Effect of a yoghurt drink containing Lactobacillus strains on bacterial vaginosis in women – a double-blind, randomised, controlled clinical pilot trial

C. Laue¹, E. Papazova¹, A. Liesegang¹, A. Pannenbeckers², P. Arendarski², B. Linnerth³, K.J. Domig⁴, W. Kneifel⁴, L. Petricevic⁵ and J. Schrezenmeir¹,⁶

¹Clinical Research Center, Schauenburgerstraße 116, 24118 Kiel, Germany; ²LADR, Lauenburger Str. 67, 21502 Geesthacht, Germany; ³Dairyfem, Tuchlauben 18/12, 1010 Vienna, Austria; ⁴Department of Food Science and Technology, BOKU – University of Natural Resources and Life Sciences, Muthgasse 18, A-1190 Vienna, Austria; ⁵Department of Obstetrics and Fetal-Maternal Medicine, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria; ⁶University Medicine, Langenbeckstraße 1, 55131 Mainz, Germany; j.schrezenmeir@crc-kiel.de

Received: 8 February 2017 / Accepted: 5 August 2017 © 2017 Wageningen Academic Publishers

Abstract

Bacterial vaginosis (BV) is characterised by a depletion of lactobacilli in favour of an overgrowth of anaerobic bacteria. It is associated with increased risk for urogenital infections and abortion. In this study we assessed the effect of a yoghurt drink containing Lactobacillus strains on BV. The strains had been isolated from healthy pregnant women and selected for acidification capacity, production of H₂O₂, glycogen utilisation, bile salt tolerance and inhibition of pathogens. Using Amsel criteria BV was diagnosed in 36 women aged ≥18 years with stable menstrual cycle or menopause. They were treated with oral metronidazole for 7 days (2×500 mg/d). Starting with the treatment, women consumed twice daily either verum or placebo during 4 weeks. Verum was 125 g yoghurt containing (besides Lactobacillus delbrueckii ssp. bulgaricus and Streptococcus thermophilus) living strains Lactobacillus crispatus LbV 88 (DSM 22566), Lactobacillus gasseri LbV 150N (DSM 22583), Lactobacillus jensenii LbV 116 (DSM 22567) and Lactobacillus rhamnosus LbV96 (DSM 22560), each 1×10⁷ cfu/ml; placebo was 125 g chemically acidified milk. After 4 weeks of intervention 0 of 17 had BV in the verum group versus 6 of 17 in the s.a. control (0.018 in Fisher Exact test). Amsel score decreased during the intervention period by 4.0 (median) (4.0; 3.0) (25th; 75th percentile) in the verum group compared to 2.0 (4.0; 0.0) in the control group (P=0.038 in Mann-Whitney test). Discharge and odour (Amsel criteria 2+3) also decreased by 2.0 (2.0; 1.0) in the verum compared to 1.0 (2.0; 0.0) in the control group (P=0.01) and differed after 4 weeks intervention between the groups 0.0 (0.0; 0.0) versus 1.0 (0.0; 2.0) (P=0.001). Nugent score decreased during the intervention period by 5.5 (7.0;2.3) in the verum compared to 3.0 (6.0;0.5) in the control group (P=0.158). Additional intake of yoghurt containing these probiotic strains improved the recovery rate and symptoms of BV and tended to improve the vaginal microbial pattern.

Keywords: probiotics, lactobacilli, vaginosis, microbiota

1. Introduction

Adult women have normally a vaginal pH of 3.8 to 4.4 (Hauth et al., 2003; Mendling et al., 2013; Platz-Christensen et al., 1993). This is maintained by the vaginal microbiota predominantly consisting of lactobacilli and of transient and commensal anaerobic and aerobic bacteria and Candida species common with skin and intestinal microbiota. Most women in the USA were colonised by Lactobacillus crispatus (32%), followed by Lactobacillus jensenii (23%), a previously undescribed species designated Lactobacillus 1086V (15%), Lactobacillus gasseri (5%), Lactobacillus fermentum (0.3%), Lactobacillus oris (0.3%), Lactobacillus reuteri (0.3%), Lactobacillus ruminis (0.3%), and Lactobacillus vaginalis (0.3%) (Antonio et al., 1999). In Japanese women, L. crispatus was the predominant vaginal Lactobacillus, followed by L. gasseri, (Song et al., 1999). These findings were corroborated using various techniques of detection (Fredricks, 2011; Fredricks et al.,...
Colonisation by L. crispatus or L. jensenii was associated with lower frequency of BV, whereas Lactobacillus iners, Gardnerella vaginalis, Clostridiales, Leptotrichia/Sneathia, Atopobium vaginae, Megasphaera, Eggerthella, Aerococcus, Leptotrichia/Sneathia, Prevotella and Papillibacter were detected in most subjects with BV (Antonio et al., 1999; Arouetcheva et al., 2001; Fredricks, 2011; Fredricks et al., 2005, 2007; Hyman et al., 2005; Ling et al., 2010; Oakley et al. 2008; Ravel et al., 2011; Srinivasan et al., 2012; Verhelst et al., 2004). BV is the most common microbiological disturbance of the vaginal environment in adult women.

BV is characterised by a depletion of lactobacilli in favour of a wide overgrowth of anaerobic bacteria (Fredricks, 2011; Fredricks et al., 2005; Verhelst et al., 2004). Different bacterial species showed different associations with the four clinical Amsel criteria of BV (Srinivasan et al., 2012): Eggerthella and Lactobacillus amnionii were the only BV-associated bacteria that were positively associated with all four clinical criteria (pH>4.5, thin, homogeneous discharge, amine odour in whiff test, and clue cells). L. crispatus was strongly correlated with the absence of BV and was the only Lactobacillus species associated with low pH, negative whiff test, absence of clue cells and normal vaginal discharge. G. vaginalis and A. vaginae were each associated with 3 criteria. G. vaginalis was not associated with abnormal vaginal odour, while A. vaginae was not associated with amine odour.

