

Effectiveness of an LLM-integrated Chatbot Intervention for Physical Activity Habit Formation in Prehypertensive Adults: A 12-Month Randomized Controlled Trial Protocol

Haoming Ma

Chinese Academy of Medical Sciences & Peking Union Medical College

Guangnan Liu

Chinese Academy of Medical Sciences & Peking Union Medical College

Runyuan Pei

Chinese Academy of Medical Sciences & Peking Union Medical College

Sijia Li

Chinese Academy of Medical Sciences & Peking Union Medical College

Zhaoqi Liu

Chinese Academy of Medical Sciences & Peking Union Medical College

Aoqi Wang

Chinese Academy of Medical Sciences & Peking Union Medical College

Xingyi Tang

Chinese Academy of Medical Sciences & Peking Union Medical College

Li Qiao

Chinese Academy of Medical Sciences & Peking Union Medical College

Rongrong Huang

Chinese Academy of Medical Sciences & Peking Union Medical College

Jianping Zhang

Chinese Academy of Medical Sciences & Peking Union Medical College

Hongzhen Cui

University of Science and Technology Beijing

Yuyang Zhang

Chinese Academy of Medical Sciences & Peking Union Medical College

Meihua Piao

`parkmihua@snu.ac.kr`

Chinese Academy of Medical Sciences & Peking Union Medical College

Study protocol

Keywords: Physical activity, Prehypertension, Chatbot, Large language model, Habit Formation

Posted Date: July 31st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-5470873/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Prehypertension is an important risk factor for hypertension and related cardiovascular diseases. Physical activity (PA) is an effective non-drug approach for managing blood pressure and preventing the progression of prehypertension; however, maintaining long-term PA habits is often challenging. Combining behavior change theories with wearable devices and large language model (LLM)-based chatbots may offer more adaptive and personalized support by integrating user profiles, real-time activity data, and conversational feedback to address individual barriers to PA. However, little is known about the effectiveness of these interventions.

Objective

This study aims to assess the effectiveness of HabitBot, an LLM-integrated PA behavior change chatbot, in promoting PA habits among adults with prehypertension. It also explores how PA habits are formed and evaluates the intervention's usability and acceptance.

Methods

This is a 12-month, two-arm parallel, pragmatic randomized controlled trial (RCT) with an embedded 3-month intensive longitudinal study. A total of 118 participants will be recruited from two companies and social media platforms in Beijing, China. Participants will be randomly assigned to one of two groups: the intervention group, which will receive a wearable blood pressure monitoring smartwatch paired with HabitBot, a chatbot powered by an LLM designed to support PA habit formation through personalized PA prescriptions, health advice, and motivational feedback; or the control group, which will receive the same smartwatch, a standard PA prescription, and access to a health app offering general education on blood pressure management. The primary outcome is daily step count, measured via the smartwatch. Secondary outcomes include physical activity levels, blood pressure, body composition, and psychological factors relevant to behavior change theories. Data will be collected at baseline, 3, 6, and 12 months.

Results

This study will provide insights into the effectiveness of HabitBot in promoting PA habits and improving health among prehypertensive adults. It will also explore the mechanisms underlying PA habit formation through detailed longitudinal data analysis.

Conclusion

The results of this study could help shape future LLM-integrated interventions aimed at long-term PA promotion and prehypertension management. By examining both reflective and reflexive pathways of habit formation, this study advances digital health strategies for reducing cardiovascular risk.

Trial Registration

Chinese Clinical Trials Registry, ChiCTR2400085073, Registered May 30, 2024

Introduction

Prehypertension represents a critical window for early cardiovascular disease prevention. Hypertension remains a major contributor to cardiovascular morbidity and mortality worldwide, with global prevalence estimates indicating that over 30% of adults are affected(1). In 2015, approximately 8.5 million fatalities were associated with elevated blood pressure, especially in low- and middle-income countries(2). Prehypertension (i.e., normal high blood pressure in China), characterized by systolic blood pressure between 120–139 mmHg and diastolic blood pressure between 80–89 mmHg, is a crucial target for early intervention to prevent progression to clinical hypertension. In China, approximately 28.5% of adults with prehypertension develop hypertension within 2 years if preventive measures are not implemented(3). The 2024 Chinese Guidelines for the Prevention and Treatment of Hypertension specifically distinguish between normal high blood pressure and normal blood pressure and emphasize lifestyle interventions for this population(4).

Physical activity (PA) is a core non-pharmacological strategy for prehypertension management, yet insufficient and unsustained PA remains common. Regular PA can reduce the risk of progression to hypertension, with average blood-pressure-lowering effects comparable to those of commonly used antihypertensive medications in some analyses(5, 6). Current guidelines generally recommend at least 150–300 minutes per week of moderate-intensity aerobic PA, or 75–150 minutes per week of vigorous-intensity aerobic PA, together with muscle-strengthening activities on at least two days per week when feasible(7). However, many adults fail to achieve or maintain recommended PA levels, and exercise advice alone may be insufficient to produce sustained cardiovascular risk reduction among individuals with prehypertension or hypertension(8). Common barriers to PA adherence include lack of time, fatigue, competing work or family responsibilities, low motivation, low self-efficacy, uncertainty about appropriate exercise intensity, lack of social support, environmental barriers, and disrupted routines(8). These barriers suggest that interventions for adults with prehypertension should not only prescribe PA but also help individuals translate PA recommendations into feasible, repeatable daily practices(9).

Although previous interventions involving adults with prehypertension have produced short-term improvements in blood pressure and related lifestyle behaviors, they have commonly relied on supervised or structured exercise programs, standardized exercise advice, and time-limited behavioral or multicomponent lifestyle counseling(10, 11). Such approaches primarily target reflective, motivational, and self-regulatory processes and can improve knowledge, risk perception, intention, and self-

efficacy(11, 12). However, they often provide limited continuous, context-sensitive support for translating intention into repeated behavior in everyday life and for maintaining PA after the intensive intervention period has ended(13). Long-term PA maintenance requires more than knowing what to do or intending to be active; it also requires repeated performance in stable contexts so that behavior becomes easier to initiate with less conscious effort(14). Therefore, a habit formation perspective is needed to complement traditional motivational and self-regulatory approaches, particularly in interventions designed to support sustained PA among adults with prehypertension(14).

Habit formation theory offers a mechanism-based approach to sustaining PA behavior. A habit is a learned process in which exposure to a stable contextual cue automatically triggers an urge to perform a behavior, with less reliance on conscious deliberation(15). In this study, “automatic cue–behavior associations” refers to learned links between recurring cues, such as time, place, preceding activity, or social context, and the initiation of PA(12, 13). Habit formation interventions aim to strengthen these associations by promoting repeated performance in stable contexts, cue selection, action planning, self-monitoring, feedback, problem-solving, and reinforcement(16). Accumulating evidence suggests that habit formation interventions can improve PA habit strength and PA behavior(16–18). By linking PA to existing routines and reinforcing repeated performance in daily contexts, these interventions may help PA become less dependent on momentary motivation and more likely to be initiated automatically(19). This evidence supports integrating habit-formation strategies into PA interventions to enhance long-term maintenance.

Digital health technologies may provide scalable and continuous support for PA behavior change and habit formation. Wearable activity trackers can provide objective monitoring and feedback, and evidence suggests that wearable-based interventions can increase PA in adults(20). Chatbot-based interventions may further support PA promotion by providing interactive education, reminders, feedback, and motivational support. Two systematic reviews found that chatbot interventions were effective in improving lifestyle behaviors, including total PA, daily step count, and moderate-to-vigorous PA(21, 22). Large language models (LLMs) may extend earlier rule-based chatbots by enabling more flexible natural-language responses, multi-turn dialogue, contextual understanding, and individualized problem solving(23). These functions may be particularly relevant for PA habit formation because users often need timely support when they miss planned exercise, encounter barriers, need to adjust goals, or need help identifying feasible cues and rewards(24, 25).

