Blended care to discontinue BZRA use in patients with chronic insomnia disorder: a pragmatic cluster randomized controlled trial in primary care.

Kristien Coteur[1], Gilles Henrard[2], Birgitte Schoenmakers[1], Annouschka Laenen[3], Kris Van den Broeck[4], An De Sutter[5], Sibyl Anthierens[4], Dirk Devroey[6], Nadine Kacenelenbogen[7], Anne-Marie Offermans[7], Marc Van Nuland[1]

1. Department of Public Health and Primary Care, Academic Centre for General Practice, KU Leuven, Leuven, Belgium

2. Department of General Practice, Research Unit Primary Care & Health, Liège Université, Liège, Belgium

3. Leuven Biostatistics and Statistical Bioinformatics Centre (L-Biostat), KU Leuven, Leuven, Belgium

4. Department of Family Medicine and Population Health, University of Antwerp, Antwerp, Belgium

5. Department of Public Health and Primary Care, Centre for Family Medicine, Ghent University, Ghent, Belgium

6. Department of Family Medicine and Chronic Care, Vrije Universiteit Brussel, Brussels, Belgium

7. Department of General Practice, Université libre de Bruxelles, Brussels, Belgium

Corresponding author:

Kristien Coteur

Kapucijnenvoer 7 box 7001, 3000 Leuven, Belgium

kristien.coteur@kuleuven.be

0032 16 19 42 65

Summary

Study Objectives

International guidelines recommend using benzodiazepine receptor agonists (BZRA) for maximally four weeks. Nevertheless, long-term use for chronic insomnia disorder remains a common practice. This study aimed to test the effectiveness of blended care for discontinuing long-term BZRA use in general practice.

Methods

A pragmatic cluster randomized controlled superiority trial compared blended care to usual care through urine toxicology screening. In the intervention, care by the general practitioner (GP) was complemented by an interactive e-learning program, based on cognitive behavioral therapy for insomnia. Adults using BZRA daily for minimally six months were eligible. Participants were clustered at the level of the GP surgery for allocation (1:1). Effectiveness was measured as the proportion of patients who had discontinued at one-year follow-up. Data analysis followed intention-to-treat principles.

Results

In total, 916 patients in 86 clusters, represented by 99 GPs, were randomized. Primary outcome data was obtained from 727 patients (79%). At one-year follow-up, 82 patients (18%) in blended care, compared to 91 patients (20%) in usual care, had discontinued. There was no statistically significant effect for the intervention (OR: 0·924; 95% CI: 0·60, 1·43). No adverse events were reported to the research team.

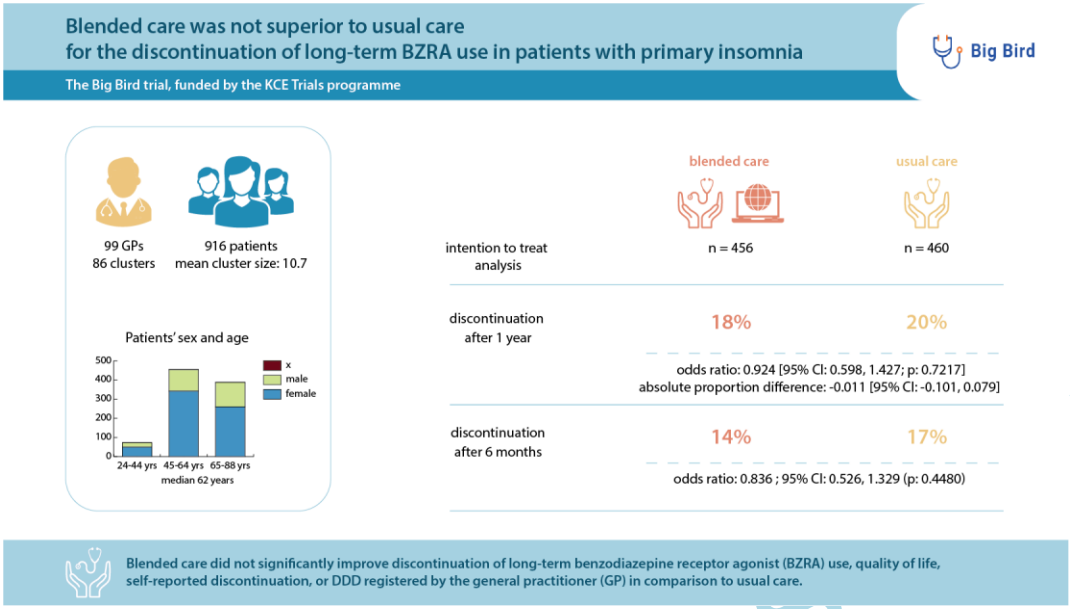
Conclusions

The findings did not support the superiority of blended care over usual care. Both strategies showed clinical effectiveness, with an average of 19% of patients having discontinued at one-year follow-up. Further research is important to study the effect of structurally implementing digital interventions in general practice.

Keywords

Benzodiazepines; Sleep Initiation and Maintenance Disorders; Internet-Based Intervention; Cognitive Behavioral Therapy; Drug Tapering; Psychosocial Intervention; Primary Health Care

Graphical abstract



Clinical trial

Big Bird trial; KCE-17016. This trial is registered at clinicaltrials.gov (NCT03937180).

Statement of significance

Research shows that digital CBT-I boosts the non-pharmacological treatment of insomnia, and that brief interventions help discontinue long-term BZRA use. This study brought both approaches together in a blended care intervention, for the first time, using a pragmatic setup in general practice. To test its effectiveness for discontinuing long-term BZRA use in patients with chronic insomnia disorder, an active control group was used, and there was no protocol for implementation. Our analyses revealed that blended care was not superior to usual care. Future research should explore the clinical effectiveness of this intervention with recent BZRA users to prevent long-term use, and how to increase the uptake of digital interventions in general practice.

Introduction

Worldwide, benzodiazepine receptor agonists (BZRA), including benzodiazepines and the related z-drugs zolpidem, (es)zopiclone, and zaleplon, are prescribed extensively to treat sleeping disorders and anxiety, or as adjuvant therapy for depression and pain management. European guidelines state that treatment with BZRA should be limited to the lowest possible dose and the shortest possible duration, maximally four weeks [1]. The rationale for these guidelines is twofold. First, there is only weak evidence for long-term BZRA use as tolerance develops over time [1]. Second, BZRA use increases the risk of accidental injuries by impairment of motor and cognitive skills [2–5], and, when using z-drugs, complex sleep behaviors [6].