The significant reduction in vaginal lactobacilli is accompanied by a decrease of H$_2$O$_2$ and of lactic acid, typically produced by lactobacilli in varying extent, and by an increase in pH above 4.5. These factors promote infections (Nardis et al., 2013; O’Hanlon et al., 2013). Indeed BV increases the risk for ascending gynaecological diseases (e.g. endometritis, salpingitis, tuboovarian abscess) (Hillier et al., 1996; Klebanoff et al., 2004; Larsson et al., 1990), sexually transmitted infections (Brotman et al., 2010; Cherpes et al., 2005; Cone, 2014; Ling et al., 2011; Myer et al., 2005; Nardis et al., 2013; Watts et al., 2005) and obstetric complications (e.g. miscarriage, preterm rupture of membranes, preterm birth, endometritis post-partum) (Donders et al., 2000; Hauth et al., 2003; Hillier et al., 1996; Leitch et al., 2003; Macones et al., 2004; Petricevic et al., 2014; Simhan et al., 2005; Swidsinski et al., 2013). Furthermore, the risk of urinary tract infections was reported to be increased (Harmanli et al., 2000; Hillebrand et al., 2002; Klebanoff et al., 2004; Larsson et al., 1990).

Although antibiotic treatment of BV is strongly recommended (Anonymous, 2012; Koumans et al., 2001; Mendling et al., 2013), attempts for improving efficacy of therapy are still a matter of debate. Gardnerella and other bacteria adherent to the vaginal epithelium are well protected against the access of antibiotics by biofilms, which are not eradicated using the recommended standard therapy (Swidsinski et al., 2008). This is regarded to be reason for the moderate cure rates and the high recurrence rate after antibiotic treatment. The cure rate after 3 months is 60-70% and much lower after 6 months (Larsson and Forsum, 2005; Verstraalen and Verhelst, 2009).

To overcome this challenge, recent scientific research has focused on the adjuvant application of probiotics (lactobacilli) in BV treatment. This concept arose from the fact that humans are inhabited with microbiota from birth and that these organisms play a role in the defence against pathogens. Probiotics are defined as ‘a preparation of or a product containing viable, defined microorganisms in sufficient numbers, which alter the microbiota (by implantation or colonisation) in a compartment of the host and by that exert beneficial health effects in this host’ (Schrezenmeir and De Vrese, 2001), and probiotic strains have already been shown in meta-analyses to prevent antibiotics associated and infectious diarrhoea, respiratory tract infections, necrotising enterocolitis and atopic dermatitis, and alleviate IBS, ulcerative colitis and pouchitis (Allen et al., 2010; Elazab et al., 2013; Goldenberg et al., 2013; Hao et al., 2015; Ritchie and Romanuk, 2012; Szajewska et al., 2016).

Oral administration of probiotic strains (Lactobacillus rhamnosus GR-1 and L. reuteri RC-14) enabled alteration of vaginal microbiota in healthy women (Reid and Bocking, 2003). In a double-blind, randomised controlled clinical trial (DB-RCT) oral administration of these strains augmented the efficacy of a 7 days standard therapy with metronidazole in Nigerian women with BV. In the probiotic group the cure rate as assessed by Nugent score was significantly higher (88%, P<0.001) after 30 days of consumption compared to the placebo group (40% BV-free women) (Anukam et al., 2006). These results are in line with data of Martinez et al. (2009). They administered the same strains orally for 28 days combined with a single dose of tinidazole (2 g) to Brazilian women with BV. Using Nugent and Amsel scoring they found a cure rate of 87.5% in the probiotic group compared to 50% in the placebo group (P<0.05) (Martinez et al., 2009). Austrian postmenopausal women with intermediate Nugent score were treated for 14 days. Women in the intervention group received probiotic capsules containing 2.5×10$^9$ cfu each of lyophilised L. rhamnosus GR-1 and L. reuteri RC-14. Women in the control group received an oral placebo once daily. Sixty percent in the intervention group and 16% in the control group showed a reduction in the Nugent score by at least two grades. The difference in the number of patients with improvement was highly significant (P=0.0001) (Petricevic et al., 2008).
The strains used in these trials, however, are not belonging to the species being most abundant in the vaginal microbiota of most healthy women and their suitability for application in a yoghurt matrix is uncertain.