Despite their potential, LLM-integrated chatbots have not yet been adequately evaluated as tools for sustained PA behavior change. Existing studies of LLM-based conversational agents in health behavior change have largely been short-term or formative and have primarily examined feasibility, usability, and user experience rather than long-term behavioral outcomes(23, 26). Moreover, how these systems may influence behavior remains poorly understood, particularly among adults with prehypertension. By delivering context-sensitive, personalized, and adaptive support, LLM-integrated chatbots may enhance engagement and help users address motivational, self-regulatory, and contextual barriers to sustained PA. However, these proposed benefits require rigorous evaluation in long-term intervention studies.

The present study addresses three specific gaps. First, few interventions for adults with prehypertension integrate quantified PA guidance, objective wearable-based monitoring, and continuous behavioral support within a single intervention. Second, existing PA interventions have predominantly targeted reflective determinants, such as knowledge, intention, and self-efficacy, with less attention to reflexive processes involved in habit development, including context stability, intrinsic reward, and habit strength. Third, limited evidence is available on whether an LLM-integrated chatbot can produce sustained improvements in PA and through which reflective and reflexive mechanisms such changes may occur. Accordingly, this protocol describes a 12-month randomized controlled trial evaluating HabitBot, an LLM-integrated chatbot designed to promote PA and facilitate PA habit formation among adults with prehypertension. Daily step count is the primary behavioral outcome because it provides an objective and continuously measurable indicator of PA, whereas PA habit strength and related reflective and reflexive psychological constructs are assessed as secondary and mechanistic outcomes. A 3-month intensive longitudinal assessment embedded within the trial will further examine the temporal processes through which PA behavior and habit strength change during the intervention.

Methods

Aims of the study

This protocol paper outlines the design, intervention components, outcome measures, and planned analyses for a 12-month randomised controlled trial of HabitBot, an LLM-integrated PA behaviour-change chatbot for adults with prehypertension. The trial aims to evaluate HabitBot's effectiveness in increasing PA, primarily assessed by average daily step count. It also seeks to explore PA habit formation and its underlying reflective and reflexive mechanisms through intensive longitudinal data collection, and to assess participants' engagement, usability, and acceptance of the intervention.

Study Design

This study is a 12-month, two-arm parallel, randomized controlled trial (RCT) with an embedded intensive longitudinal study lasting 3 months. Given the controlled recruitment procedures and research-specific measurement schedule, this trial is more accurately characterized as an efficacy trial aimed at establishing proof of concept for the HabitBot intervention under standardized conditions. This study was reported in accordance with the Guidelines for Reporting Outcomes in Trial Protocols (SPIRIT-Outcomes 2022)(27). A total of 118 participants will be recruited from two companies and through social media platforms for the intervention. This study was registered in the Chinese Clinical Trials Registry (Registration Number: ChiCTR2400085073) and received ethical approval from Peking Union Medical College (Ethics Approval Number: PUMCSON-2024-23). This manuscript reflects the protocol version 1.0 (dated November 24, 2024). The Chinese Academy of Medical Sciences and Peking Union Medical College (School of Nursing) sponsored the trial. Figure 1 illustrates the overall study design and timeline, including the baseline assessment, 3-month intensive intervention period, follow-up

assessments at 3, 6, and 12 months, and daily measurements of psychosocial constructs during the intensive phase.

Note

EMA, ecological momentary assessment. The green bar indicates the 3-month intensive EMA period during months 1–3. The enlarged 24-hour panel illustrates a representative day within this period, during which participants complete a single evening EMA survey via the study WeChat mini program.

Study Setting and Recruitment

The study is being conducted in Beijing, China, starting on October 8, 2024, and will be completed by December 15, 2025. Both companies are large technology corporations in Beijing, where the workforce is predominantly sedentary office workers engaged in computer-based work with extended screen time and low physical activity levels—a demographic with a documented elevated prevalence of prehypertension(28). Employees in this sector represent an increasingly common urban occupational profile at elevated cardiovascular risk, making them a high-priority target population for early intervention. Both active and passive recruitment strategies were employed to enhance sample representativeness. Passive recruitment involved posting digital flyers on social media, allowing interested individuals to contact the research team via WeChat, a widely used Chinese messaging and social media platform. Active recruitment involved direct engagement with companies, in which the research team introduced the study to management and, upon approval, posted recruitment notices on company premises. Participants from both recruitment channels were required to attend two in-person blood pressure measurements conducted by registered nurses and to complete a baseline questionnaire to determine eligibility. A flowchart of participant recruitment and study implementation is shown in Fig. 2.

Participants

Participant recruitment will commence in October 2024 and will be completed by January 15, 2025. Before randomization, eligibility screening will be conducted in three steps. First, potential participants will complete a structured self-report health questionnaire covering medical history, medication use, smoking and alcohol status, pregnancy or lactation status, and history of major cardiovascular diseases or severe health conditions, including myocardial infarction, stroke, heart failure, renal failure, cancer, and other conditions that may affect safe participation in physical activity. Second, registered nurses will review the questionnaire and conduct two in-person blood pressure measurements on different days to confirm prehypertension status. Third, participants who report relevant cardiovascular or other severe health conditions, present with abnormal findings, or have uncertain exercise suitability will be referred for a cardiologist's assessment before enrollment. Participants considered unsuitable for exercise intervention after clinical review or cardiologist assessment will be excluded.

Participants must meet the following criteria before randomization: 1) aged 18–60 years with prehypertension (systolic BP 120–139 mmHg or diastolic BP 80–89 mmHg, measured twice on different days). 2) physically inactive, defined as International Physical Activity Questionnaire-Short Form score indicating < 150 minutes of moderate-to-vigorous physical activity per week during the past month; 3) willing and able to follow the intervention program; 4) able to read Chinese; and 5) capable of using chatbots and related technologies.

The exclusion criteria were as follows: 1) diagnosed hypertension or use of antihypertensive medication; 2) major cardiovascular diseases or other severe health conditions, including recent myocardial infarction, stroke, heart failure, renal failure, or cancer; 3) acute health conditions that may affect safe participation in physical activity; 4) cognitive impairment or communication disorders; 5) participation in other intervention studies; 6) pregnancy or lactation; and 7) being considered unsuitable for exercise intervention after clinical review or cardiologist assessment.

All participants will provide written informed consent before enrollment. They will be informed of their right to withdraw from the study at any time without penalty. Participants will receive a HUAWEI WATCH D smartwatch at enrollment as the primary data collection device throughout the trial. Upon successful completion of all 12-month follow-up assessments, the participant will retain the smartwatch as compensation for their time and contribution. Participants will be instructed to continue their usual medical care and lifestyle during the trial and to avoid enrolling in any new formal exercise or lifestyle programs outside the study. Any changes in health status or medication during the trial should be reported to the research team. Changes in health status or medication use during the trial will be captured systematically through follow-up questionnaires at 3, 6, and 12 months, as well as participant-initiated reports. Participants will be instructed to report any new diagnosis, medication change, adverse event, or condition affecting exercise safety to the research team as soon as possible. All reported changes will be documented in the study records and reviewed by the research team.

To promote participant retention, the research team will maintain regular contact with participants through WeChat or phone. Reminders about upcoming follow-up assessments will be sent, and scheduling will be made as convenient as possible for participants. Participants who discontinue or deviate from the intervention will be encouraged to attend all remaining follow-up assessments unless they withdraw their consent entirely.