Moreover, long-term use is associated with serious health problems, like cognitive impairment, increased fall risk, paradoxical symptoms like insomnia and anxiety, drug dependence, depression, and suicidal ideation [2–5,7,8]. Despite the guidelines, and although long-term use may be medically justified for some patients, prescription of a low, steady dosage on a continuous basis for sleeping disorders remains a common practice [9–11]. Because of the many adverse effects, BZRA use should be stabilized or reduced on a large scale to positively affect public health.

Evidence-based discontinuation strategies for long-term BZRA use rely on the effective intervention of gradual tapering [12]. Experimental trials have shown that a combination of dose-tapering and non-pharmacological interventions, such as psychotherapy interventions, self-help instructions, and patient education, is more effective than a stand-alone strategy [13,14].

To offer such combined interventions, blended care is a new and promising approach. It combines an interactive educational e-tool with face-to-face clinical consultations with the care provider [15]. Blended care has already proven to be successful in the treatment of multiple psychiatric and somatic conditions, including sleeping disorders, and substance use disorders [16–18].

In 2015, a small descriptive pilot study confirmed that blended care is effective for discontinuing BZRA use for sleeping disorders, more so than a minimal intervention and equal to face-to-face interventions that combine tapering and patient education [19]. To confirm these findings, a properly powered and controlled study was necessary. A cluster design was chosen to prevent treatment contamination in the control group. This multicenter, pragmatic, cluster randomized, controlled, superiority trial was aimed at establishing an evidence-based blended care approach for the discontinuation of chronic BZRA use in adult patients with chronic insomnia disorder in primary care.

Methods

Study design

This pragmatic, cluster randomized controlled trial (CRCT) was set in general practice in Belgium to test the effectiveness of blended care for the discontinuation of chronic BZRA use for a primary indication of insomnia among community-dwelling adults. Pragmatic trials assess both the applicability of an intervention and its effectiveness [20]. The intervention effect was assessed by the toxicological screening of urine samples at baseline, six months, and twelve months after the start of the intervention. It was a multicenter study, conducted by general practitioners (GPs) in both Dutch and French regions of Belgium. Study conduct at all sites was monitored by both the Sponsor, UZ/KU Leuven, and researchers at the general practice departments of six Belgian universities: KU Leuven, ULiège, UAntwerpen, UGent, VUB, and ULB. The trial was approved by the Ethics Committee for Research of UZ/KU Leuven in March 2019 (ref. S61194).

A more detailed method description is available in the protocol publication [21]. The final version, with substantial amendments due to COVID-19, was published in the trial registry. Changes pertained to the use of telemedicine. Findings are reported following the CONSORT guidelines for pragmatic and cluster trials [22,23].

Patients

Adult patients, minimally 18 years old, who had been prescribed and using BZRA continuously (more than 80% of days) for more than six months as hypnotics, without any severe psychiatric or neurological comorbidities, or other contra-indications for discontinuation, were eligible for participation. Patients were also required to have access to the internet, and show a basic level of digital literacy, operationalized as "using e-mail, and using Google to perform a targeted information search". If they acknowledged having these skills, they were deemed fit for the trial. Substance use disorders, terminal illness, and not having their patient record – or Global Medical File – managed by the GP, as is common practice in Belgium, were reasons for exclusion.

Patients were recruited in a usual care setting by the GPs. Flyers and posters in the waiting room informed patients of the upcoming trial. Most patients were screened during routine appointments. Some GPs screened their patient records for eligible patients and contacted them pro-actively, by e-mail, a phone call, or a discontinuation letter, to discuss their hypnotics' use. Eligible patients were informed of the trial requirements. Those willing to participate signed informed consent with the GP.

Randomization and masking

Randomization occurred at the level of the GP's surgery. Eighty-six clusters were randomized to blended or usual care in a 1:1 ratio using a block randomization system, stratified per language to guarantee that allocation was balanced between Dutch and French communities. Two block sizes were used, four and six GPs, so that the allocation process could not be predicted. This was necessary because of the pragmatic nature of the trial, which made it impossible to uphold masking after randomization.

Surgeries were randomized chronologically, depending on when the GPs had recruited sufficient patients (min. five, max. twenty) to start the trial. Randomization and concealment were centralized at the KU Leuven and conducted by a staff member not involved in data collection or delivering the

intervention. The randomization sequence was created using an electronic random numbers generator, in cooperation with AL, a statistician at the Leuven Biostatistics and Statistical Bioinformatics Centre.

The research team, except the above-mentioned staff member and statistician, was masked to the randomization sequence. To avoid cluster heterogeneity and post-randomization selection bias, patients could no longer be recruited after allocation. The statistician remained masked to the allocation until analyses were complete.

Procedures

GPs were recruited via the network of the six universities involved, mainly during training sessions on usual care when discontinuing hypnotics use. These trainings were an initiative to reach GPs across Belgium. At the end of every training, the trial was presented and GPs could express their interest in participation. Recruited GPs that could not attend these live training sessions, received a self-study package with the presentation and tools. In the control arm, GPs did not receive any additional tools. Their patients received usual care, left at the discretion of the GP. In the intervention arm, an interactive e-learning program was offered to supplement usual care, also referred to as blended care.

The program aimed to empower patients by improving their knowledge, self-efficacy, and confidence concerning discontinuation. Six informational chapters offered the patient an evaluation of their current motivation, linked to a suggestion about where to start in the program, information about types of hypnotics, associated effects and risks of hypnotic use, sleep and sleep hygiene, non-pharmacological alternatives, and tapering of BZRA. (Supplement 1) The sleeping tips, alternative approaches, and coping strategies provided in this program were based on the principles of cognitive behavioral therapy for insomnia (CBT-I) which has been proven to be effective in the management of insomnia, also when offered digitally [24]. Throughout the program, text was alternated with interactive elements, such as quizzes, checklists, self-tests, and cognitive exercises. It also included a

digital sleeping diary with a visual representation of progress, the possibility of adding a personal tapering schedule, and a library with links to additional information.

The program could be viewed by both patients and GPs so that they could discuss progress together. Nevertheless, patients could also decide for each exercise separately to not share their answers with the GP. There were no guidelines on how to deliver or adhere to the intervention. The e-learning program was available in the first six months after randomization.

Patients were followed up for one year. GPs were asked to consult at least once with the patient in the first six months, which is a minimal request, compatible with usual care when prescribing hypnotics. During these first six months, they completed an electronic Case Report Form (eCRF) for all consultations in which they discussed sleep or hypnotic use with the patient. In the eCRF, the following data was collected: BZRA use at six and three months before the trial (completed once at baseline visit), any changes in BZRA use, including product name and defined daily dose (DDD), chronic co-morbidities, discontinuation interventions used by the GP, having discussed any withdrawal symptoms, and whether or not there was a follow-up visit planned.