The Lactobacillus strains used in this study were isolated from healthy pregnant women in the late first trimester of pregnancy (Kiss et al., 2007) and were selected out of the Lactobacillus species dominating the vaginal microbiota in healthy women. Out of a pool of 168 presumptive lactobacilli pure isolates were cultured followed by phenotypic and genotypic tests (including Gram-staining, catalase and oxidase activity, growth under aerobic and anaerobic atmospheres, acidification capacity, production of extracellular hydrogen peroxide, glycosidase and arylamidase activity). Finally, 4 strains were selected: L. crispatus LBV 88 (DSM 22566), L. gasseri LBV 150N (DSM 22583), L. jensenii LBV 116 (DSM 22567) and L. rhamnosus LBV96 (DSM 22560) (Domig et al., 2014). These strains already have been tested in a RCT in 60 male to female transsexual women with penile linked neovagina, which is regarded as a model for dysbiosis (Weyers et al., 2009). Oral supplementation with these strains resulted in an improvement of neo-vaginal microbial pattern towards lactobacilli microbiota (Kaufmann et al., 2014). In an open randomised controlled trial in 60 pregnant women with herpes virus infection these strains together with 84 mg fructooligosaccharides were orally administered twice daily for 7 days (Anoshina, 2016). Compared to the control group without supplementation a lower proportion of women with abnormal vaginal pH and with positive amine test of vaginal discharge and a lower proportion of placental insufficiency (23.3 vs 40%) and foetal distress (16.7 vs 36.7%) were found after the 7 days intervention. Based on the clinical studies showing an effect on alteration of vaginal or neo-vaginal microbiota and cure rates of BV by oral administration of probiotics, respectively (Anoshina, 2016; Anukam et al., 2006; Kaufmann et al., 2014; Martinez et al., 2009; Petricic and Witt, 2008; Reid and Bocking, 2003) and based on the particular properties of the strains L. crispatus LBV 88, L. gasseri LBV 150N, L. jensenii LBV 116 and L. rhamnosus LBV96 (Domig et al., 2014) we investigated whether the oral intake of yoghurt drink containing these four probiotic strains improves the cure rate for BV and the symptoms of BV, such as vaginal discharge or fishy odour. Furthermore, we examined whether the vaginal microbial pattern may be influenced by the oral intake of yoghurt containing these strains.

2. Materials and methods

Design

The study followed the design of a single-centre, double-blind, placebo-controlled, randomised (1:1) clinical trial with two parallel arms (Figure 1). One group received yoghurt containing living strains of L. crispatus, L. gasseri, L.
rhamnosus, and L. jensenii (verum), the other group received chemically acidified milk without bacterial strains (placebo). The study duration was 4 weeks. During these 4 weeks, subjects were asked for abstaining from fermented dairy products (except the study products), probiotic-containing foods or supplements, OTC medication, and vitamin and mineral supplements. This alimentary restriction was maintained until the end of the study. Placebo treatment consisted of daily 2×125 g chemically (with H₃PO₄) acidified milk without bacterial strains (identical in flavour, colour, texture, appearance, vessel with verum). The product consumption period (4 weeks) lasted from visit 1 (V1) to visit G2 (see below). Over the whole study duration each subject attended 3 scheduled visits at the site (V1, D and V2), whereby D served delivery of test products during intervention. Furthermore there were 2 appointments G1 and G2 with the gynaecologist, the screening examination (G1) for diagnosis of BV (Amsel criteria and Nugent score) and G2 for control of the therapeutic success 4 weeks after starting metronidazole treatment (based on Amsel criteria and Nugent score). At study visit D and V2, subjects had to return any non-used test products, if applicable, to the study site for assessing compliance.

The study was approved from an independent ethics committee (The Ethical Committee of the Medical Council of Schleswig-Holstein, Bad Segeberg, Germany) and was conducted in line with the principles of the current version (2013) of the Declaration of Helsinki (WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects) adopted by the 18th WMA (World Medical Association) General Assembly, Helsinki, Finland, June 1964, and amended by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013, the recommendations for Good Clinical Practice (ICH E6), and in accordance with European and National regulatory requirements.

The study was registered at ClinicalTrials.gov (identifier: NCT02744638).

Subjects

Female volunteers aged ≥18 years at reproductive age with stable menstrual cycles or postmenopausal women presenting with newly diagnosed bacterial vaginosis (episode) were enrolled, if they did not have menses at G1, and if they were complying with a standard oral antibiotic treatment (metronidazole (Arilin®) for 7 days 2×500 mg), willing to consume for 4 weeks the study product two times daily, complying with the dietary restrictions (s. a.) and had given written informed consent. Diagnosis of BV was based on Amsel criteria (Amsel et al., 1983), i.e. at least three of four criteria had to be present: (1) vaginal pH1 above 4.5; (2) thin, homogenous discharge; (3) release of amine (‘fishy’) odour after the addition of 10% KOH (whiff test) to vaginal smear; (4) clue cells on saline wet mount of vaginal smear (in phase contrast microscopy).

Any of the following was regarded as a criterion for exclusion from enrolment into the study: current enrolment in another clinical study; enrolment in another study within the last 4 weeks before inclusion; infection caused by Chlamydia trachomatis, Neisseria gonorrhoea, Trichomonas vaginalis, or Candida albicans or any other vaginal mycosis; leucocytes present in the vaginal smear; history of PAP-testing ≥3; vulvovaginal inflammation identified macroscopically by the gynaecologist; dyspareunia; pregnancy and breastfeeding; chemically based contraceptives (e.g. suppositories, salves, foam, gel); irregular cycles (e.g. polymenorrhoea, metrorrhagia); dysuria; infection of the urinary tract; chronic or sporadic abdominal pain with exception of dysmenorrhoea; any anal or rectal infection, disease, surgery in the medical history or current anus praeter; hypersensitivity; allergy or idiosyncratic reaction to metronidazole or to any similar active substances; hypersensitivity, allergy or idiosyncratic reaction to any component of the yoghurt (e.g. lactose intolerance, allergy against milk protein); any disease or condition which might compromise significantly the hematopoietic, renal, endocrine, pulmonary, hepatic, gastrointestinal, cardiovascular, immunological, central nervous, dermatological or any other body system with the exception of the conditions defined by the inclusion criteria; history of hepatitis B and C, or of HIV infection; regular medical treatment including OTC, which may have impact on the study aims (e.g. probiotics, antibiotic drugs, laxatives, etc.); major cognitive or psychiatric disorders; subjects who were scheduled to undergo hospitalisation during the study period; eating disorders (e.g. anorexia, bulimia) or special diets (e.g. vegan, vegetarian); present or previous drug abuse or alcoholism; legal incapacity.

