant role in the early result differences between the groups: in the control group, the start of therapy was significantly delayed by waiting times until the first therapy session with the physiotherapist. No earlier than after time point T2 (6 weeks) all physiotherapy patients had also started their therapy. Thus, data also suggests patient related process improvements being produced by the app that are superior to TAU. Superiority is characterized by a significantly earlier onset of Vivira therapy when compared with TAU.

Pain score efficacy subgroup analysis by anatomic localization of disease, by age or by gender did not reveal significant subgroup differences.

[37] Prof. Dr. Weise

Using the SF-36 for the determination of QoL this improved in both groups. A statistically significant overall QoL improvement, however, was observed only in the Vivira group (T2: p=0.0252, T3: p=0.0180) suggesting superiority of Vivira over the control. When the mental score (MCS) was determined the significant treatment difference (T2: p=0.0186) between both groups at T2 supports the superiority of Vivira. When investigating the physical score (PCS) significance of improvement was given only in the Vivira group (T2: p=0.0170; T3: p=0.0184) also suggesting superiority of Vivira over the control.

Movement analyses revealed that mobility, stability (strength) and the total functional score improved significantly in both groups (T3: p<0.0001). Treatment differences between both groups (Vivira vs. physiotherapy) were not statistically significant. Coordination Score in both study groups showed only a statistically insignificant increase from baseline to weeks 6 and 12. Group comparison movement analysis revealed similar results for both study groups i.e. data shows a statistically significant efficacy but no superiority of Vivira.

Using the consented Medication Qualification Scale (MQS) there was a statistically significant non-linear decrease in analgesic consumption in the Vivira group by – 1.79/-67% (SD=2.24, p=<.0001) between weeks 1 and 12. In the control group analgesic consumption is also reduced, which is -0.57/-27% (SD: 2.07, p=0.1268) but statistically not significant. The group difference is significant (p=0.0210). Thus, Vivira is not only effective in reducing analgesic consumption, but superior to control physiotherapy. This drug data may be strongly influenced by the delayed start of physiotherapy therapy, whereas Vivira patients were able to start their exercises without considerable time delay. Thus, this data also suggests further patient related process improvements being produced by the app that are superior to TAU.

18. SAFETY EVALUATION

Extent of Exposure

From Nov 16, 2020 to March 31, 2021, N=213 ITT patients were exposed to exercises. 199/213 patients adhered to the study over 12 weeks.

Burden

An active surveillance of burden was performed. Vivira produced the following burden:

Burden	N
Muskelkater / Sore muscles	10
starkes Schwitzen / Sweating	1

Table 31: Burdens reported by ITT Vivira Users

For physiotherapy, the following burden was described:

Burden	N
Muskelkater / Sore muscles	9
Übungen anstrengend / Exhausting	1
Seitenstechen / Stitch	1

Table 32: Burdens reported by ITT patients by physiotherapy.

Adverse Events (AEs)

An active surveillance of AEs was performed. For a total of 68 (31.8 %) patients adverse effects including burden, AEs and ARs were described. In the Vivira App group 37 patients reported adverse effects (34.3 %) compared to 31 patients in the Physiotherapy group (29.3 %). The details of the adverse effects are given in the next table. Also, the verbatim terms in German of the adverse effects are presented in the table.

Of the total 37 patients who experienced adverse events during the use of the Vivira app, 28 (75.6%) experienced side effects that were caused by the use of the Vivira app (ARs). The following list tabulates the ARs produced by the app. All ARs were transient:

- Transient cervical blockage
- Transient thigh/calve clamp
- Isolated, transient muscle clamp
- Transient leg stiffness
- Transiently impaired hip movement
- Transient cervical tension
- Transient dizziness
- Transient nausea
- Transient strength loss
- Transient numb toes
- Transient movement decrease
- Transient pain/pain increase