Patients were invited to complete a self-report questionnaire five times and provide three urine samples. The questionnaire at baseline comprised an evaluation of alcohol use (Audit-C) [25], BZRA dependency (Bendep-SRQ) [26], quality of life (EQ-5D-3L) [27], insomnia severity (ISI) [28], and health literacy (HLS-EU-Q16) [29]. BZRA dependency and health literacy were measured once, as they were not expected to alter during the trial. At weeks 6, 12, 26, and 52, the questionnaire re-evaluated alcohol use, quality of life, and insomnia severity. At this time, patients were also questioned about their use of BZRA and antidepressants, injurious falls, and use of medical services. In June 2020, three questions about the impact of COVID-19 on patients' substance use, and psychosocial symptoms associated with any changes in their use, were added to all subsequent questionnaires. An invitation to complete the questionnaire within two weeks was automatically e-mailed one week in advance. After one week, all participants who did not respond received a reminder. These continued every

week until their response or the deadline. The GP could also contact the patient to help remind them. The deadline was set at four weeks for the questionnaire at weeks 6 and 12, and eight weeks for the questionnaire at weeks 26 and 52. Patients in the intervention group were redirected to the welcome page of the e-learning after having completed their questionnaire, at weeks 6 and 12.

To objectify BZRA use, toxicological screening of urine samples was done. BZRA are detectable up to six days or longer after ingestion of a single dose with liquid chromatography-tandem mass spectrometry (LC-MS/MS). Chronic use for months or years can extend excretion times up to six weeks after discontinuation. LC-MS/MS was the most sensitive method available, which could detect low-dose BZRA use and multiple components in one assay. Toxicological analyses were performed by an external lab, being "Algemeen Medisch Labo" (AML) in Antwerp, Belgium. Because urine analysis is not part of usual care in Belgium, the results of these analyses were not disclosed to the GPs.

Patients provided a urine sample at baseline, six months, and one year. The first was requested during the baseline visit with the GP. For the other samples, there was no obligation to consult the GP. However, patients were requested to produce their urine samples at the surgery to limit adulteration. Similar to the questionnaires, an invitation was sent one week in advance of weeks 26 and 52, to deliver the sample within two weeks. Reminders were sent every week until compliance or the deadline, eight weeks later.

Outcomes

The primary outcome was the discontinuation of BZRA use at twelve months, assessed by toxicological urine analysis. Secondary outcomes were: discontinuation of BZRA use at six months, assessed by toxicological urine analysis, quality of life, self-reported discontinuation, and the number of DDD as registered by the GP. All data was processed by the statistician and lead research team at KU Leuven. For safety monitoring, the Belgian standard procedures were followed, with GPs directly reporting adverse effects of medication to the Federal Agency for Medicines and Health Products

(FAMHP). Because no medication or new treatment protocols were tested, this trial was considered low risk. No additional safety monitoring measures were required.

Statistical analysis

The study was powered to detect a statistically significant difference in the primary outcome measure of 10% between both groups, assuming a discontinuation rate of 15% in the control group. A sample size of 756 patients was required with α=0·05 at 80% power, considering an attrition rate of 10%, based on an earlier, similar study [14]. To account for clustering effects, an intracluster correlation coefficient of 0·04 was used. For a cluster size of 10 patients, the number of required patients was multiplied by 1·99, corresponding to the cluster design effect (DE=1+ICC(size of the cluster-1)).

For the urine toxicology test results, logistic regression was used, with a binary outcome (negative/positive) and intervention group as a factor. A random intercept for GP dealt with clustering. The group effect was reported as an odds ratio (OR) with a 95% confidence interval (CI). The remaining secondary outcomes were measured longitudinally. For binary outcomes, multilevel logistic regression analyses were performed, including random intercepts for patient and GP, with intervention effects presented as OR with 95% CIs. Continuous outcomes were analyzed using linear mixed models, similarly modelling random intercepts for patient and GP, with intervention effects reported as the mean difference with 95% CIs. The fixed effects models in all these analyses included the intervention group, time, and group-by-time interaction. In case of a significant group-by-time interaction, the group effect was reported separately for each time point. In case of a non-significant group-by-time interaction, a group main effect was reported. No correction for multiplicity was performed for the secondary analyses, as the study was not powered for these analyses, and hence, its results were considered as hypothesis-generating. The EQ-5D-3L data was summarized based on the principles of a Pareto improvement in Welfare Economics - the Paretian Classification of Health

Change (PCHC) [30]. The EQ-5D-3L scores were analyzed as (1) better versus other, and (2) worse versus other. The registered DDD of all BZRA was converted to milligrams diazepam.

All analyses followed the intention-to-treat (ITT) principle unless otherwise stated. In case the urine analysis data was not collected, the outcome was classified as failure or continued benzodiazepine use, as to not overestimate the treatment effect. Although controversial to some authors, this is a simple method that has previously been deemed acceptable if drop-out rates remain below 20% [31]. For missing data in the questionnaires or eCRFs, there was no imputation done. After withdrawal, no further patient data was collected. Post-hoc, the primary outcome data was descriptively analyzed per-protocol (PP), for which missing data and data collected out of window were excluded. Finally, self-reported discontinuation was descriptively compared to discontinuation as defined by toxicological analysis.

A data monitoring committee reviewed data collection as it was ongoing. This committee also prepared the data for analysis, in cooperation with the statistician. All analyses were done with SAS/STAT® software, version 9·4 of the SAS system for Windows. The trial was registered prospectively with clinicaltrials.gov, NCT03937180.

Role of the funding source

The funder of the study, KCE Trials, had a limited role in the study design. Their scientific board critically reviewed the protocol and commented on the sample size calculations, and timing of the baseline questionnaire. They had no further role in data collection, data analysis, data interpretation, or writing of this report. KCE also has no conflict of interest concerning (non)prescription of BZRA.