The study population was recruited in the screening examination (G1) by gynaecologists in the region of Schleswig-Holstein. At this screening examination BV was diagnosed based on Amsel criteria (Amsel, 1983). The eligible volunteers received brief information about the study and had to give written informed consent for screening examinations. The gynaecologist offered to participate in the study and to contact the study site. Interested women were offered an appointment for the visit V1 by telephone call. Prior to any study-related procedure, the women were provided by the investigator with detailed information on the study including benefits and risks. After enrolment, subjects were randomly assigned to one of the two test groups: verum or placebo. Randomised assignment started in February 2015 and was planned to be terminated end of July 2016 irrespective of the number of included subjects.
The study participants were free to withdraw from the study at any time without prejudice to their continued care. Specific reasons for discontinuing the study were defined as: safety reasons as judged by the investigator, development of specific exclusion criteria during the study, which have an impact on subject’s safety; incorrect enrolment or randomisation of the subject; wish of the subject to withdraw prematurely from the study; severe non-compliance to protocol as judged by the investigator. Individuals withdrawing or discontinuing prematurely before finishing all study visits were not replaced. For those subjects withdrawing prematurely from the study or had to discontinue due to any reason, adverse events were followed up, questionnaires were completed and diaries returned by the subjects, if possible.

**Intervention**

**Randomisation and allocation concealment**

To avoid selection bias the randomisation scheme was generated by data managers of Tecura GmbH, Kiel, Germany in line with the Cochrane guidelines (Cochrane, 2011) using a software implementation by G.E. Dallal of the pseudo-random number generator of Wichmann and Hill (1982) as modified by McLeod (1985) (www.randomizer.org). The randomisation list was kept confidential at the premises of Tecura GmbH and remained confidential with the exception of those involved in production of the test products (NÖM AG, Baden, Austria) and statistical managers (after the first part of data locking was performed). Code-breaking systems were available in case of occurrence of a serious adverse event for which the medical personnel needed to be aware of what the participant received (verum or placebo product). Raw data were also blinded during the blind review. The code was broken after the database was locked.

**Test products and blinding of participants and personnel**

- **Verum.** A fermented dairy drink was produced using pasteurised whole milk containing 3% (w/v) of milk powder and the yoghurt starter cultures *S. thermophilus* 95% and *L. delbrueckii* subsp. *bulgaricus* 5% (JOINTEC X3 Centro Sperimentale Del Latte, Zalo Buon Persico, Italy). It contained living strains of *L. crispatus, L. gasseri, L. rhamnosus, L. jensenii, L. crispatus* LbV 88 (DSM 22566), *L. gasseri* LbV 150N (DSM 22583), *L. jensenii* LbV 116 (DSM 22567) and *L. rhamnosus* LbV 96 (DSM 22560). The concentration of each strain was adjusted for maintenance to at least $1 \times 10^7$ cfu/ml.

- **Placebo.** The control product was chemically (H$_3$PO$_4$) acidified milk without bacterial strains. Portions of 125 g were applied as yoghurt drink filled into mini-bottles.

The test products (verum and placebo) were identical in flavour, colour, texture, appearance and packaging (white bottles with plastic screw top) to avoid performance bias (Cochrane, 2011) by ensuring the blinding of both study participants and key study personnel including the outcome assessors. The nutritional values per unit (125 g) were: 82.5 kcal, 4.1 g protein, 7.0 g carbohydrates (2.5 g glucose), 4.8 g fat (2.9 g saturated fat), 0.06 g sodium.

NÖM AG produced, packaged, labelled and delivered the test products to the study site once every week by one-day courier service with icepacks in sealed cool boxes to keep the temperature of the products during the transport between 0 and 8°C (as monitored by a temperature logger). The test products were stored in a refrigerator (between 2 to 6°C) upon arrival in the study site. Temperature was continuously monitored by a temperature logger system.

Based on previous technology and stability tests, shelf-life with storage was 20 days after production. Methods and results of these pre-trials are given in Supplementary Material S1 – Table S1. In a further analysis the study products were tested directly before study conduct. A freshly produced sample and a sample stored until the end of shelf-life were assessed for their content of lactobacilli, *L. rhamnosus* and of further lactobacilli (comprising *L. crispatus, L. jensenii* and *L. gasseri*). Results are also provided in the Supplementary Material S1.

During the intervention period of 28 days each study participant consumed two yoghurt drinks daily after V1. One bottle should be consumed in the morning and one in the evening at any time which was convenient.

**Dietary restriction**

Subjects were asked to abstain from V1 to G2 from consumption of fermented dairy products (except the study products), probiotic-containing foods or supplements (over-the-counter (OTC), vitamin and mineral supplements). The subjects were also asked to avoid changes to their normal diet during the study period (i.e. do not start any special diet, etc.).

**Metronidazole treatment**

According to the guidelines of the German Association of Gynaecology and Obstetrics (Mendling et al., 2013), BV was treated with metronidazole 500 mg (Arilin®, Dr. August Wolff GmbH & Co. KG, Bielefeld, Germany) twice a day for 7 days. Based on findings of sexually transmitted *Gardnerella* biofilms (Swidsinski et al., 2010b,c) the patients were offered metronidazole treatment of their partners.
Objectives

The primary objective of this trial was to compare the effect of oral intake of yoghurt containing the pertinent four probiotic strains with that of placebo on the rate of recovery from BV after standard treatment. The secondary objectives were: (1) to compare the effect on the symptoms of BV, such as vaginal discharge or fishy odour of oral intake of yoghurt containing the pertinent four probiotic strains with that of placebo when accompanying standard treatment; (2) to compare the effect on the vaginal microbial pattern (dysbiosis) of oral intake of yoghurt containing the four probiotic strains with that of placebo when accompanying standard treatment.