Randomization and Blinding

In this study, randomization will be performed at the cluster level for company-recruited participants for logistical and feasibility reasons, given the organizational structure of company-based recruitment. While cluster randomization in co-located workgroups does not fully prevent contamination—as coworkers may interact and share information about the intervention—we consider this risk manageable given that the primary outcome (step count) is measured objectively. Participants will be unaware of the study hypothesis. Clusters will be defined based on pre-existing workgroups within each participating

company. Each company will contribute approximately 11 clusters, for a total of 22, with each cluster comprising 4–5 participants. These clusters will include individuals working in the same location and functional department, as verified by the company's HR department. Teams with 4–5 members will be included; teams smaller than 4 or larger than 6 will be excluded or split accordingly. This structure ensures that randomization is applied to groups rather than to individuals within a company setting.

For participants recruited via social media, individual randomization will be used to assign them to the study without the constraints imposed by clustering. Social media recruits will be randomly assigned to either the intervention or control group using the same computer-generated sequence used for the cluster assignments. The rationale for using both recruitment channels is to enhance the representativeness of the sample among occupational and community-dwelling prehypertensive adults: company-based recruitment targets sedentary office workers with a high prevalence of prehypertension, while social media recruitment captures a broader community population.

Randomization for both company-based clusters and socially recruited individuals will be conducted by an independent statistician using the PROC PLAN procedure in SAS (version 9.4). Participant blinding cannot be guaranteed given the digital nature of the intervention, but participants will be unaware of the study hypothesis to minimize expectancy effects. To further reduce performance bias, the control group will receive a health management app compatible with the HUAWEI D smartwatch that also provides general health education on physical activity, diet, and sleep.

Sample Size

The sample size was calculated for the primary between-group comparison of objectively measured physical activity. Two design considerations were incorporated. First, the active control group receives a wearable activity tracker and access to a health application, both of which may independently increase physical activity, thereby reducing the between-group effect compared with studies using passive control conditions. Second, most participants are recruited through participating companies and randomized at the company level, requiring adjustment for within-cluster correlation.

Gal et al. reported a standardized mean difference of 0.51 for digital physical activity interventions(29). To account for the expected attenuation of the treatment effect due to the active control condition, this estimate was reduced by approximately 20%, yielding a conservative anticipated standardized mean difference of 0.40. This value is also broadly consistent with the effect size of 0.44 reported by Singh et al. for chatbot-based physical activity interventions(21).

For a two-group comparison with a two-sided significance level of 0.05, 80% statistical power, and a standardized mean difference of 0.40, approximately 200 evaluable participants are required under individual randomization. For the company-recruited component, the design effect was estimated using $1 + (m - 1) \times \text{ICC}$. Assuming a mean cluster size of 4.5 and an intraclass correlation coefficient of 0.05, the cluster-specific design effect was 1.175. Because approximately 80% of participants will be

randomized by company and 20% will be randomized individually, a weighted design effect of approximately 1.14 was applied.

After adjustment for the partially clustered design, approximately 228 evaluable participants are required. Allowing for an anticipated dropout rate of 20%, the final recruitment target was set at 286 participants. For operational feasibility and balanced allocation, the target may be rounded to 288 participants, with 144 participants allocated to each study group.

Intervention Arm

Participants in the intervention arm will receive a Huawei Watch D smartwatch and use HabitBot for 12 months. HabitBot is an LLM-integrated conversational intervention delivered via a WeChat mini program, a lightweight application embedded within WeChat. Its development, technical architecture, theory-informed design, and safety procedures are described in detail in our previous development study(24). Briefly, HabitBot was developed using a user-centered, multidisciplinary approach and was informed by the Health Action Process Approach (HAPA) and Habit Formation Theory to support both the reflective and automatic processes involved in PA behavior change(15, 30). These theory-informed components are delivered through a structured prompting approach in which each LLM-supported interaction incorporates participant profile information, baseline assessment data, current PA goals, recent PA plan and completion status, wearable-derived activity and blood pressure data, user questions or feedback, and predefined safety and scope-control instructions.

In the system design, the HAPA is operationalized through functions that support risk awareness, outcome expectancy, self-efficacy, action planning, coping planning, and self-regulation. Specifically, HabitBot helps participants understand the relevance of PA for prehypertension management, establish individualized PA goals, receive tailored exercise recommendations, identify anticipated barriers, generate coping strategies, and review progress through interactive dialogue. Habit Formation Theory is operationalized through repeated PA planning, context-specific reminders, stable cue-behavior routines, self-monitoring, and reinforcement of repeated PA performance in daily life. These theory-informed functions are intended to help participants move from intention formation and goal setting to repeated PA performance and gradual habit development.

Based on the Behaviour Change Technique Taxonomy v1, the core behavior change techniques embedded in HabitBot include goal setting (behavior), goal setting (outcome), action planning, problem solving, feedback on behavior, self-monitoring of behavior, social support, instruction on how to perform the behavior, information about antecedents, behavioral practice/rehearsal, habit formation, and self-reward. The detailed mapping of HabitBot components to theoretical constructs and behavior change techniques has been reported in our previous development study(24).

As shown in Fig. 3, HabitBot includes four modules: My Coach, My Plan, My Data, and User Center. My Coach offers personalized PA goal setting, exercise recommendations, and interactive support to

overcome barriers. My Plan displays scheduled activities and sends reminders and motivational prompts. My Data presents wearable-derived measures, including step count, heart rate, and blood pressure, to support self-monitoring and progress tracking. User Center lets participants manage personal settings, review their PA history, and customize intervention preferences.

Participants interact with HabitBot through both scheduled and user-initiated conversations. Structured interactions occur during the initial assessment, weekly goal setting and review, scheduled PA reminders, follow-up for missed or incomplete activities, and scheduled follow-up assessments. Participants may also initiate conversations at any time to ask questions, report difficulties, request PA adjustment, or seek motivational support. Participants complete an initial assessment of their health status, injury history, baseline PA level, exercise preferences, available time, and perceived barriers. Based on these data, HabitBot recommends an individualized weekly PA goal and generates a weekly PA plan. Participants may accept or modify the recommended goal through multi-turn dialogue before confirming. The confirmed weekly goal is translated into scheduled daily PA activities and reviewed weekly. Reminders are delivered via WeChat mini program notifications and are triggered by scheduled PA activities, missed or incomplete planned activities, and weekly goal review time points.

During the intervention, personalized recommendations are further adjusted based on each participant's current PA goal, recent PA plan, goal completion status, wearable-derived activity and blood pressure data, user-reported difficulties, and previous interaction history. Plan modifications are triggered by predefined conditions, including insufficient goal completion, missed or incomplete scheduled activities, wearable-derived PA levels substantially below the planned target, user-reported barriers or discomfort, changes in health status, or participant requests for adjustment.

For each LLM-supported interaction, the input includes a structured system prompt, the participant's profile, current PA goals, recent PA plan and completion status, relevant wearable-derived data, the user's question or feedback, and predefined safety and scope-control instructions. The LLM then generates context-sensitive responses, such as motivational feedback, exercise adjustment suggestions, barrier-solving strategies, and action or coping planning support. Multilayered safety mechanisms restrict responses to the PA behavior change and prehypertension management domain and escalate potentially unsafe or out-of-scope health-related queries for professional review. The intervention is delivered in Simplified Chinese.

Note

Module labels in the interface reflect the Chinese-language version; English translations used in the text (My Coach, My Plan, My Data, User Center) correspond to the primary functional descriptions of these modules.

Active Control Arm

Participants in the control group will receive a one-time, individually tailored PA prescription delivered in a single 30-minute face-to-face session with an advanced practice registered nurse, following the American College of Sports Medicine (ACSM) guidelines for prehypertensive individuals. Participants will also be instructed to contact the researchers with any health-related questions during the intervention. Additionally, to isolate the net effect of traditional mobile health, the control group will receive a wearable smartwatch and a compatible health app, CoHealth, which will provide visualization of smartwatch-collected data and general health education covering PA, diet, and sleep.