Results

In total, 99 GPs, working in 86 surgeries participated. Thirteen GPs ran a solo practice. Sixty-seven GPs worked in a group practice, with twenty of them (30%) working in a multidisciplinary group practice. Nineteen GPs were employed in a community health center. Patients were recruited

between May 23 and December 20, 2019. During this time, 1814 patients were screened, of which 898 were not eligible or declined participation. In total, 916 patients were randomized, with 456 allocated to blended care (n=44 clusters, mean cluster size of 10·4), and 460 to usual care (n=42 clusters, mean cluster size 10·9). The trial ended prematurely for 173 patients, representing 19% of the total sample. Among the remaining 743 patients, primary outcome data was not obtained from sixteen participants. All randomized patients were included in the intention-to-treat analysis. (Figure 1)

Baseline characteristics were comparable across groups (Table 1). The patients' mean age was 61 years (SD 11·5, range 24-88), and most of them were female (71%). At least one co-morbidity was documented at baseline for 66% of all patients. Depression was the most common (27%). Although most patients used one BZRA at baseline, for 14% multiple BZRA were prescribed, with one patient receiving prescriptions for four different BZRA. On average, a DDD of 10·47mg (SD 8·90, range 0-110) diazepam was used. A small proportion of patients had discontinued their BZRA use between enrolment and baseline visit – which was scheduled after randomization -. According to the registration by the GP, this was 3%, while the toxicology results of the baseline samples revealed it was 8% of patients. In total, the baseline sample of 82% of all patients tested positive or contained a trace of BZRA, while the sample of 10% of patients was missing.

With regard to benzodiazepine dependency, an average score of two in the category *problematic use* showed that the patients were aware that their hypnotic use was problematic but did not yet make any changes. Scoring two or higher in the category *preoccupation* revealed that the patients were obsessed with having their hypnotic available to use when they feel they need it. Patients with high scores in *preoccupation* are hard to motivate towards discontinuation as they perceive it necessary support to have the hypnotic available. The scores in the third category, *lack of therapy compliance*, were ambiguous. Having a score higher than one revealed low therapy compliance, which makes it hard to guide patients in their discontinuation process [26]. Although the mean score in the sample

rounded up to one, the median was at 0, suggesting that for at least half of the sample, lack of compliance was not a barrier to discontinuation. The results in both the control and intervention group were similar, resulting in the same clinical interpretation for each category. Interpreting the ISI scores, the current treatment for insomnia appeared to be effective for 17% of participants at baseline. However, a large proportion of the sample (66%) experienced mild to severe symptoms of insomnia. For 31% of participants, additional evaluation of the insomnia severity and treatment would be recommended according to the questionnaire guidelines. Concerning alcohol, up to 40% of patients showed signs of hazardous use. The overall quality of life was rated at 70%. The most, and most severe, problems were reported in the domains *pain* and *anxiety*. Finally, health literacy levels were comparable to the general population in Belgium [32,33]. (Table 1)

For none of the outcomes, a significant group-by-time interaction was found. Therefore, only the main effect of intervention versus control was reported. Analysis of the toxicology data at six months and one year did not reveal a statistically significant intervention effect. At six months, 15% of patients had discontinued. This increased to 19% at one-year follow-up. The OR of 0·924 (95% CI: 0·60; 1·43), and the absolute proportion difference of -0·011 (95% CI: -0·101; 0·079), implied a lower probability of a negative urine test in the intervention group after one year but was not significant (p-value OR=0·7217). (Tables 2 and 3) When limiting the analysis to the patients with BZRA in their sample at baseline (N=751), on average 16% had discontinued their use after one year. (Table 4) For the quality of life, self-reported discontinuation and number of DDD registered by the GP, there was also no statistically significant difference in outcomes between intervention and control. (Table 3)

Post-hoc PP analysis of the primary outcome data resulted in an equal discontinuation rate between groups, namely 24%. Self-reported discontinuation was lower at six and twelve months, and in both groups, than discontinuation shown by toxicological analysis, with an underestimation of 4% to 8%. (Table 5) No adverse events were reported to the research team.

Discussion

This study was the first trial to assess the effectiveness of blended care for the discontinuation of long-term BZRA use for chronic insomnia disorder in general practice. It could not confirm the superiority of blended care over usual care, as in both groups, on average, one in five patients had discontinued their BZRA use by one-year follow-up. Results imply that most of them had already achieved discontinuation at six months. Any differences between groups were not found to be statistically significant.

Our study sample showed to be a realistic patient population, as baseline characteristics were similar to Belgian patients that receive three or more prescriptions for BZRA per year, as documented in Intego, a general practice registration network [34]. Sex and age characteristics and depression rates were comparable across both samples. Hazardous alcohol use was more prevalent in this study sample, 40% versus 9%, which is probably due to underreporting of alcohol use in patients' health records [35]. Both co-morbidities, depressive and alcohol use disorder, contraindicate BZRA use because of their depressogenic and addictive effects, which persist even after tolerance for the anxiolytic and hypnotic effects is developed [8].

The gold standard for treating insomnia is psychological and behavioral therapies [8]. Also digitally delivered CBT-I is efficacious in reducing insomnia symptoms [18]. Two studies reported an additional positive association, with self-reported discontinuation of hypnotic use [24,36]. However, this trial could not confirm that pragmatic implementation of such tools in general practice leads to a statistically significant reduction in the use of hypnotics. Comparing trial specifics reveals two main differences due to the highly pragmatic character of our study [20]. First, the patient sample was meant to be a true reflection of long-term hypnotic users in primary care, so GPs were stimulated to invite all eligible patients chronologically during routine consultations. This is in contrast to the study of Vedaa et al., who explicitly selected participants based on their literacy profile and motivation for behavioral change [24]. Also Kaldo et al. thoroughly screened participants in two phases, including

their motivation and ability to participate [36]. Second, there was no protocol on the implementation of blended care for GPs to follow, nor guidelines on how to use the platform with the patient. Only limited instructions on follow-up (at least once in the first six months after randomization) were provided. GPs and patients determined how they wanted to use the intervention, which resulted in different implementation scenarios across the trial, comparable to the diversity in clinical practice. This is in contrast to most blended care trials in which a timeline and implementation protocol are followed [37–39]. Moreover, to compare the intervention to an active control group, all participating GPs were trained on how to provide optimal usual care for the discontinuation of long-term BZRA use, with brief interventions for daily practice, which have proven effectiveness [12]. This strong control condition could have influenced the results. In 2018, Nesvåg and colleagues showed that just slightly more than half of studies with a control group, to test digital interventions for substance use disorders, had positive results. Effects were described as small to moderate. The less complex the intervention, the more the strength of the control condition influenced the results [16]. Moreover, one of the hypotheses in our trial, raised in discussion between researchers and GPs, was that GPs in the usual care group put in more effort to effectively use brief interventions because their patients were denied access to the intervention in the study.