Assessments/parameters

Primary parameter

Rate of BV-free women at the end of the yoghurt consumption period (day 28) as assessed by Nugent criteria assuming presence of BV at a Nugent score 7 to 10 (Nugent et al., 1991). In order to ascertain utmost quality of the primary parameter, Nugent score was assessed by one investigator (P.A.) in an external laboratory (LADR) at the end of the study conduct (maintaining concealment for double-blindness).

Secondary parameters

1. Alteration (G2-G1) of Nugent score (Nugent et al., 1991).
2. Alteration (G2-G1) of symptom score based on Amsel criteria 2 (thin, homogeneous discharge) (Yes=1; No=0) and 3 (release of amine (‘fishy’) odour after the addition of 10% KOH (Whiff test) to vaginal smear) (Yes=1; No=0) resulting in a score ranging from 0 to 2.

Exploratory parameters

- Rate of BV-free women at G2 as assessed by Amsel criteria (Amsel et al., 1983).
- Alteration (G2-G1) of Amsel criteria 1 to 4.
- Alteration of vaginal pH.
- Vaginal pH at G2.
- Occurrence of antibiotic-associated diarrhoea (AAD) (as defined by number of subjects with diarrhoea within the intervention period of 4 weeks (after starting antibiotic treatment). Diarrhoea was defined after WHO: three or more loose or watery stools (Bristol Stool Scale 5-7) per day for at least 1 day.
- Duration of AAD: number of consecutive days of AAD (start and end-date included), fulfilling the characteristics of diarrhoea definition.
- Average frequency of bowel movements within a period of AAD (median of number of bowel movements per day during an AAD period).
- Cumulative severity of AAD: sum of transformed Bristol scale values of all stools during an episode (transformation Bristol scale: 1 to 4 = 0; 5 = 1; 6 = 2 and 7 = 3).
- Gastrointestinal symptom rating scale (GSRS) (Dimenäs et al., 1995 (Table II), Questionnaire after Revicki (1998) (Appendix) and Svedlund (1988) (Appendix)) at V1, day 7, day 14, day 21 after starting intervention and day 28 at the end of intervention.
- Dietary habits of both groups at study start (V1) and end (G2) based on the EPIC-food-frequency-questionnaires (FFQ).
- Adverse events. An adverse event was defined by the appearance or worsening of any undesirable sign, symptom, or medical condition occurring in a subject even, if the event is not considered to be related to study products. Medical conditions/diseases present before starting study products were only considered adverse events, if they worsen after starting study products. Abnormal laboratory values or test results were regarded as adverse events only, if they induced clinical signs or symptoms, if they were considered clinically significant or required therapy. The occurrence of adverse events was sought by non-directive questioning of the subject at each visit during the study. Adverse events might also have been detected when they were addressed by the subject during or between visits or detected through physical examination, laboratory test, or other assessments.

Compliance. The subjects documented test product consumption daily and during the first 7 days also the consumption of the antibiotics in their personal diary. They also recorded consumption of forbidden products and possible concomitant medication including dietary supplements and OTC. The subjects were requested to return their subject diary at each visit in order to review recorded information by the study team. The investigator or authorised study staff counted the number of non-used study products, and documented intake of forbidden products and possible concomitant medication. In addition subjects completed a questionnaire concerning antibiotics compliance (at VD) and product compliance (both according Morisky et al., 1986) at visit V2.

Statistical analysis

Estimation of sample size

The study was designed as pilot trial, since no data were available from studies in BV in which the pertinent strains were used. For technical reasons (expenses for test yoghurt production over time) the sample size was limited to the number of study participants who had finalised the study period until end of August 2016.
**Preventing bias**

In order to meet the recommendations of the Cochrane Collaboration for preventing detection bias (Cochrane, 2011) blinding of outcome assessment was ensured by a blind review of raw data and by un-blinding only after data base was locked, and by conducting statistical analysis in compliance with the statistical analysis plan. In order to avoid attrition bias (Cochrane, 2011), distribution of potentially missing data across intervention groups was assessed. Missing values were equally distributed between intervention groups and occurred in less than 5 (4.38%) for primary and secondary outcomes and therefore were considered to have negligible impact and to be missing completely at random. This was corroborated by a sensitivity analysis: number of subjects with missing values based on Nugent score at visit G2 did not differ (P=1.000) between the two interventional groups in Chi square test. The full analysis set (FAS), which was *a priori* defined to comprise the analysis set for effectiveness evaluation, implied missing values for the primary parameter in one case and not at all for the secondary parameters. Reporting bias by selective outcome reporting (Cochrane, 2011) was prevented by the availability of the study protocol including its amendments and pre-specification of (primary and secondary) outcomes and by adhering to these specifications.

**Sets analysed**

The Intention-To-Treat (ITT) collective was defined to comprise all subjects randomised and having taken at least one dose of the test products (first day after visit 1). Since compliance with the ITT principle necessitates complete follow-up of all randomised subjects for study outcomes and since this is often difficult to achieve, the evidence for an effect was examined in the full analysis set (FAS) according to ICH E9 Guideline. It is as complete as possible and as close as possible to the ITT-set (FAS). Elimination of subjects was considered to be justified according to ICH E9 Guideline in the following cases: violation of an essential, before randomisation objectively measurable inclusion criterion, taking not a single dose of the test substance (without knowledge of the assigned test group), lack of any dates for the judgment of the effectiveness after randomisation. Additionally, a per-protocol (PP) analysis was carried out to check the robustness of the product effect. For this analysis all subjects randomised, who had any major protocol deviation were excluded.