Outcomes and Measures

Overview of Assessment Schedule

As depicted in Table 1, outcomes and related variables will be assessed at baseline and at 3, 6, and 12 months, with additional daily or continuous data collection during the intervention period. An intensive 3-month ecological momentary assessment (EMA) period will also be conducted during months 1–3 of the intervention. During this period, participants in both groups will complete daily evening EMA surveys, approximately between 20:00 and 22:00, via the study WeChat mini program to assess reflective and reflexive pathway variables related to PA habit formation. Each survey includes the Health Behavior Questionnaire (HAPA-based items assessing the reflective pathway: PA intention, outcome expectancy, risk perception, self-efficacy, social support, action planning, and coping planning) and the Self-Report Behavior Automaticity Index (SRBAI; assessing the reflexive/automatic pathway: habit strength, context stability, intrinsic reward, intrinsic motivation, PA affect, and PA identity).

Primary Outcomes

The primary outcomes are average daily step count and physical activity level over the specified assessment periods. Step count will be collected continuously throughout the 12-month intervention using the HUAWEI WATCH D smartwatch. For outcome assessment at each time point, average daily step count will be calculated over a predefined 7-day assessment window at baseline and immediately before the 3-, 6-, and 12-month follow-up assessments. Only valid wear days, defined as days with at least 12 hours of smartwatch wear, will be included in the calculation. The study will use step-count values generated by the manufacturer's proprietary algorithm based on the smartwatch's embedded accelerometer data; the research team will not access or process raw acceleration data to derive step count. The processed step-count data will be transmitted in real time to a cloud server via an application programming interface (API) that relays data directly from the user's wearable device. Although direct validation evidence for step-count measurement using the HUAWEI WATCH D is limited, a daily-life validation study of the Huawei GT2 smartwatch reported acceptable accuracy for measuring physical activity, including step count, among adults⁽³¹⁾. Therefore, smartwatch-derived step count will be used as an objective indicator of PA in this trial. To ensure consistency in between-group comparisons, all participants in both groups will use the same HUAWEI WATCH D device model and follow the same wear instructions. Physical activity level will be assessed using the International Physical Activity

Questionnaire (IPAQ) at baseline and 12 months. Participants will be asked to wear the smartwatch on the non-dominant wrist for at least 12 hours per day and to manually initiate smartwatch-based blood pressure measurements at least three times per week during the trial. Wearable-derived heart rate, blood pressure, step count, and wear compliance data will be collected continuously throughout the 12-month intervention period. The HUAWEI WATCH D, an oscillometric wrist blood pressure monitor, was selected because it provides wrist blood pressure measurements that meet the requirements of the AAMI/ESH/ISO Universal Standard(32). The questionnaire-based measures used in this study, including the IPAQ(33), HAPA-based Questionnaire(34), SRBAI(35), and Sinicized Chatbot Usability Scale(36), were selected because they have been widely used in previous health behavior and digital health research and have demonstrated acceptable reliability and validity.

Table 1
Assessment schedule by time point

Measures	Baseline	Daily / continuous for 12 months	Daily EMA during months 1–3 months	3 months	6 months	12 months
Sociodemographic characteristics	✓					
Medical history, medication use, smoking, alcohol status, and family history	✓			✓*	✓*	✓*
Screening for major cardiovascular diseases and severe health conditions	✓					
Office blood pressure measured by registered nurses	✓			✓	✓	✓
Wearable-derived blood pressure		✓				
Daily step count		✓				
Heart rate		✓				
Smartwatch wear compliance		✓				
Physical activity level measured by IPAQ	✓					✓
Height	✓					
Waist circumference	✓			✓	✓	✓
Body weight, body fat percentage, muscle mass, bone mass, BMI, metabolic age, and visceral fat grade	✓			✓	✓	✓
Reflective pathway variables: PA intention, outcome expectancy, risk perception, self-efficacy, social support, action	✓		✓	✓	✓	✓
Note. *Changes in medical history, medication use, and health status will be recorded during follow-up assessments if applicable. EMA, ecological momentary assessment; IPAQ, International Physical Activity Questionnaire; PA, physical activity; BUS, Chatbot Usability Scale.						

Measures	Baseline	Daily / continuous for 12 months	Daily EMA during months 1–3 months	3 months	6 months	12 months
planning, and coping planning						
Reflexive pathway variables: habit strength, context stability, intrinsic reward, intrinsic motivation, PA affect, and PA identity	✓		✓	✓	✓	✓
Chatbot engagement: active usage frequency, feature utilization, and interaction duration		✓				
Chatbot usability and acceptance measured by BUS				✓		
Note. *Changes in medical history, medication use, and health status will be recorded during follow-up assessments if applicable. EMA, ecological momentary assessment; IPAQ, International Physical Activity Questionnaire; PA, physical activity; BUS, Chatbot Usability Scale.						

Secondary Outcomes

1) Blood pressure: Registered nurses on the research team used Omron U10K automatic blood pressure monitors for measurements. The day before the measurement, participants were advised to avoid caffeine, nicotine, and strenuous exercise. Before the formal measurement, the participants were asked to rest for at least 10 minutes in a quiet room. For the first blood pressure measurement, the left and right arms are measured separately, and the higher value is used as the first reading. BP should be measured a second time at 1–2 minute intervals, and the average of the two readings should be taken. If the difference between the two measurements exceeds 10 mmHg, the blood pressure will be measured once, and the average of the three measurements will be the final blood pressure value.

2) Height and waist circumference will be measured twice using certified instruments.

3) Body weight, body fat percentage, muscle density, bone mineral density, BMI, metabolic age, and visceral fat grade: all measurements were performed using the TANITA Inner Scan V multifunction body composition analyzer. By analyzing the impedance values of the current passing through different tissues in the body, the device automatically calculates and displays parameters such as body weight, body fat percentage, muscle mass, bone mass, and BMI. In addition, the device extrapolated participants' metabolic age and visceral fat grade from these indicators.

- 4) Reflective pathway: We used the Health Behavior Questionnaire (based on the HAPA model) to assess participants' cognitive processes related to physical activity, including PA intention, outcome expectancy, risk perception, PA self-efficacy, social support, PA action planning, and PA coping planning.
- 5) Reflexive pathway: We used the Self-Report Behavior Automaticity Index (SRBAI), which includes PA habit strength (instigation), PA context stability, PA intrinsic reward, PA intrinsic motivation, PA affect, and PA identity, to evaluate the reflexive pathway of developing PA habits over time.
- 6) Usability evaluation: The Sinicized and culturally adjusted Chatbot Usability Scale (BUS) will be provided to the participants for evaluation at the third month of the study.
- 7) Sociodemographic characteristics: age, sex, education, income levels, and medication use; history of disease, smoking, and alcohol use; and family history of hypertension or cardiovascular disease were collected and used as confounders in the analyses.
- 8) Engagement, usability, and adherence: Chatbot engagement will be operationalized as: (a) daily active usage frequency (number of user-initiated sessions per day), (b) feature utilization (proportion of My Coach, My Plan, and My Data modules accessed), and (c) total interaction duration per session, all automatically logged by the HabitBot server. Usability will be assessed at 3 months using the Sinicized BUS, which evaluates participants' perceived ease of use, usefulness, interaction quality, and overall acceptability of the chatbot. Intervention adherence will be tracked using multiple indicators, including smartwatch wear compliance, completion of planned PA activities, frequency of smartwatch-based blood pressure measurements, EMA survey completion during the 3-month intensive period, and chatbot usage patterns. Smartwatch wear compliance will be monitored via the device's continuous wear log. Participants who wear the device for fewer than 12 hours per day on more than 30% of monitored days will be flagged for follow-up contact and adherence support.