Although the pragmatic set-up decreased comparability to previous trials, it is considered this study's major strength. Evaluation with the PRECIS-2 [20] showed that the trial was highly pragmatic in all domains. (Supplement 2) The benefit of this approach is that the true effectiveness of the intervention is shown in the current general practice culture in Belgium. Second, our trial meant to map the long-term effects of blended care by following patients until one year after randomization. To verify if discontinuation would be sustained, the online intervention was only available in the first six months. Third, hard outcome measures (toxicological analysis) were used, which ruled out recall bias by patients. Contrary to expectations, self-reported discontinuation seemed to underestimate true discontinuation rates. Finally, a process evaluation, including a survey about patient recruitment, focus groups and interviews about the delivery of and response to the intervention, and

an analysis of the usage data of the e-learning program, was nested within this trial. The results of this qualitative project will be reported separately. This trial also knew limitations. Initially, the ICC was set at 0·11 for the sample size calculation per request of the funder, resulting in a minimum sample of 1182 participants. During review of the protocol paper, it became clear that the correct ICC to be used was 0·04. After careful consideration, the sample size was recalculated and reported accordingly. Next, some sampling bias occurred, despite pragmatic recruitment having been the norm. On the one hand, bias was related to e-literacy, as the need for basic e-literacy skills was inherent to our trial. On the other hand, GPs mentioned in the process evaluation (published elsewhere) that they did not invite all eligible patients because they expected that the patient would refuse participation. However, as Sirdifield and colleagues described, GPs tend to generalize their experience with BZRA withdrawal and not explore the expectations of the patient [40], which is counterproductive for screening in a pragmatic trial. Another limitation was the high drop-out rate, 19%, which impacted the missing data rate for the primary outcome, 21%. Although this is an important aspect of the trial, we do not expect this to substantially have influenced the results of the study. The amount of drop-out could be associated with COVID-19, which interfered with the conduct of the study by increasing the daily tasks of the GPs and raising stress among the patients, as described during patient interviews. An overall decline in mental health, with steep increases in insomnia, anxiety, and depression, implies that the past years have been a difficult period for discontinuing BZRA use [41]. Finally, we should have documented the uptake of the e-learning program in detail, such as patient completion and GP use. This would have provided more insight into the role of online intervention in blended care, as well as its potential mechanisms of action, for the discontinuation of long-term BZRA use.

For future studies, we recommend standardization of the concepts *long-term use* and *long-term follow-up*. For use, a minimal period of six months or longer in one year, as was also suggested by Kurko et al [9]. For follow-up, we would suggest 21 months, as this was found to be a predictor for abstinence after ten years by de Gier et al [42]. Next, we recommend using toxicological assessment

due to the underestimation in self-reporting. LC-MS/MS is deemed to be the most appropriate technique because of the wide detection window. A single DDD can be detected up to one week after ingestion, while chronic use of (z-)BZD can be detected for four to six weeks after cessation. Meaning that even if a patient succeeds to alter their long-term BZRA use for the assessment, it reflects high self-management skills, rather than dependence. Finally, we would recommend to use a non-inferiority design in future trials that assess non-pharmacological discontinuation interventions in general practice because of their possible added value due to the vast diversity among both patients and GPs. In clinical practice, long-term BZRA use is a treatment that often is not tailored to the needs of the patient. This is supported by the high percentage of moderate to severe insomnia symptoms in our study sample, and the high rates of co-morbid depression and hazardous alcohol use. Discontinuing long-term BZRA use can be extra challenging due to these co-morbidities which make behavior change a complex but necessary process. Digital interventions that are tailored to the patients' needs could be part of the solution, as digitalization of care increases. Screening and differential diagnosis could decrease health risks, and interventions aimed at mental health and pain relief could increase the quality of life in this population. Moreover, the rise in insomnia prevalence to 71% and the twofold increase in anxiety and depression compared to 2018, set the stage for policy changes that increase accessibility and implementation of non-pharmacological treatment [41]. Finally, interventions aimed at patient empowerment for better medication and health management, could complement the efforts of GPs to boost non-pharmacological strategies and discontinuation.

In conclusion, blended care was not superior to usual care in discontinuing long-term BZRA use for chronic insomnia disorder in general practice when implemented pragmatically. GPs trained in providing optimal usual care for BZRA discontinuation could achieve a stop with 19% of patients at one-year follow-up. Although previous studies have shown that more discontinuation could be achieved with structured interventions, the implementation process of such interventions, including a potential increase in workload due to training or change management in general practice, needs to be considered.

Acknowledgments

We first want to express gratitude to prof. dr. Catharina Matheï†, who was the driving force behind this project. She inspired and lead the entire team throughout this challenging trial. We would have loved to share the results with her by our side. Next, we are grateful to all general practitioners and patients who have participated in the trial. Also, thank you to our research partners, being An De Sutter from UGent, Sibyl Anthierens and Kris Van den Broeck from UAntwerp, Marc Vanmeerbeek and Gilles Henrard from Université de Liège, Nadine Kacenelenbogen and Anne-Marie Offermans from Université Libre de Bruxelles and Dirk Devroey from VUB, and to our colleagues at the Academic Center for General Practice at KU Leuven for supporting this project. Special thanks to the Clinical Trial Center at UZ Leuven for ensuring that the study was conducted in accordance with ICH-GCP. And finally, we highly appreciate the feedback of Jillian Harrison and France Vrijens from KCE with regard to the clinical study report.

Contributors

Catharina Matheï and MVN conceived the study. All authors, except BS, and the independent funder KCE contributed to study design. AL did all analyses masked to group allocation. KC produced the first draft of the manuscript with input from BS and MVN. All authors assisted in drafting of the final, submitted version of the manuscript and all authors have approved this version.

Data sharing statement

All individual participant data (IPD) that underlie the results in this Article will be made available upon request to the corresponding author, starting six months after publication of this Article. The IPD set will consist of deidentified participant data. Reuse of the IPD set is only allowed for non-profit research, with correct reference to the original research and Sponsor. For

any requests later than September 2023, follow-up by the Sponsor cannot be guaranteed. To ensure compliance with GDPR, data processing must be covered by the European Commission's standard contractual clauses for the transfer of personal data, which must be signed by the data requesters.

Financial disclosure

This study (KCE-17016) is independent research (KCE-17016) which was funded by Belgian Health Care Knowledge Centre under the KCE Trials Programme. The views expressed in this publication are those of the author(s) and not necessarily those of Belgian Health Care Knowledge Centre. The funder had no role in writing this Article. KC and BS also report having received project funding by EBPracticenet for a participatory project on insomnia, which took place parallel to the study reported in this Article.

Non-financial disclosure

GH, BS, ADS, SA, DD, NK, and MVN are practicing general practitioners. There are no other potentially competing interests to disclose.