**Tests**

The baseline characteristics of the 2 groups were compared using Mann-Whitney U test. Fisher exact test was used for comparing recovery rates based on Nugent criteria in placebo versus *verum* group (at *P*<0.05 level a significant difference is assumed) and for BV cure rates as assessed by Amsel criteria. The alteration of Nugent score and symptom score (G2-G1) was compared between *verum* and placebo group using Mann-Whitney U test. The occurrence of AAD, defined as subjects who had AAD according to the WHO criteria, was compared with Fisher exact test. Other parameters, such as mean duration and cumulated severity of AAD episodes as well as mean frequency of bowel movement at AAD days were compared with Mann-Whitney test. The questionnaires according to the GSRS were evaluated in categories such as total score, reflux, indigestion, constipation and diarrhoea at following time points: at V1, day 7, day 14, day 21 and day 28. Both groups were compared with Mann-Whitney test. The α values of the secondary parameters were adjusted after Bonferroni-Holm.

Descriptive statistics are given in arithmetic mean, standard deviation (SD), standard error of the mean (SEM), median, 95% confidence interval (95% CI) and frequency of observation.

**3. Results**

**Subjects’ demographics**

As shown in the flow diagram (Figure 2) of the 93 screened women, 48 had BV and 48 were recommended for therapy by gynaecologists, 36 were screened by the investigators at CRC and considered suitable for inclusion and continued to randomisation. Of the 36 randomised women, 36 received the study products with 18 women in the *verum* group and 18 women in the placebo group, and thus comprise the ITT population. Among the 36 randomised women, 2 (1 in the *verum* group and 1 in the placebo group) dropped out. In the one of the placebo group primary and secondary outcomes were assessed after 2 weeks intervention. Thus, the FAS comprised n=33 women for the primary parameter and n=34 women for the secondary parameter.

Four (11.1%) subjects (n=2 in the *verum* vs n=2 in the placebo group) revealed major deviations so that n=32 women comprise the PP-population. The two deviations in the *verum* group were a complete non-compliance with the study protocol in one case and a too late G2 examination at the gynaecologist 3 month after end of intervention in another case. The two deviations in the placebo group were a premature G2 examination at the gynaecologist (two weeks after start of intervention) in one case and withdraw consent caused by an adverse event in another case. Consecutively, there were only few missing data, which were balanced across the groups making reasons for missing data unlikely to be related to true outcome and unlikely to have a clinically relevant impact on the intervention effect estimate indicating low risk of attrition bias according to the Cochrane Collaboration (Cochrane, 2011). The baseline characteristics for the ITT population are shown in Table 1. The randomisation was well-balanced, and there was no

---

**Beneficial Microbes**

Please cite this article as ‘in press’
relevant difference between these 2 groups at baseline in terms of demographics, medical history, or current disease. In the ITT population BV was diagnosed for the first time in n=20 and n=16 had a history of BV episode. In the FAS n=19 had no history and n=15 had a history of BV. In the verum group n=11 (FAS: n=10) had no history and n=7 (FAS: n=7) a history, in the placebo group n=9 (FAS: n=9) had no history and n=9 (FAS: n=8) a history of an episode.

Study products and compliance

The results of the product stability testing are provided in the Supplementary Materials S1. Compliance with taking the study products was high. It was 98.1±10.1% in the verum group and 99.6±8.1% in the placebo group based on counting (at visit D and V2) the study products consumed during intervention and it did not differ between both groups (P=0.835 in Mann-Whitney test). Morisky score confirmed the result with 3.8±0.6 in the verum group and 3.7±0.6 in the placebo group (P=0.696 in Mann-Whitney test). Compliance with antibiotic treatment was identical in both groups (98.3±5.4%, P=1.000 in Mann-Whitney test) based on counting the (not) consumed antibiotics at visit D. Compliance with adhering to antibiotic treatment according to Morisky was similar with 3.8±0.5 in the verum group and 3.9±0.5 in the placebo group (P=0.575 Mann-Whitney test).

Recovery rate of bacterial vaginosis

FAS: Out of the 17 women in whom BV was diagnosed based on Amsel criteria in the verum group before intervention (at G1) all 17 were BV-free based on Amsel criteria after 4 weeks intervention (at G2), whereas only 11 out of 17 became BV-free in the placebo group (Table 2). The two groups differed after 4 weeks intervention with P=0.018 in Fisher Exact test. Out of 16 women assigned for verum, in whom Nugent score was available during intervention and BV was diagnosed at G1 based on Amsel criteria, BV criteria were fulfilled after Nugent in 13 women at G1. Out of 17 women assigned to placebo, in whom Nugent score was available during intervention and BV was diagnosed at G1 based on Amsel criteria, BV criteria were fulfilled after Nugent in 13 women at G1. After intervention (at G2) all (16/16) women were BV-free based on Nugent criteria in the verum group compared to only 13/17 BV-free in the placebo group (Table 2). The two groups differed after 4 weeks intervention with P=0.103 in Fisher Exact test (primary parameter). The PP and ITT populations showed alterations in alignment with these results (see Supplementary Materials S2 – Tables S2 and S3).

In the verum group in 9 out of 17 women the partner were also treated with metronidazole; in the placebo group this was the case in 8 out of 16 women. Thus, this potential bias did not significantly differ between the groups (P=0.866 in Chi² test, resp. P=1.000 in Fisher exact test). Recovery failures were seen only in the placebo group. Two of them had no history of BV and three had a history of an episode.