Data Management

Data management will adhere to standardized protocols to ensure data quality and participant confidentiality throughout the study. All participants will be assigned unique identification codes to maintain anonymity and allow tracking in the case of missing follow-up data. During data collection, a double-entry checking procedure will be implemented to minimize data entry errors and enhance accuracy. Data will be entered into Epidata 3.1, allowing for cross-verification of entries to ensure consistency and reliability before analysis. If abnormal or inconsistent data are detected, the original questionnaires or records will be reviewed and verified before proceeding to the next step of the analysis. All data will be stored securely in a password-protected electronic database, with access restricted to the authorized members of the research team. Regular data audits will be conducted to ensure the integrity of the data. Any changes to the data will be logged and documented to maintain transparency.

Statistical Analysis

All analyses will be conducted on an intention-to-treat basis, including all participants in the groups to which they were randomized, regardless of adherence or dropout (if the outcome data are available)(37). Prior to breaking the randomization code for analysis, the data will be cleaned and verified as described. Statistical significance will be evaluated at the 0.05 two-sided level for the primary outcomes.

Baseline characteristics of the two groups will be summarized to assess comparability. Continuous variables will be reported as mean $\pm$ standard deviation (or median and interquartile range if not normally distributed), and categorical variables will be reported as frequencies and percentages. Any notable baseline imbalances will be noted and adjusted for in the analyses if needed. Baseline differences between groups will be evaluated using independent t-tests or Mann-Whitney U tests for continuous variables and chi-square tests or Fisher's exact tests for categorical variables, as appropriate. Formal hypothesis testing at baseline is for descriptive purposes only; covariates will be pre-specified in the statistical analysis plan based on clinical and theoretical relevance.

The primary outcomes (daily step count and IPAQ-measured PA) will be analyzed using mixed-effects models to account for repeated measures over time (baseline, 3, 6, and 12 months) and cluster-based randomization for a subset of participants. These models will include fixed effects for group, time, and group-by-time interaction, and will include appropriate random effects to account for within-participant repeated measures and cluster-based randomization for the subset of participants randomized by workgroup cluster. Continuous secondary outcomes such as blood pressure and weight will be analyzed using generalized estimating equations (GEE) or linear mixed models(38), as with the primary outcomes, to account for repeated measures and clustering.

To address the second aim, we will explore how the intervention influences habit formation pathways. Multilevel modeling will be employed to analyze the intensive longitudinal data (e.g., daily HAPA questionnaire items and SRBAI scores collected each evening via EMA during the first 3 months)(39). These models (e.g., multilevel regression or time-series analyses) capture within-subject changes in habit strength and determine whether these changes mediate improvements in PA behavior. We may conduct mediation analyses where changes in reflective (e.g., self-efficacy) and reflexive (habit automaticity) measures are tested as mediators between group assignment and PA outcomes at 12 months. Subgroup analyses may be performed to determine whether the intervention effects differ by factors such as age, sex, baseline activity, or recruitment source, although these will be considered exploratory.

Missing data will be handled carefully to minimize any bias. We will make every effort to follow up with participants who discontinue the intervention but not the study. If outcome data are missing for some participants at certain time points, we will examine the extent and patterns of missing data. If data are missing at random (MAR assumption is plausible), we will use mixed-model analyses that can handle incomplete repeated measures under MAR or apply multiple imputations for sensitivity analysis. We plan to perform multiple imputations for missing outcome data using relevant covariates and baseline values to create several imputed datasets, then analyze each and pool the results. The primary analysis will be

compared with a complete case analysis to ensure robustness. We will also conduct sensitivity analyses, such as per-protocol (including only those with high adherence), to determine whether the results are consistent.

No formal interim efficacy analysis is planned because of the relatively short duration and minimal risks of the trial; thus, no early stopping rules for efficacy have been defined. All hypothesis tests will be two-tailed with $\alpha = 0.05$, and confidence intervals will be provided for the effect size estimates. Statistical analyses will be performed using IBM SPSS Statistics version 27.0 and R (version 4.3.1). All analysis procedures (especially for primary outcomes) will be documented in a statistical analysis plan, which will be finalized before locking the database.

Data Monitoring

A Data Monitoring Committee (DMC) was established to oversee participant safety and data integrity during the trial. The DMC was composed of two members from the Biomedical Ethics Committee of Peking Union Medical College School of Nursing and one senior clinical researcher, none of whom were otherwise involved in this study. The DMC members have no financial or professional conflicts of interest with the trial outcomes and are independent of the investigators and sponsor. The DMC will review the progress of the trial and any safety issues at regular intervals (approximately every 3–6 months). They will evaluate the recruitment rates, data quality, and any reported adverse events.

Adverse Events Reporting

Participants are given access to report adverse events by calling the study telephone line. All researchers will be trained to record the details of any adverse events, and a participant will be followed until the event is resolved. All potential adverse events will be reported and reviewed by the principal investigator. Possible adverse effects may include injuries resulting from increased physical activity.

Furthermore, if there are any technical problems (such as a broken connection or an error in data transmission), the technical team will fix it in time within a day. If there are any significant deficiencies or performance errors of the HabitBot during this trial, the intervention must be suspended or terminated to correct the artificial intelligence program.

Ethical Considerations

All participants will provide written informed consent before enrollment after a thorough explanation of the study objectives and procedures. This study will comply with all relevant data protection regulations to ensure participant privacy and data security. Ethical approval was obtained from the Peking Union Medical College School of Nursing Ethics Committee.

Dissemination

The results of this trial will be disseminated to maximize its reach and impact. Upon completion of the data analysis, the investigators plan to publish the primary findings in a peer-reviewed scientific journal

specializing in digital health or physical activity intervention research. The manuscript will follow the CONSORT guidelines for reporting RCT results.

In addition to academic dissemination, we will communicate the results to stakeholders and participants in accessible formats. Each participant will be offered a summary of the study results (in lay language) sent via email or WeChat after the study concludes so they can learn about the overall outcomes and significance. We will also share our findings with the participating companies in the form of a brief report or presentation, as they are interested in workplace health initiatives.

To support transparency and future research, de-identified individual participant data (IPD) and statistical codes will be made available upon reasonable request after the publication of the primary results. Interested researchers can contact the corresponding author to arrange for data access. Data sharing will be subject to a data use agreement that protects participant privacy (since the dataset will be de-identified, the re-identification risk is low, but we will ensure that requesters commit to using the data responsibly).

Discussion

The HabitBot intervention represents an advancement in digital health by integrating health psychology theories with LLM-driven chatbots and wearable technology to address the challenge of long-term PA change in individuals with prehypertension. Using LLMs, HabitBot can deliver a highly customized experience by providing real-time, tailored PA recommendations that dynamically adapt to the user's needs and preferences. Recent studies have underscored the effectiveness of such personalized interventions. For instance, Gao et al. demonstrated that a mobile health intervention, combining wearables, personalized exercise prescriptions, and social media, significantly improved daily step counts and physical health among cancer survivors compared to standard care(40).

The dual-phased design of this study, with a 12-month observation period and a more intensive 3-month phase, provides a robust framework for examining the mechanisms of PA habit formation. This approach is consistent with the findings of Kwasnicka et al., who emphasized the importance of assessing habit strength and context stability over time to understand behavior maintenance(41). Moreover, by combining subjective self-reports with objective data from wearable devices, this study minimized potential measurement bias(42).

In summary, the findings of this study are expected to offer meaningful insights into optimizing AI-driven chatbot interventions for managing chronic conditions such as prehypertension. By elucidating the processes underlying habit formation in PA adherence, this study will inform future improvements in digital health tools, enhancing their adaptability, personalization, and long-term effectiveness. These contributions have the potential to advance the use of AI in chronic disease prevention and to support sustained health behavior change across diverse populations.