References

1. Riemann D, Baglioni C, Bassetti C, et al. European guideline for the diagnosis and treatment of insomnia. *J Sleep Res*. 2017;26(6):675-700. doi:10.1111/JSR.12594

2. Zhong GC, Wang Y, Zhang Y, Zhao Y. Association between benzodiazepine use and dementia: A meta-analysis. *PLoS One*. 2015;10(5):1-16. doi:10.1371/journal.pone.0127836

3. Glass J, Lanctot KL, Herrmann N, Sproule BA, Busto UE. Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits. *BMJ-BRITISH Med J*. 2005;331(7526):1169-1173.

4. Griffin CE, Kaye AM, Rivera Bueno F, Kaye AD. Benzodiazepine pharmacology and central nervous system-mediated effects. *Ochsner J*. 2013;13(2):214-223.

5. Jung ME, Metzger DB, Hall J. The long-term but not short-term use of benzodiazepine impairs motoric function and upregulates amyloid β in part through the suppression of translocator protein. *Pharmacol Biochem Behav*. 2020;191(January):172873. doi:10.1016/j.pbb.2020.172873

6. FDA. FDA adds Boxed Warning for risk of serious injuries caused by sleepwalking with certain prescription insomnia medicines. Drug Safety Communications. Published 2019. Accessed April 15, 2022. https://www.fda.gov/drugs/drug-safety-and-availability/fda-adds-boxed-warning-risk-serious-injuries-caused-sleepwalking-certain-prescription-insomnia

7. Ashton H. Drug dependency: benzodiazepines. In: Ayers S, Baum A, McManus C, et al., eds. *Cambridge Handbook of Psychology, Health and Medicine.* 2nd editio. Cambridge University Press; 2007:675-678.

8. Guina J, Merrill B. Benzodiazepines I: Upping the Care on Downers: The Evidence of Risks, Benefits and Alternatives. *J Clin Med 2018, Vol 7, Page 17*. 2018;7(2):17.

doi:10.3390/JCM7020017

9. Kurko TAT, Saastamoinen LK, Tähkäpää S, et al. Long-term use of benzodiazepines: Definitions, prevalence and usage patterns - A systematic review of register-based studies. *Eur Psychiatry*. 2015;30(8):1037-1047. doi:10.1016/j.eurpsy.2015.09.003

10. Weymann D, Gladstone EJ, Smolina K, Morgan SG. Long-term sedative use among community-dwelling adults: a population-based analysis. *C open*. 2017;5(1):E52-E60. doi:10.9778/CMAJO.20160056

11. Soyka M. Treatment of Benzodiazepine Dependence. *N Engl J Med*. 2017;376(12):1147-1157.

12. Lynch T, Ryan C, Hughes CM, et al. Brief interventions targeting long-term benzodiazepine and Z-drug use in primary care: a systematic review and meta-analysis. *Addiction*. 2020;115(9):1618-1639. doi:10.1111/add.14981

13. Tannenbaum C, Martin P, Tamblyn R, Benedetti A, Ahmed S. Reduction of inappropriate benzodiazepine prescriptions among older adults through direct patient education: the EMPOWER cluster randomized trial. *JAMA Intern Med*. 2014;174(6 CC-Effective Practice and Organisation of Care):890-898. doi:10.1001/jamainternmed.2014.949

14. Vicens C, Bejarano F, Sempere E, et al. Comparative efficacy of two interventions to discontinue long-term benzodiazepine use: cluster randomised controlled trial in primary care. *Br J psychiatry*. 2014;204(6 CC-Common Mental Disorders):471-479. doi:10.1192/bjp.bp.113.134650

15. Wentzel J, Van der Vaart R, Bohlmeijer ET, Van Gemert-Pijnen JEWC. Mixing Online and Face-to-Face Therapy: How to Benefit From Blended Care in Mental Health Care. *JMIR Ment Heal 2016;3(1)e9 https//mental.jmir.org/2016/1/e9*. 2016;3(1):e4534. doi:10.2196/MENTAL.4534

16. Nesvåg S, McKay JR. Feasibility and Effects of Digital Interventions to Support People in

Recovery From Substance Use Disorders: Systematic Review. *J Med Internet Res 2018;20(8)e255 https//www.jmir.org/2018/8/e255*. 2018;20(8):e9873. doi:10.2196/JMIR.9873

17. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. *Cogn Behav Ther*. 2018;47(1):1-18. doi:10.1080/16506073.2017.1401115

18. Hasan F, Tu YK, Yang CM, et al. Comparative efficacy of digital cognitive behavioral therapy for insomnia: A systematic review and network meta-analysis. *Sleep Med Rev*. 2022;61:101567. doi:10.1016/j.smrv.2021.101567

19. Klaasen G, Helsen C, Klaasen B, Van Nuland M. Afbouwen van chronisch benzodiazepinegebruik. *Huisarts Nu*. 2015;44(5):186-191. doi:10.1007/s40954-015-0079-4

20. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350. doi:10.1136/BMJ.H2147

21. Coteur K, Van Nuland M, Vanmeerbeek M, et al. Effectiveness of a blended care programme for the discontinuation of benzodiazepine use for sleeping problems in primary care: Study protocol of a cluster randomised trial, the Big Bird trial. *BMJ Open*. 2020;10(2):1-8. doi:10.1136/bmjopen-2019-033688

22. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 2008;337(7680):1223-1226. doi:10.1136/BMJ.A2390

23. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345(7881). doi:10.1136/BMJ.E5661

24. Vedaa Ø, Kallestad H, Scott J, et al. Effects of digital cognitive behavioural therapy for insomnia on insomnia severity: a large-scale randomised controlled trial. *Lancet Digit Heal*. 2020;2(8):e397-e406. doi:10.1016/S2589-7500(20)30135-7

25. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care. Published online 2001.

26. Kan CC, Breteler MHM, Timmermans EAY, Van Der Ven AHGS, Zitman FG. Scalability, reliability, and validity of the benzodiazepine dependence self-report questionnaire in outpatient benzodiazepine users. *Compr Psychiatry*. 1999;40(4):283-291. doi:10.1016/S0010-440X(99)90129-3

27. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1990;16(3):199-208. doi:10.1016/0168-8510(90)90421-9

28. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2(4):297-307. doi:10.1016/S1389-9457(00)00065-4

29. Sørensen K, Van Den Broucke S, Pelikan JM, et al. Measuring health literacy in populations: Illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BMC Public Health*. 2013;13(1):1-10. doi:10.1186/1471-2458-13-948/TABLES/4

30. Devlin N, Parkin D, Janssen B. Analysis of EQ-5D Profiles. *Methods Anal Report EQ-5D Data*. Published online 2020:23-49. doi:10.1007/978-3-030-47622-9_2

31. Armijo-Olivo S, Warren S, Magee D. Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: a review. *Phys Ther Rev*. 2009;14(1):36-49. doi:10.1179/174328809X405928

32. Charafeddine R, Demarest S, Berete F. *Gezondheidsenquête 2018: Gezondheidsvaardigheden. [Health Interview Survey 2018: Health Literacy.]*.