Amsel score, symptoms and vaginal pH

The results for the FAS are shown in Table 3. Before intervention Amsel score did not significantly (P=0.901) differ between verum (3.59±0.62) (mean ± SD) and placebo group (3.65±0.49). After 4 weeks intervention Amsel score differed (P=0.006 in Mann-Whitney) between verum (1.18±0.39) and placebo (1.71±1.83). During the intervention period it decreased by -3.41±0.71 in the verum group compared to -1.94±1.95 in the placebo group (P=0.037 in Mann-Whitney test). Symptoms discharge and odour, as assessed by Amsel criteria 2 and 3 before
Beneficial Microbes

Probiotic yoghurt in bacterial vaginosis

Table 1. Baseline characteristics (mean ± standard deviation) of the intention-to-treat population.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Total population (n=36)</th>
<th>Verum (n=18)</th>
<th>Placebo (n=18)</th>
<th>Verum vs placebo²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>35.8±12.1</td>
<td>32.6±11.2</td>
<td>39.0±12.3</td>
<td>0.112</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.1±4.4</td>
<td>25.2±4.4</td>
<td>24.9±4.5</td>
<td>0.849</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.69±0.1</td>
<td>1.68±0.1</td>
<td>1.69±0.1</td>
<td>0.503</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>71.5±15.2</td>
<td>72.3±13.9</td>
<td>70.7±16.8</td>
<td>0.752</td>
</tr>
<tr>
<td>BP syst (mmHg)</td>
<td>111.1±11.2</td>
<td>117.7±9.9</td>
<td>110.6±12.7</td>
<td>0.771</td>
</tr>
<tr>
<td>BP diast (mmHg)</td>
<td>72.2±6.0</td>
<td>72.2±7.9</td>
<td>72.2±7.9</td>
<td>1.000</td>
</tr>
<tr>
<td>Heart rate (min⁻¹)</td>
<td>69.4±6.2</td>
<td>70.2±6.1</td>
<td>68.6±6.3</td>
<td>0.442</td>
</tr>
<tr>
<td>Reproductive (number)</td>
<td>31</td>
<td>17</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Postmenopausal (number)</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0.338³</td>
</tr>
</tbody>
</table>

¹ BPsys = systolic blood pressure; BPdiast = diastolic blood pressure; BMI = body mass index.
² P-value in student t-test.
³ P-value in Fisher Exact test.

Table 2. Women with bacterial vaginosis (BV) in full analysis set.

<table>
<thead>
<tr>
<th></th>
<th>Verum</th>
<th>Placebo</th>
<th>Verum vs placebo²</th>
</tr>
</thead>
<tbody>
<tr>
<td>BV/n rate (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amsel criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>17/17</td>
<td>17/17</td>
<td>1.000</td>
</tr>
<tr>
<td>G2</td>
<td>0/17</td>
<td>6/17</td>
<td>0.018</td>
</tr>
<tr>
<td>Nugent criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>13/16</td>
<td>11/17</td>
<td>0.438</td>
</tr>
<tr>
<td>G2</td>
<td>0/16</td>
<td>4/17</td>
<td>0.103</td>
</tr>
</tbody>
</table>

¹ G1 = examination by the gynaecologist before intervention; G2 = examination by the gynaecologist after 4 weeks intervention.
² P-value in Fisher Exact test.

intervention, did also not significantly (P=1.00) differ between verum (1.71±0.47) and placebo group (1.72±0.46) and decreased significantly by -1.71±0.47 in the verum compared to -0.82±1.07 in the placebo group (P=0.01) (secondary parameter; α=0.017) and differed after 4 weeks intervention between the groups (0.00±0.00 vs 0.88±0.93; P=0.01) (Table 3). The single Amsel criteria also did not differ before intervention and differed after 4 weeks intervention with P<0.05 for criterion 2 (discharge), criterion 3 (odour) and criterion 4 (clue cells) and with p=0.171 for criterion 1 (pH above 4.5). Vaginal pH was 5.31±0.64 in the verum group compared to 5.22±0.39 in the placebo group before intervention (P=0.63 in Student t test). It decreased by -1.12±0.88 in the verum group compared to -0.65±0.76 in the placebo group (P=0.109 in Student t test) during intervention and tended to be lower (4.19±0.50) in the verum group compared to 4.57±0.74 in the placebo group after 4 weeks intervention (P=0.091 in Student t test) (Table 3). The PP and ITT populations showed alterations in alignment with these results (Supplementary Materials S3 – Tables S4 and S5).

Nugent score

Data for the full analysis set are shown in Table 3. Before intervention Nugent score did not significantly (P=0.437 in Mann-Whitney) differ between verum (7.06±2.72) (mean ± SD) and placebo group (6.65±2.5). After 4 weeks intervention Nugent score was somewhat lower (P=0.444 in Mann-Whitney) in verum (2.44±1.71) compared to placebo (3.82±3.57). During the intervention period it decreased by -4.65±2.85 in the verum group compared to -2.82±3.59 in the placebo group (P=0.158 in Mann-Whitney test; α=0.025 (secondary parameter)) (Table 3). The PP and ITT populations showed alterations in alignment with these results (Supplementary Materials S4).
Gastrointestinal symptoms

AAD (as defined by WHO) occurred in 4/17 women of the verum group during the 4 weeks intervention compared to 2/17 women of the placebo group (P=0.656 in Fisher Exact test). Seven AAD episodes were seen in each group. Mean duration of AAD was 1.43±0.53 days in the verum group compared to 1.71±1.89 in the placebo group. Cumulated duration was 10 days in the verum group compared to 12 in the placebo group. Mean severity of AAD episodes was 9.29±4.6 days in the verum group compared to 11.0±5.5 in the placebo group. The average frequency of bowel movement was 3.2±0.42 stool per day in the verum group compared to 3.42±0.9 in the placebo group.