Side Effect	AE	AR All transient
Schmerzen bei Übungen	19	19
Schmerzverstärkung, bei manchen Übungen	3	3
Krämpfe, Oberschenkel, Waden	3	3
Schwindel, bei manchen Übungen, bei Vierfüsslerstand	2	2
Blockade HWS, Kopf ließ sich nicht mehr drehen	2	2
Bewegung zeitweilig sehr schmerhaft, App hat Schmerzen gelindert	1	0
Übelkeit	1	1
Leichte Hüftbewegungseinschränkun nach einer Übung	1	1
Hexenschuss	1	0
Verspannungen HWS	1	1
Treppensturz	1	0
Lumbago aufgrund täglichen Tragen eines Tieres	1	0
Beinversteifung	1	1
Kraftverlust	1	1
Eingeschlafene Zehen	1	1
Nierensteine	1	0
Pancreaskarzinom	1	0

Table 33: AEs and transient ARs reported by Vivira Users (ITT patients).

Side Effect	AE	AR All transient
Schmerzen, nach Physiotherapie, bei manchen Übungen, Übungen z.T. nicht möglich	10	10
Schmerzverstärkung, manche Übungen verstärken Schmerz	9	8
Blockade, HWS, Genickschmerzen, Verspannung im Rücken	4	3
Schwindel, Ohrgeräusche, nach Massage, gelegentlich	2	2
Kopfschmerzen, nach Behandlung	2	2
verstärkte Verspannungen	1	1
Migräne	1	0
Treppensturz	1	0
Knieprobleme	1	1
Schmerzen in der Hüfte	1	1
Herzarrythmie	1	0

Table 34: AEs and transient ARs reported by control group (ITT patients).

Further comments

Active surveillance also included the search for serious adverse reactions (SARs[38]. Specifically, these are adverse events that

1. lead to death,
2. lead to serious deterioration in the health of the subject, that either resulted in
    a) a life-threatening illness or injury, or
    b) a permanent impairment of a body structure or a body function, or
    c) in-patient or prolonged hospitalization, or
    d) medical or surgical intervention to prevent life-threatening illness or injury or permanent impairment to a body structure or a body function,
3. lead to foetal distress, foetal death or a congenital abnormality or birth defect.

As a result, no SAR was observed.

Serious adverse reactions (SARs)
None

Table 35: Tabulation of SARs.

Safety Conclusions

For safety issues an active surveillance of burden and of AEs/ARs including SAEs/SARs was performed allowing the subsequent burden analysis and risk/benefit analysis:

Burden analysis

Vivira produced the following burden each of which is fully acceptable in relation to the benefits. Furthermore, burden produced by Vivira is similar to that by physiotherapy.

Burden	Commentary	Conclusion
Sore muscles	Transient	This burden is fully acceptable in relation to the benefits.
Sweating	Transient	This burden is fully acceptable in relation to the benefits.

Table 36: Burden of Vivira

[38] As defined by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, document E2A (available at www.ich.org)

AR analysis

Using active surveillance, the present study revealed the following transient ARs:

- Transient cervical blockage
- Transient thigh/calve clamp
- Isolated, transient muscle clamp
- Transient leg stiffness
- Transiently impaired hip movement
- Transient cervical tension
- Transient dizziness
- Transient nausea
- Transient strength loss
- Transient numb toes
- Transient movement decrease
- Transient pain/pain increase

No SAR was observed. All ARs are acceptable in relation to the benefits (see next chapter).

Risk/benefit analysis

Any undesirable side effect must constitute an acceptable risk when weighed against the performances intended. The benefits and risks of a medical device can only be assessed in interrelation by "weighing any benefit to health from the use of the device against any probable risk of injury or illness from such use"[39]. Hence, they are to be understood as relative terms: a balanced consideration based on valid scientific evidence in making risk and benefit determinations, including the critical issue of identifying benefits and residual risks is essential (equipoise). The subsequent equipoise is based on the present AR data addressing Vivira and physiotherapy ARs. Furthermore, equipoise includes surgery as a further therapeutical alternative. Incidence of complications after spinal surgery was reported to be 16.4%[40] and as high as 80% for complex procedures[41] and include SARs and mortality.
To allow equipoise the benefits and residual risks of the Vivira app are summarized in the following table:

Evidence based benefits	**ARs**	**ARs of conventional physiotherapy or surgery**
• Reduction of pain and/or the • Improvement of movement capabilities • Improvement of QoL • Home care setting	• Transient cervical blockage • Transient thigh/calve clamp • Isolated, transient muscle clamp • Transient leg stiffness • Transiently impaired hip movement • Transient cervical tension • Transient dizziness • Transient nausea • Transient strength loss • Transient numb toes • Transient movement decrease • Transient pain/pain increase	**Physiotherapy** ARs are similar to those of the app

[39] Guidance for Industry and Food and Drug Administration Staff, Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications, 2012.
[40] Nasser R, Yadla S, Maltenfort MG, et al. Complications in spine surgery. J Neurosurg Spine 2010;13(02):144–157
[41] Yadla S, Maltenfort MG, Ratliff JK, Harrop JS. Adult scoliosis surgery outcomes: a systematic review. Neurosurg Focus 2010;28(03):E3
Carreon LY, Puno RM, Dimar JR II, Glassman SD, Johnson JR. Perioperative complications of posterior lumbar decompressionand arthrodesis in older adults. J Bone Joint Surg Am 2003;85(11):2089–2092.

	SARs	SARs of conventional physiotherapy or surgery
	None	**Physiotherapy** None **Surgery** ARs may be more severe than those of the app and frequent and include SARs and mortality

Table 37: Overview of benefits and residual risks of Vivira in comparison to physiotherapy and surgery.

Safety data resulting from active surveillance of the present study clearly indicate with high evidence strength that the inherent rate and severity of Vivira complications is very low. No SAR was observed. Furthermore, complication rate and AR quality equal those of conventional physiotherapy. Moreover, complication rate and AR quality are significantly lower than those of surgery. Specifically, in contrast to surgery no SARs has been observed.

Thus, based on the findings from literature, clinical data as well as risk analysis it can be inferred that the probability of a patient experiencing a substantial benefit when using the Vivira App significantly outweighs the probability of suffering harm due to a risk of the device.

19. DISCUSSION AND OVERALL CONCLUSIONS

N=213 ITT-patients were analyzed of which N=108 were randomly allocated to the experimental group and N=105 to the control group. From the 213 ITT patients N=199 remained as PP population, which corresponded to the a priori case number calculation of N=198. Both groups revealed to be balanced by age and by pain localization. Slightly more males than females were enrolled in the Vivira App group. The majority of cases suffered from lumbar spine pain in both groups (Vivira App: lumbago/lumboichalgia 40.74 %/40.74 %, physiotherapy: 42.86 %/ 31.43 %), followed by thoracic spine pain and its variants in the ITT population (Vivira App: 18.53 %, physiotherapy: 25.71 %). Demography and clinical baseline findings for the PP population did not differ from those of the ITT population.

Vivira PP patients (N=98) produced an adherence of 82,4% (mean 5.77 days of use per week vs. 7 optional days per week). In the control group all PP patients (N=101) received physiotherapy with an adherence to the prescribed 714 sessions of 97.33% (695 sessions).

Using a German language validated verbal numerical rating scale (VNRS) as primary efficacy variable for pain both groups of the ITT population produced a mean pain decrease from baseline after 2, 6 and 12 weeks. In the Vivira group pain decreased by -2.47 score points

(-38.01 %, week 2), -2.92 score points (-45.73 %, week 6) and -3.35 score points (-53.11 %, week 12). In the physiotherapy group this was no more than -0.33 score points (- 2.45 %, week 2), - 0.58 score points (7.14 %, week 6) and -0.91 score points (14.62 %, week 12). Pain score changes in both groups the Vivira group and the physiotherapy group were statistically significant. T-tests at T1, T2 and T3 all revealed $p < 0.001$ for the intragroup changes in the Vivira group. For the physiotherapy intragroup changes produced significant p-values at T2 and T3 of 0.015 and 0.0003.