Limitations

Several limitations of this study design should be acknowledged. First, participant blinding is not feasible given the digital nature of the intervention, which may introduce performance or expectancy bias; however, using objectively measured smartwatch-derived step counts as a primary outcome may reduce reliance on self-reported PA. Second, although all participants will use the same smartwatch model, wearable-derived step counts and blood pressure data may still contain measurement error, and questionnaire-based measures, including IPAQ and EMA responses, may be affected by recall bias, response burden, or social desirability bias. Third, recruiting participants from large technology companies and social media platforms in Beijing may limit generalisability to the broader prehypertensive adult population, particularly individuals with lower digital literacy, different occupational backgrounds, or limited access to mobile health technologies. Fourth, although cluster randomisation is used for company-recruited participants to reduce contamination, information sharing between coworkers cannot be completely ruled out. Fifth, because the intervention and follow-up period lasts 12 months, attrition higher than the anticipated 20% dropout rate could reduce statistical power. Finally, the use of a single LLM platform, ChatGPT-4o, may limit the generalisability of findings to other LLM architectures or future model versions.

Abbreviations

PA

physical activity

BCTs

behavior change techniques

LLM

large language models

RCT

randomized controlled trial

HAPA

health action process approach

HFT

habit formation theory

API

application programming interface

AI

artificial intelligence

IPAQ

the International Physical Activity Questionnaire

ACSM

American College of Sports Medicine

SRBAI

Self-Report Behavior Automaticity Index

Declarations

This manuscript has not been published or presented elsewhere in full, but a preliminary abstract of this work was presented at the Nursing Informatics 2024(43).

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the School of Nursing, Peking Union Medical College (approval number: PUMCSON-2023-10). All the participants provided informed consent.

Consent for publication

The authors consent to publication.

Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Disclosure statement

The authors declare no conflicts of interest.

Funding

This study was funded by the Non-Profit Central Research Institute Fund of the Chinese Academy of Medical Sciences (Grant NO. 2023-RC320-01).

Authors' contributions

Every author significantly contributed to the study. H.M. led the conceptualization and design of the study, drafted the manuscript, and coordinated the development of the protocol. G.L., R.P., and S.L. contributed to the study design and critically reviewed and revised the manuscript. Z.L., A.W. and X.T. supported the development of the intervention framework and provided input on the methodology. L.Q., R.H. and J.Z. assisted in refining the study design and prepared the materials for ethical approval and trial registration. H.C. and Y.Z. contributed to the statistical analysis plan and reviewed the protocol for methodological rigor. M.P. supervised the entire study, critically revised the manuscript for intellectual content, and served as the corresponding author. All authors reviewed and approved the final manuscript and agreed to be accountable for all aspects of the work.

Acknowledgements

We would like to thank the Engineer Yuanhang Zhang from Shijiazhuang Cloud Technology Co., LTD for providing constant technical support.

References

1. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. *Nature Reviews Nephrology*. 2020;16(4):223-37. doi:10.1038/s41581-019-0244-2
2. Zhou B, Perel P, Mensah GA, Ezzati M. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. *Nature Reviews Cardiology*. 2021;18(11):785-802. doi:10.1038/s41569-021-00559-8
3. Li Z, Cao L, Zhou Z, Han M, Fu C. Factors influencing the progression from prehypertension to hypertension among Chinese middle-aged and older adults: a 2-year longitudinal study. *BMC Public Health*. 2023;23(1):339. doi:10.1186/s12889-022-14410-3
4. Committee CHPaCGR. Chinese Guidelines for the Prevention and Treatment of Hypertension (2024 revision). *Chinese journal of hypertension*. 2024;32(07):603-700. doi:10.16439/j.issn.1673-7245.2024.07.002
5. Williamson W, Foster C, Reid H, et al. Will Exercise Advice Be Sufficient for Treatment of Young Adults With Prehypertension and Hypertension? A Systematic Review and Meta-Analysis. *Hypertension*. 2016;68(1):78-87. doi:doi:10.1161/HYPERTENSIONAHA.116.07431
6. Fu J, Liu Y, Zhang L, et al. Nonpharmacologic Interventions for Reducing Blood Pressure in Adults With Prehypertension to Established Hypertension. *Journal of the American Heart Association*. 2020;9(19):e016804. doi:10.1161/JAHA.120.016804
7. Organization WH. WHO guidelines on physical activity and sedentary behaviour: World Health Organization; 2020.
8. Deslippe AL, Soanes A, Bouchaud CC, et al. Barriers and facilitators to diet, physical activity and lifestyle behavior intervention adherence: a qualitative systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*. 2023;20(1):14. doi:10.1186/s12966-023-01424-2
9. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJ, Martin BW. Correlates of physical activity: why are some people physically active and others not? *The lancet*. 2012;380(9838):258-71.
10. Li X, Chang P, Wu M, et al. Effect of Tai Chi vs aerobic exercise on blood pressure in patients with prehypertension: a randomized clinical trial. *JAMA network open*. 2024;7(2):e2354937.
11. Fu J, Liu Y, Zhang L, et al. Nonpharmacologic interventions for reducing blood pressure in adults with prehypertension to established hypertension. *Journal of the American Heart Association*. 2020;9(19):e016804.
12. Hagger MS. Habit and physical activity: Theoretical advances, practical implications, and agenda for future research. *Psychology of Sport and Exercise*. 2019;42:118-29.

13. Englert C, Rebar A, Rhodes RE, Pfeffer I. New developments in the intention-behavior gap for physical activity—recent trends, controversies, and a critical outlook. *Frontiers in psychology*. 2023;14:1119973.
14. Gardner B, Rebar AL, Zhang CQ. Making health a habit: Progress in habit theory and its application in real-world settings. *Wiley Online Library*; 2025. p. e70064.
15. Gardner B, Rebar AL. Habit formation and behavior change. *Oxford research encyclopedia of psychology* 2019.
16. Ma H, Wang A, Pei R, Piao M. Effects of habit formation interventions on physical activity habit strength: meta-analysis and meta-regression. *International Journal of Behavioral Nutrition and Physical Activity*. 2023;20(1):109. doi:10.1186/s12966-023-01493-3
17. Baretta D, Gillmann N, Edgren R, Inauen J. HabitWalk: A micro-randomized trial to understand and promote habit formation in physical activity. *Applied Psychology: Health and Well-Being*. 2025;17(1):e12605.
18. Singh B, Murphy A, Maher C, Smith AE. Time to Form a Habit: A Systematic Review and Meta-Analysis of Health Behaviour Habit Formation and Its Determinants. *Healthcare*. 2024;12(23):2488.
19. Kaushal N, Rhodes RE, Spence JC, Meldrum JT. Increasing Physical Activity Through Principles of Habit Formation in New Gym Members: a Randomized Controlled Trial. *Ann Behav Med*. 2017;51(4):578-86. doi:10.1007/s12160-017-9881-5
20. Ferguson T, Olds T, Curtis R, et al. Effectiveness of wearable activity trackers to increase physical activity and improve health: a systematic review of systematic reviews and meta-analyses. *The Lancet Digital Health*. 2022;4(8):e615-e26.
21. Singh B, Olds T, Brinsley J, et al. Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. *npj Digital Medicine*. 2023;6(1):118. doi:10.1038/s41746-023-00856-1
22. Oh YJ, Zhang J, Fang M-L, Fukuoka Y. A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss. *International Journal of Behavioral Nutrition and Physical Activity*. 2021;18:1-25.
23. Jörke M, Sapkota S, Warkenthien L, et al. Supporting Physical Activity Behavior Change with LLM-Based Conversational Agents. *arXiv preprint arXiv:2405.06061*. 2024.
24. Ma H, Cui H, Yu H, et al. Development and evaluation of an large language model-integrated chatbot intervention for physical activity habit formation in adults with prehypertension. *Digital Health*. 2026;12:20552076261421367.
25. Zhu Y, Long Y, Wang H, Lee KP, Zhang L, Wang SJ. Digital behavior change intervention designs for habit formation: systematic review. *Journal of medical Internet research*. 2024;26:e54375.
26. Mantena S, Johnson A, Oppezzo M, et al. Fine-tuning LLMs in behavioral psychology for scalable health coaching. *NPJ Cardiovascular Health*. 2025;2(1):48.
27. Butcher NJ, Monsour A, Mew EJ, et al. Guidelines for Reporting Outcomes in Trial Protocols: The SPIRIT-Outcomes 2022 Extension. *JAMA*. 2022;328(23):2345-56. doi:10.1001/jama.2022.21243