33. Avalosse H, Verniest R, Vancorenland S, De Cock S, Gérard F, et al. *Littératie En Santé et Sources d'information. [Health Literacy and Information Sources.]* .; 2017.

34. Coteur K, Mamouris P, Vaes B, Van Nuland M, Matheï C, Schoenmakers B. Evolution of benzodiazepine receptor agonist prescriptions in general practice: A registry-based study. *Front public Heal*. 2022;10. doi:10.3389/FPUBH.2022.1014734

35. Abidi L, Oenema A, van den Akker M, van de Mheen D. Do general practitioners record alcohol abuse in the electronic medical records? A comparison of survey and medical record data. *https://doi.org/101080/03007995201814246 23*. 2018;34(3):567-572. doi:10.1080/03007995.2018.1424623

36. Kaldo V, Jernelöv S, Blom K, et al. Guided internet cognitive behavioral therapy for insomnia compared to a control treatment - A randomized trial. *Behav Res Ther*. 2015;71:90-100. doi:10.1016/j.brat.2015.06.001

37. van der Zweerde T, Lancee J, Slottje P, et al. Cost-effectiveness of i-Sleep, a guided online CBT intervention, for patients with insomnia in general practice: Protocol of a pragmatic randomized controlled trial. *BMC Psychiatry*. 2016;16(1):1-10. doi:10.1186/S12888-016-0783-Z/TABLES/2

38. Branquinho M, Canavarro MC, Fonseca A. A Blended Cognitive–Behavioral Intervention for the Treatment of Postpartum Depression: Study Protocol for a Randomized Controlled Trial. *Int J Environ Res Public Heal 2020, Vol 17, Page 8631*. 2020;17(22):8631. doi:10.3390/IJERPH17228631

39. Ford ME, Geurtsen GJ, Groet E, et al. A blended eHealth intervention for insomnia following

acquired brain injury: a randomised controlled trial. *J Sleep Res*. Published online May 31, 2022:e13629. doi:10.1111/JSR.13629

40. Sirdifield C, Anthierens S, Creupelandt H, Chipchase SY, Christiaens T, Siriwardena AN. General practitioners' experiences and perceptions of benzodiazepine prescribing: systematic review and meta-synthesis. *BMC Fam Pract*. 2013;14:191. doi:10.1186/1471-2296-14-191

41. Demarest Stefaan, Berete Finaba, Braekman Elise, et al. Negende COVID-19-Gezondheidsenquête. Published online 2022.

42. de Gier NAH, Gorgels WJMJ, Lucassen PLBJ, et al. Discontinuation of long-term benzodiazepine use: 10-year follow-up. *Fam Pract*. 2011;28(3):253-259.

Figure 1. Trial profile

Table 1. Baseline characteristics

	Blended care (n=456)	Usual care (n=460)
Age, years	61·72 (11·57)	61·10 (11·44)
Missing data	2 (<1%)	0
Gender		
Female	329 (72%)	324 (70%)
Male	124 (27%)	135 (29%)
X	3 (<1%)	1 (<1%)
Missing data	0	0
Co-morbidity		
None registered	98 (21%)	142 (31%)
Depression	115 (25%)	135 (29%)
Other	201 (44%)	151 (33%)
Missing data	42 (9%)	32 (7%)
# BZRA used		
0 BZRA	14 (3%)	18 (4%)
1 BZRA	339 (74%)	336 (73%)
2 BZRA	53 (12%)	65 (14%)
3 BZRA	5 (1%)	8 (2%)
4 BZRA	1 (<1%)	0
Missing data	44 (10%)	33 (7%)
Defined Daily Dose		
Mg diazepam	10·10 (8·95)	10·83 (8·84)
Range	0-110	0-70
Missing data	44 (10%)	33 (7%)
Bendep-SRQ*		
Problematic use	1·92 (1·43)	1·87 (1·41)
Preoccupation	2·69 (1·43)	2·61 (1·35)
Lack of compliance	0·77 (1·11)	0·82 (1·17)
Missing data	83 (18%)	67 (15%)
ISI*		
No insomnia	78 (17%)	80 (17%)
Mild insomnia	153 (34%)	173 (38%)
Moderate insomnia	120 (26%)	126 (27%)
Severe insomnia	21 (5%)	14 (3%)
Missing data	84 (18%)	67 (15%)
Audit-C*		
Positive	189 (41%)	181 (39%)
Negative	183 (40%)	211 (46%)
Missing data**	84 (18%)	68 (15%)
HLS-EU-Q16*		
Adequate	212 (47%)	221 (48%)
Problematic	71 (16%)	83 (18%)
Inadequate	32 (7%)	31 (7%)
Missing data***	141 (31%)	125 (27%)
EQ-5D-3L*		
Any problems with:		
Mobility	82 (18%)	87 (19%)
Self-care	21 (5%)	19 (4%)
Activity	99 (22%)	107 (23%)

Pain	286 (63%)	297 (65%)
Anxiety	225 (49%)	202 (44%)
VAS-score	69·57 (16·08)	69·92 (16·82)
Missing data	82 (18%)	67 (15%)
Toxicology urine		
Positive	361 (79%)	362 (79%)
Out of window****	54 (15%)	68 (19%)
Trace	11 (2%)	17 (4%)
Out of window****	0	0
Negative	31 (7%)	43 (9%)
Out of window****	6 (19%)	16 (37%)
Missing data	53 (12%)	38 (8%)

Data presented as mean (SD) or n (%). *At baseline, 69 questionnaires were completed out of window and therefore excluded from analysis. For accurate representation, these were included in the missing data numbers. **767 patients completed the Audit-C but three scores could not be interpreted because the gender of the patient was X. These were included in the missing data numbers. Two out of three scored three points (one in blended care, one in usual care), which is the cut-off value to consider the test as positive for women. ***115 questionnaires could not be considered valid according to the guidelines for analysis, because they contained less than fourteen out of sixteen answers. These were included in the missing data numbers. ****Percentage calculated based on number of collected urine samples.