Confirmative group comparison demonstrated that in the Vivira app group the average pain decrease was significantly more pronounced than in the physiotherapy group (week 2: -2.47 score points (-38.01 %, Vivira App) versus -0.33 score points (- 2.45 %, physiotherapy), week 6: -2.92 score points (-45.73 %, Vivira App) versus 0.58 score points (7.14 %, physiotherapy), week 12: -3.35 score points (-53.11 %, Vivira App) versus -0.91 score points (14.62 %, physiotherapy), respectively). At all measurement times group differences were highly significant (t-tests: $p<0.001$/ $p<0.001$/ $p<0.001$) indicating superiority of the Vivira app. This was supported by the sensitivity analysis (Mann-Whitney test: $p<0.0001$/$p<0.0001$/ $p<0.0001$)

For a quantitative and metric-free estimation of the effect size Cohen's d was used. With Cohen's d absolute values of 1.26 (relative value=1.31) between T0 and T2 and 1.37 (relative value=1.40) between T0 and T3 the present investigation reveals a large effect size difference between the Vivira group and the physiotherapy group. Confirming superiority of Vivira over TAU these statistically significant effect sizes represent a high clinical effect strength[42].

Similar results were obtained for the mITT and the PP populations.

There were significantly more non-responders (e.g. after 12 weeks: n=42 vs. n=9) and exacerbations in the physiotherapy group than in the Vivira group. While among the Vivira patients only N=3 patients (2.8%) reported an exacerbation of pain (after 6 weeks), in the TAU group the number of patients with pain exacerbations was N=24 (22.9%, after 2 weeks), increasing to N=26 (24.8%) after 6 weeks and then decreasing until 12 weeks, but still amounting to N=18 (17.1%). Furthermore, shift analysis demonstrates that after 12 weeks N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower (N=63, 60.0%).

Using the SF-36 for the determination of QoL this improved in both groups. A statistically significant overall QoL improvement, however, was observed only in the Vivira group (T2: p=0.0252, T3: p=0.0180) suggesting superiority of Vivira over the control. When the mental score (MCS) was determined the significant treatment difference (T2: p=0.0186) between both groups at T2 supports the superiority of Vivira. When investigating the physical score (PCS) significance of improvement was given only in the Vivira group (T2: p=0.0170; T3: p=0.0184) also suggesting superiority of Vivira over the control.

[42] Prof. Dr. Weise

There were significantly more non-responders (e.g. after 12 weeks: n=42 vs. n=9) and exacerbations in the physiotherapy group than in the Vivira group. Shift analysis revealed that the higher number of non-responders together with a higher number of pain exacerbations in the physiotherapy group may contribute to explain the statistically significant superiority of Vivira. While among the Vivira patients only N=3 patients (2.8%) reported an exacerbation of pain (after 6 weeks), in the TAU group the number of patients with pain exacerbations was N=24 (22.9%, after 2 weeks), increasing to N=26 (24.8%) after 6 weeks and then decreasing until 12 weeks, but still amounting to N=18 (17.1%).

Furthermore, shift analysis demonstrates that after 12 weeks N=99 (91.7%) in the Vivira group reported an improvement in pain of at least 1 score point. In the control group, the number was significantly lower (N=63, 60.0%). At the earlier time points T1 (2 weeks) and T2 (6 weeks), the differences were even more pronounced. This notwithstanding after 2 weeks T0/T1 VNRS difference for Vivira was already significant whereas for physiotherapy pain score difference was not yet significant. Data suggest that structural and procedural influences may play a significant role in the early result differences between the groups: in the control group, the start of therapy was significantly delayed by waiting times until the first therapy session with the physiotherapist. No earlier than after time point T2 (6 weeks) all physiotherapy patients had also started their therapy. Thus, data also suggests patient related process improvements being produced by the app that are superior to TAU. Superiority is characterized by a significantly earlier onset of Vivira therapy when compared with TAU.

Pain score efficacy subgroup analysis by anatomic localization of disease, by age or by gender did not reveal significant subgroup differences.

Movement analyses revealed that mobility, stability (strength) and the total functional score improved significantly in both groups (T3: $p<0.0001$). Treatment differences between both groups (Vivira vs. physiotherapy) were not statistically significant. Coordination Score in both study groups showed only a statistically insignificant increase from baseline to weeks 6 and 12. Group comparison movement analysis revealed similar results for both study groups i.e. data shows a statistically significant efficacy but no superiority of Vivira.