28. Zhao W, Du L, Qiu W, Wang G, Zhou D. Analysis of Influencing Factors of Prehypertension and Its Development in Occupational Population. *International Journal of Public Health*. 2025;70:1608206.
29. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: a Systematic Review and Meta-analysis. *Sports Medicine - Open*. 2018;4(1):42. doi:10.1186/s40798-018-0157-9
30. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). *Rehabil Psychol*. 2011;56(3):161.
31. Mei L, He Z, Hu L. Accuracy of the Huawei Gt2 smartwatch for measuring physical activity and sleep among adults during daily life: instrument validation study. *JMIR Formative Research*. 2024;8:e59521.
32. Wang L, Xian H, Guo J, et al. A novel blood pressure monitoring technique by smart HUAWEI WATCH: A validation study according to the ANSI/AAMI/ISO 81060-2: 2018 guidelines. *Frontiers in cardiovascular medicine*. 2022;9:923655.
33. Macfarlane DJ, Lee CCY, Ho EYK, Chan KL, Chan DTS. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). *Journal of Science and Medicine In Sport*. 2006;10(1):45-51.
34. Steca P, Pancani L, Cesana F, et al. Changes in physical activity among coronary and hypertensive patients: A longitudinal study using the Health Action Process Approach. *Psychology & Health*. 2017;32(3):361-80.
35. Gardner B, Abraham C, Lally P, de Bruijn G-J. Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *International Journal of Behavioral Nutrition and Physical Activity*. 2012;9(1):102.
36. Ma H, Pei R, Li S, Wang A, Tang X, Piao M. Chatbot Usability Scale in Chinese Users: Cross-Cultural Adaptation and Validation Study. *JMIR Human Factors*. 2026;13:e84971.
37. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC medicine*. 2010;8(1):18.
38. Zeger SL, Liang K-Y, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988:1049-60.
39. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu. Rev. Clin. Psychol*. 2008;4(1):1-32.
40. Gao Z, Ryu S, Zhou W, et al. Effects of personalized exercise prescriptions and social media delivered through mobile health on cancer survivors' physical activity and quality of life. *Journal of Sport and Health Science*. 2023;12(6):705-14.
41. Kwasnicka D, Naughton F. N-of-1 methods: A practical guide to exploring trajectories of behaviour change and designing precision behaviour change interventions. *Psychology of Sport and Exercise*. 2020;47:101570.

42. Loney T, Standage M, Thompson D, Sebire SJ, Cumming S. Self-report vs. objectively assessed physical activity: which is right for public health? *Journal of Physical Activity and Health*. 2011;8(1):62-70.

43. Ma H, Pei R, Li S, Wang A, Tang X, Piao M. Systematic Development of an AI Chatbot for Physical Activity Habit Formation in Prehypertension Individual Integrating Needs, Theories, and Evidence. *Studies In Health Technology and Informatics*. 2024;315:581-2. doi:10.3233/SHTI240226

Figures

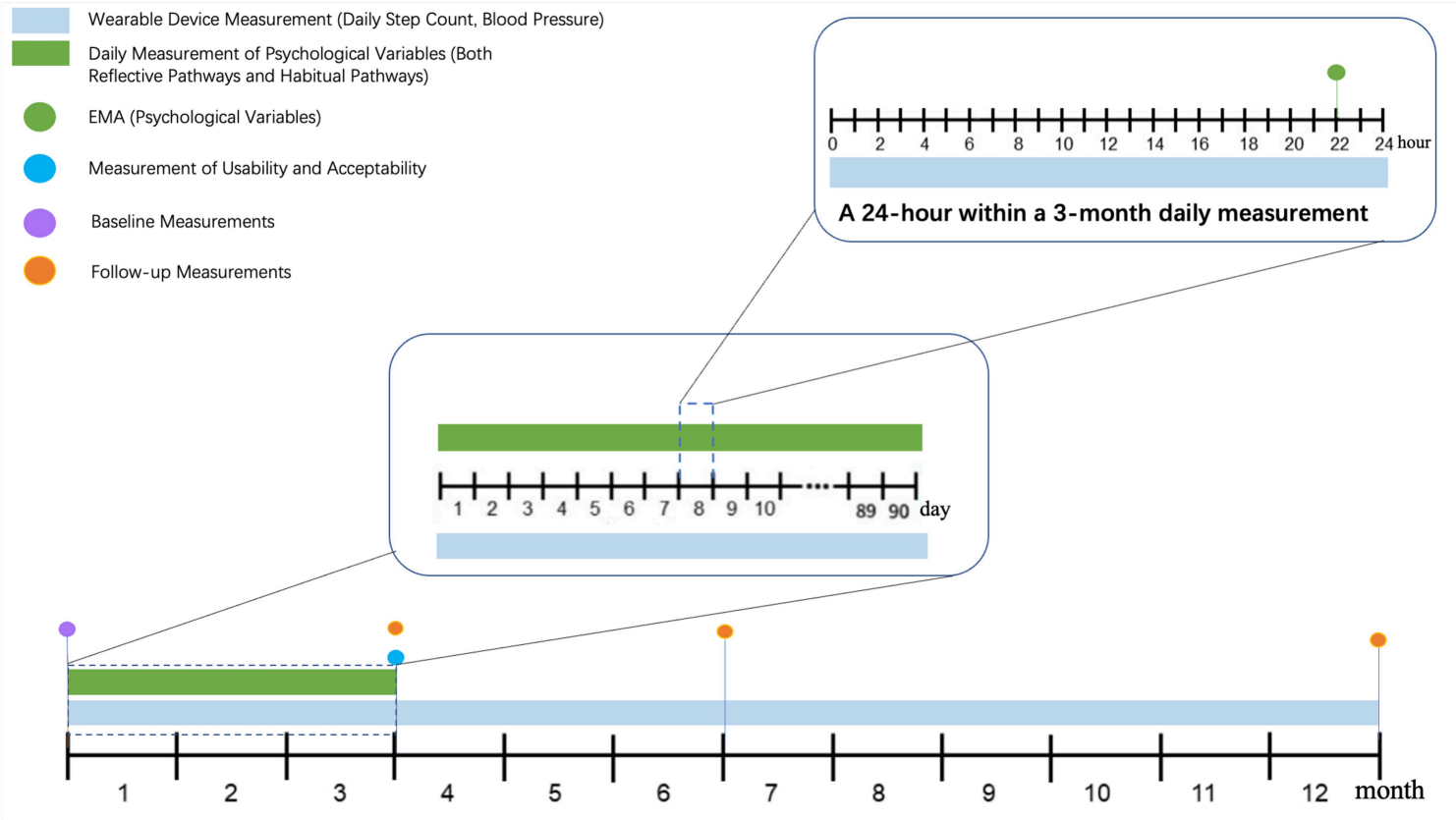


Figure 1

Study design for the 12-month two-arm randomized controlled trial embedded 3-month intensive longitudinal study

Note: EMA, ecological momentary assessment. The green bar indicates the 3-month intensive EMA period during months 1–3. The enlarged 24-hour panel illustrates a representative day within this period, during which participants complete a single evening EMA survey via the study WeChat mini program.