Table 2. ITT results

	Blended care (n=456)	Usual care (n=460)
Urine toxicology results - one year		
Positive	250 (55%)	263 (57%)
Out of window*	5 (2%)	9 (2%)
Trace	17 (4%)	24 (5%)
Out of window*	0	0
Negative	82 (18%)	91 (20%)
Out of window*	1 (1%)	3 (3%)
Missing data	107 (24%)	82 (18%)
Urine toxicology results – six months		
Positive	292 (64%)	293 (64%)
Out of window*	25 (9%)	28 (10%)
Trace	9 (2%)	10 (2%)
Out of window*	0	1 (10%)
Negative	62 (14%)	76 (17%)
Out of window*	5 (8%)	8 (11%)
Missing data	93 (20%)	81 (18%)

Data presented as mean (SD) or n (%). *Percentage calculated based on number of collected urine samples in this category (positive, trace, or negative).

Table 3. ORs of primary and secondary outcomes from baseline to one-year follow-up

Main effect: blended care VS usual care	OR (95% CI)	P-value
One-year toxicology*	0·924 (0·598;1·427)	0·7217
Six months toxicology	0·836 (0·526;1·329)	0·4480
Quality of life improvement	1·450 (0·942;2·231)	0·0911
Quality of life worsening	0·694 (0·411;1·173)	0·1724
Self-reported discontinuation	1·264 (0·569;2·809)	0·5647
	Mean difference (95% CI)	P-value
DDD registered by the GP**	-0·820 (-2·421;0·782)***	0·3156

OR: odds ratio, CI: confidence interval

OR<1: lower probability of negative urine test for blended care compared to usual care

*The absolute proportion difference for the primary outcome was -0.011 (95% CI: -0.101;0.079).

**From baseline to last follow-up, maximally at six months.

***Mean difference<0: lower prescribed DDD in blended care compared to usual care.

Table 4. Sample with BZRA present at baseline (N=751)

Urine toxicology results – one year	Blended care (n=372)	Usual care (n=379)
Ended prematurely	38 (10%)	36 (10%)
Positive/trace	263 (71%)	276 (73%)
Negative	60 (16%)	63 (17%)
Missing data	11 (3%)	4 (1%)

Table 5. Discontinuation

		One-year follow-up		Six months follow-up	
		Blended care (n=456)	Usual care (n=460)	Blended care (n=456)	Usual care (n=460)
Toxicology results	ITT PP	82/456 (18%) 81/343 (24%)	91/460 (20%) 88/366 (24%)	62/456 (14%) 57/333 (17%)	76/460 (17%) 68/342 (20%)
Self-reported discontinuation	ITT PP	64/456 (14%) 64/339 (19%)	55/460 (12%) 55/359 (15%)	46/456 (10%) 46/331 (14%)	42/460 (9%) 42/365 (12%)

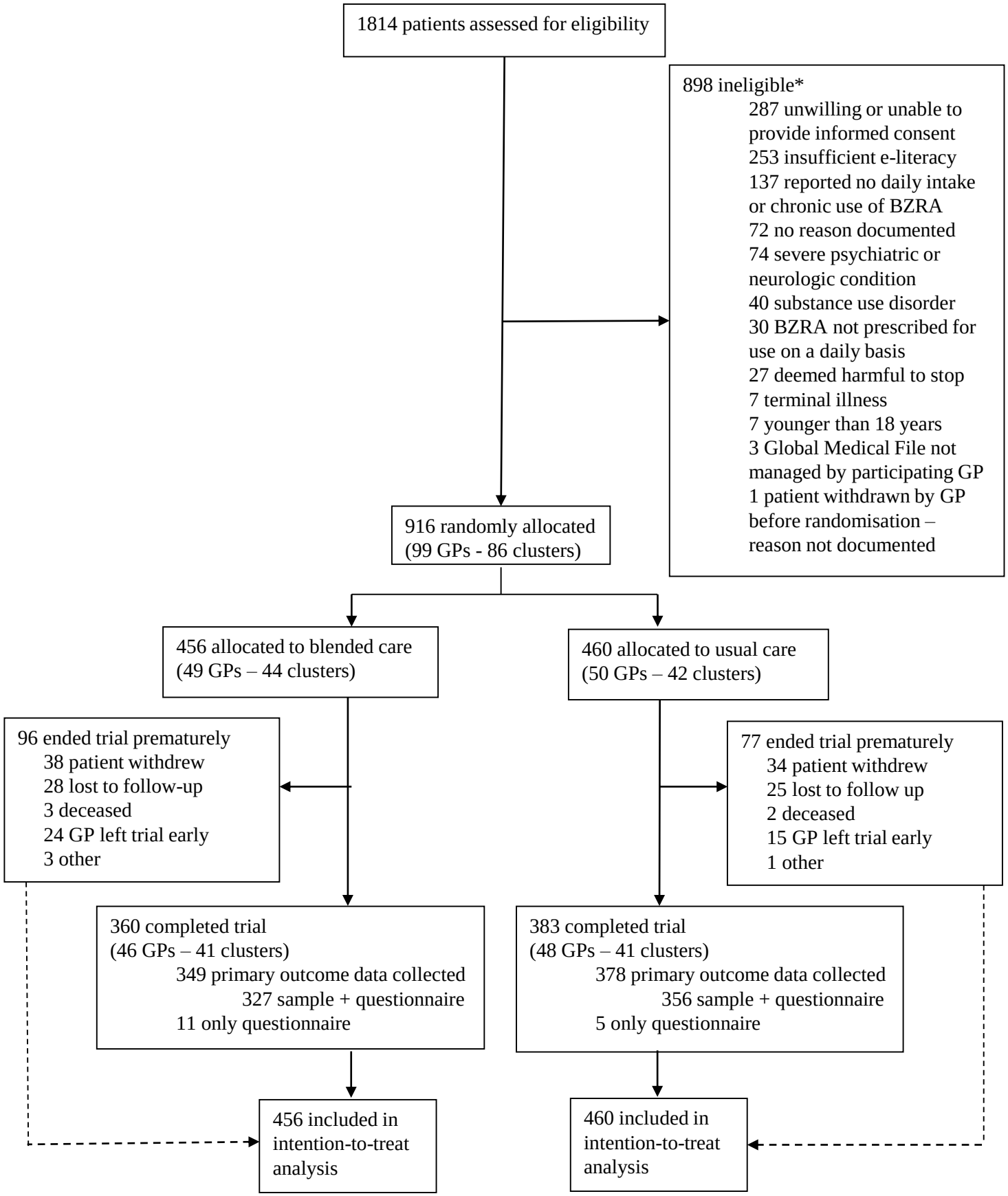


Figure 1: Trial profile
*Patients could be deemed ineligible based on more than one in- or exclusion criteria.
Primary outcome data = urine sample at week 52