Using the consented Medication Qualification Scale (MQS) there was a statistically significant non-linear decrease in analgesic consumption in the Vivira group by – 1.79/-67% (SD=2.24, p=<.0001) between weeks 1 and 12. In the control group analgesic consumption is also reduced, which is -0.57/-27% (SD: 2.07, p=0.1268) but statistically not significant. The group difference is significant (p=0.0210). Thus, Vivira is not only effective in reducing analgesic consumption, but superior to control physiotherapy. This drug data may be strongly influenced by the delayed start of physiotherapy therapy, whereas Vivira patients were able to start their exercises without considerable time delay. Thus, this data also suggests further patient related process improvements being produced by the app that are superior to TAU.

Safety data resulting from active surveillance clearly indicate with high evidence strength that the inherent rate and severity of Vivira complications is very low. Vivira produced burden each of which is fully acceptable in relation to the benefits. Furthermore, burden produced by Vivira is lower than that by physiotherapy. Moreover, complication rate and AR quality equal those of conventional physiotherapy. Finally, complication rate and AR quality are significantly lower than those of surgery. Specifically, in contrast to surgery no SARs have been observed.

Thus, based on the findings it can be inferred that the probability of a patient experiencing a substantial benefit when using the Vivira App significantly outweighs the probability of suffering harm due to a risk of the device.

Conclusion. As a conclusion pain score analyses showed

- A statistically significant reduction of pain over time through the use of the App.
- A statistically significant superiority in pain reduction of the app use in comparison to TAU.
- According to Cohen Vivira achieved a large effect size superior to that of TAU
- The effect size represents a high clinical effect strength superior to that of TAU

Furthermore, statistically significant QoL improvements were observed suggesting superiority of Vivira over the control. Group comparison movement analyses revealed similar results for both study groups i.e. it shows a statistically significant efficacy but no superiority of Vivira. Moreover, Vivira was effective in reducing transient analgesic consumption being superior to the control physiotherapy. Safety findings reveal that the probability of a patient experiencing a substantial benefit when using the Vivira App significantly outweighs the probability of suffering harm due to a risk of the device.

20. REFERENCES

See footnotes.

21. APPENDIX

List and description of investigators

Dipl.-Stat. Michael Bulitta	CRM Biometrics GmbH	Statistics and statistical report
Professor K. Weise, M.D., Dr. med. Medical specialist in Orthopedic Surgery Professor of Orthopedic Surgery & Traumatology Dr. Dr. H. Weise, M.D., M.DD. Medical Specialist in Maxillofacial Surgery	Institute of Orthopedic Surgery and Traumatology Evaluation, Tübingen, Germany.	Study report
Professor H.P. Zenner, M.D., Dr. med. Medical specialist in ORL, Head&Neck Surgery Professor of Head&Neck Surgery	Clinical Research Organization Prof. Dr. H.P. Zenner Clinical GmbH	Study report

Curriculum Vitae Dipl.-Stat. Michael Bulitta

Academic Education. Dipl.-Stat., University of Dortmund, Germany
Professional Career. Biostatistician of Boehringer Mannheim GmbH, Mannheim (1984-1989), Head of Biometrics of Madaus AG, Cologne (1989-2000), Vice Head of Medical Research of Madaus AG, Cologne (1996-2000), CEO of CRM Biometrics GmbH, Rheinbach (since 2000), CEO of Topcro GmbH, Rheinbach (since 2014), Head of Biometrics and Data Management of Venn Life Sciences GmbH, Germany (2018-2020), Freelance Senior Biostatistician for Novartis Vaccines, Marburg (2013-2015), GSK, Marburg (2015), Parexel Early Phase, Berlin (2015-2017).
Clinical Study Statistics. More than 500
Publications. 23 Medline listed publications