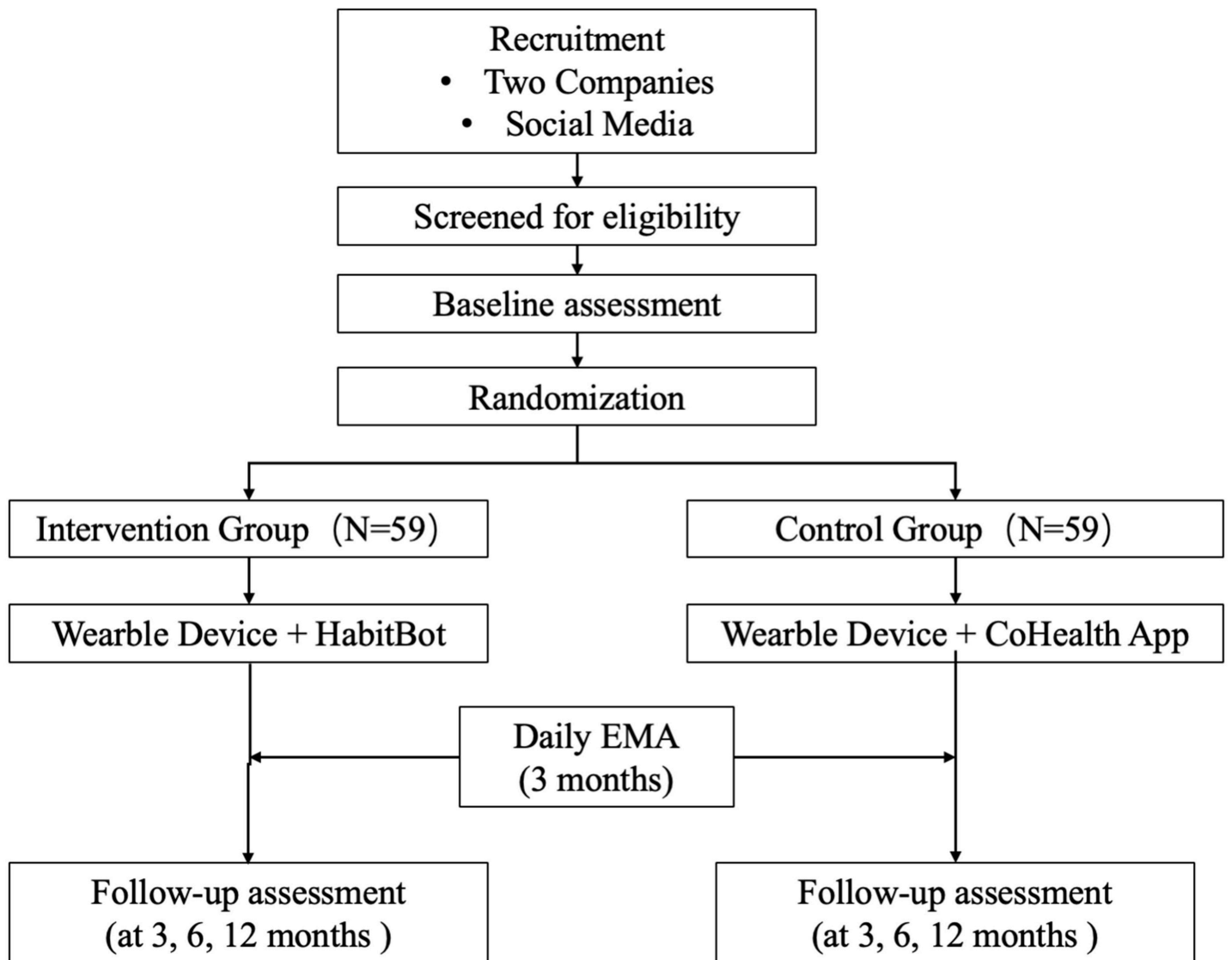


Figure 2

Flow Chart of Participant Recruitment

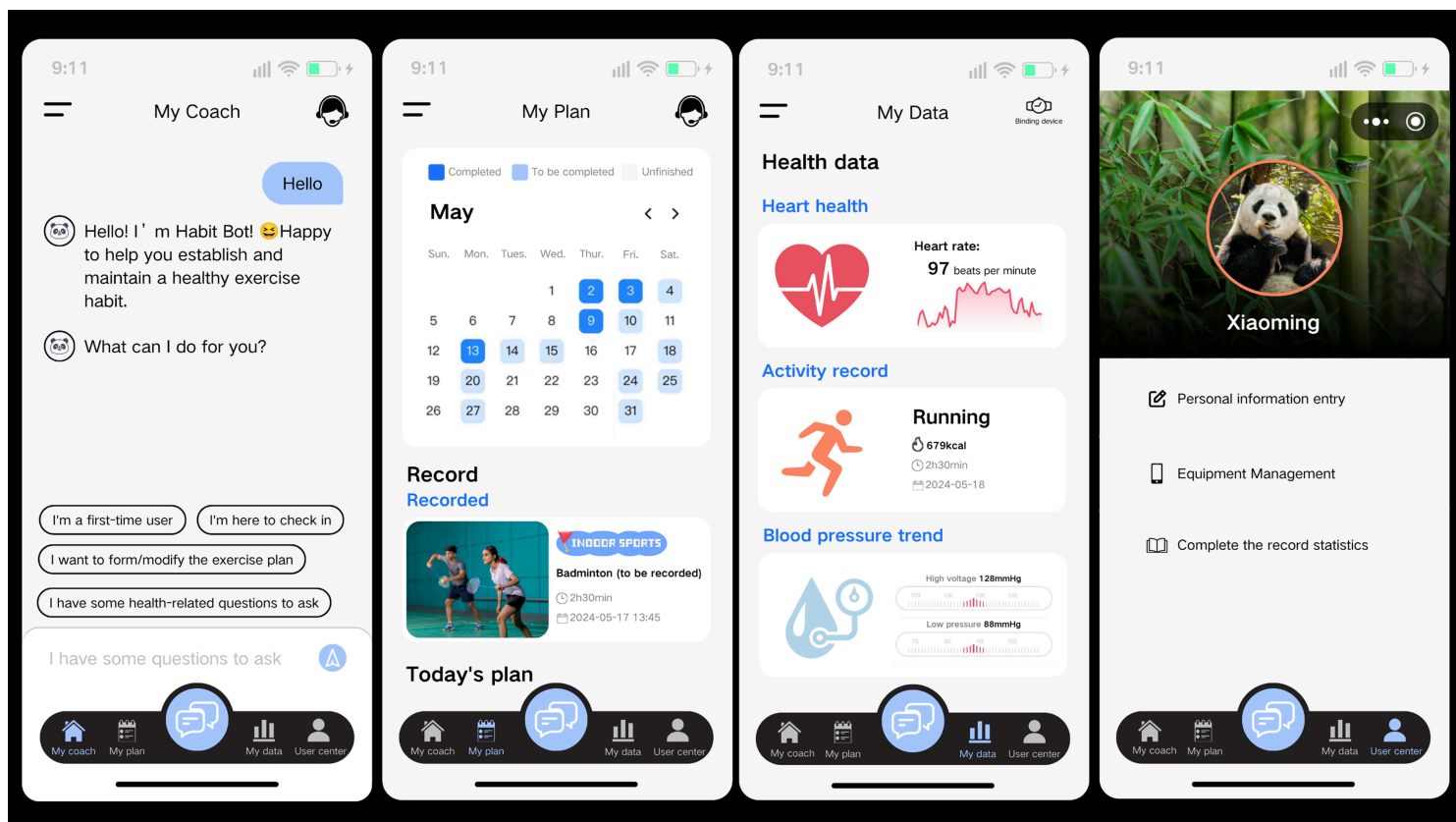


Figure 3

Screenshot from HabitBot chatbot intervention

Note: Module labels in the interface reflect the Chinese-language version; English translations used in the text (My Coach, My Plan, My Data, User Center) correspond to the primary functional descriptions of these modules.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarymaterialSPIRIT2022.pdf](#)