Curriculum Vitae Prof. Dr. med. Kuno Weise

Clinical Education. M.D., Dr. med., University of Tübingen, Germany.
Clinical Career. Residency General Surgery followed by a residency in traumatology. 1977 Board Certified General Surgeon, 1982 Board Certified Orthopedic Surgeon, 1982 Consultant (Oberarzt) Orthopedic Surgery BG Unfallklinik Tübingen. 1990 Vice Chairman BG Unfallklinik Tübingen, 1993 Chair Orthopedic Surgery and Traumatology, University of Leipzig, 1996 Chair Orthopedic Surgery and Traumatology, University of Tübingen and Chairman BG Unfallklinik Tübingen, 2010 -today Director Institute of Orthopedic Surgery and Traumatology Evaluation, Tübingen, Germany.
Honorary Offices. 2007 President German Association of Orthopedic Surgery & Traumatology.
Evaluation. Examiner Clinical Trials biocontact prostheses, Director and Examiner Clinical Trials chondrocyte transfer, 2010 – today Director Institute of Orthopedic Surgery and Traumatology Evaluation, Tübingen, Germany.
Publications. More than 100 Medline listed publications

Curriculum Vitae Dr. Dr. Hannes Weise

Clinical Education. M.D., Salzburg University, Austria; D.M.D. Tübingen University, Germany; Dr. med., and Dr.med.dent., University of Tübingen, Germany
Clinical Career. 2012 Residency Maxillofacial Surgery; 2017 Board Certified Maxillofacial Surgeon, 2018 - today Consultant (Oberarzt) Maxillofacial Surgery, University Hospital Tübingen, Germany and Institute of Orthopedic Surgery and Traumatology Evaluation, Tübingen, Germany.
Evaluation. Examiner in more than 10 Clinical Trials, Principal Investigator of two Clinical Trials.
Publications. 11 Medline listed publications

Curriculum Vitae Professor Hans P. Zenner

Clinical Education. 1972 M.D. University of Mainz, Germany; Studies in Paris, France. 1974 Dr. med.

Clinical Carreer. 1973 Medical Assistant at University of Heidelberg, Germany. 1974-1976 Post-Doc at the Dept. of Biochemistry, University of Würzburg, Germany. 1976 Residency Dept. of ORL, Head&Neck Surgery, University of Würzburg; 1981 Habilitation in Oncology, Oberarzt (Consultant) and Docent University of Würzburg. 1985 Visiting Scientist at Michigan University Ann Arbor, USA, and at Washington University St. Louis, USA. 1986 Professor of ORL, Head&Neck Surgery, University of Würzburg; 1988-2016 Distinguished Professor of ORL, HeadNeck Surgery, and Chairman Dept. of ORL, Head&Neck Surgery, University of Tübingen, Germany.

Clinical Studies. 1986-2021 Director Clinical Trials according to MPG or AMG. 2017-2018 Chairman AMG/MPG Ethics Commission Tübingen, before that member.

Honorary Offices. 2009 and 2010 President German Association of Physicians and Scientists (GDNÄ). Since 2010-2021 Board Member, German Natl. Acad. Sci. Leopoldina.

Awards. Various awards including Leibniz Prize, Humboldt Medalist. Several International Memberships of Honor and several Doctorates honoris causa.

Clinical Research Focus. Group leader Development and Clinical Trials Medical Devices.

Publications. More than 300 Medline listed publications, h index 61, 8 books, around 70 patent publications.

Publications based on the study

No paper has yet been published.

Patient Data Listings

Discontinued patients
See Chapter 12.

Protocol deviations
See Chapter 12

Individual Efficacy Response data
Not displayed for data protection reasons.

Patients excluded from the efficacy analysis
See chapter 12

Demographic data
See Chapter 13

Adherence Data
See Chapter 14

Compliance Data
See Chapter 14

Burden listings
See Chapter 18

Adverse event listings
See Chapter 18
For data protection reasons individual response data are not displayed.

SIGNATURES

Name	Function		The authors agree with the contents of the report
Prof. Dr. med. K. Weise, M.D., Professor of Orthopedic Surgery & Traumatology	Author	Oct. 8th, 2021	
Dr. Dr. H. Weise, M.D., M.DD. Medical Specialist in Maxillofacial Surgery	Author	Oct. 8th, 2021	
Prof. Dr. med. H.P. Zenner, M.D., Professor of Head&Neck Surgery	Author	Oct. 8th, 2021