The average defecation frequency during the 4 weeks intervention period did also not differ between the groups. It was 1.24±0.58 per day in the verum group compared to 1.3±0.66 in the placebo group (P=0.589). Bristol stool form score defined as the sum of products of each Bristol stool form count multiplied by the number of type (1 to 7) for each type 1 to 7 was 124 (93.0; 167) in the verum group compared to 124 (71.5; 171) in the placebo group (P=0.971 in Mann-Whitney U test).

The total score according to GSRS did neither differ between both groups in the week before intervention (1.20 (1.02; 1.73) in the verum and 1.20 (1.02; 1.63) in the placebo group (P=0.870 in Mann-Whitney U test)), nor during the 4 week intervention. In week 1 it was 1.47 (1.23; 1.70) in the verum and 1.47 (1.23; 1.67) in the placebo group (P=0.871);
in week 2 it was 1.53 (1.07; 1.87) in the verum and 1.23 (1.02; 1.38) in the placebo group (P=0.11); in week 3 it was 1.47 (1.17; 1.77) in the verum and 1.20 (1.08; 1.40) in the placebo group (P=0.051); in week 4 it was 1.33 (1.10; 1.70) in the verum and 1.13 (1.00; 1.37) in the placebo group (P=0.077).

**Adverse events**

In 11 women of the verum group compared to 12 women of the placebo group adverse events were seen (P=0.71 in Chi² test). The number of adverse events was 20 (3 AAD, 7 gastroenteritis, 6 pain, 4 other) in the verum group compared to 27 (3 respiratory tract infection, 6 gastroenteritis, 11 pain, 7 other) in the placebo group (P=0.55 in Mann-Whitney U test).

**Dietary intakes**

As assessed by FFQ (German Institute of Nutrition, Potsdam-Rehbrücke, Germany) both study groups did not differ in their dietary habits neither at start nor at the end of the intervention period (Supplementary Materials S5 – Table S6).

**4. Discussion**

The results were obtained in a study following the design of a double-blind randomised placebo controlled trial and taking principles of good clinical practice into account. Furthermore, the design and conduct was dedicated to avoidance of selection, performance, detection, attrition and reporting bias following the recommendations of Cochrane Collaboration (Cochrane, 2011). The population met inclusion criteria. Baseline characteristics, including dietary intake of relevant nutrients, did not differ between the two intervention groups. Compliance with intake of the study products, with antibiotic treatment of BV and with dietary restriction was high and did not differ between the groups.

Dietary intakes as assessed by FFQ did not differ between the groups neither before nor during intervention. The number of drop-outs, missing values and major deviations from the protocol was low and equally distributed between the intervention groups. Even though in most published studies with similar design (Anukam et al., 2006; Martinez et al., 2009; Petricevic and Witt, 2008) the sample size was below n=100, the sample size of this pilot trial was rather small with n=36 in the ITT and n=34 in the FAS population. The population was also heterogeneous in respect to the reproductive status.

In spite of this a significant difference between verum and placebo was found in one of the predefined primary and secondary parameters even after adjustment for multiple testing providing evidence for an effect. The reduction of symptoms of BV (homogenous, thin discharge and odour in whiff test), as assessed by Amsel criteria 2 and 3, was significantly more pronounced in the verum compared to the placebo group. This was in line with the difference in alteration of the total score based on Amsel criteria and was associated with a more pronounced reduction of clue cells. This may indicate that bacteria adherent to the vaginal epithelium, which are well protected in biofilms against the access of antibiotics (Swidsinski et al., 2013) were more efficiently diminished by the addition of these probiotics to antibiotic treatment than by antibiotics alone.

BV recovery rate by metronidazole, as assessed by Amsel criteria, was 64.7% (11/17) in the placebo arm. This is in line with results reported by others (Anukam et al., 2006; Larsson and Forsum, 2005; Martinez et al., 2009; Petricevic and Witt, 2008; Verstraeten and Verhelst, 2009). In the verum arm the recovery rate was higher with 100% (17/17). The recovery rates in studies using other probiotics in a similar study design ranged between 83 and 88% (Anukam et al., 2006; Martinez et al., 2009; Petricevic et al., 2008). The high recovery rate in this trial may indicate that the Lactobacillus mixture applied in this trial is particularly efficacious in BV. This may be due to the fact that the 4 strains belong to the species predominating vaginal microbiota. Particularly L. crispatus and L. jensenii were found to be most abundant in vaginal microbiota of healthy women (Antonio et al., 1999; Arouicheva et al., 2001; Fredricks et al., 2005, 2007; Liu et al., 2010; Song et al., 1999) and to be associated with lower frequency of BV (Antonio et al., 1999; Arouicheva et al., 2001; Fredricks et al., 2005, 2007; Liu et al., 2010). These species, however, were not part of the probiotics used in other studies with similar design (Anukam et al., 2006; Martinez et al., 2009; Petricevic and Witt, 2008; Reid and Bocking, 2003). Furthermore the selection of the 4 strains for bile salt tolerance, growth under anaerobic and anaerobic atmospheres, glycogen utilisation, acidification capacity, production of extracellular hydrogen peroxide, which is decreased in BV, and growth inhibition of pathogens, like G. vaginalis (De Seta, 2014; Domig et al., 2014), may have facilitated not only survival of gastrointestinal transit, but also colonisation of the vagina and regression of bacteria involved in the pathogenesis of BV.

BV recovery rate, as assessed after Nugent showed a trend towards the same direction with again 100% in the verum arm and 76.5% in the placebo arm. Likewise, the reduction of Nugent score during intervention tended to be more pronounced in the verum group compared to the placebo group. The differences between verum and placebo were statistically less clear for the Nugent score than those based on Amsel criteria in this trial with limited sample size. This may be due to the fact that the study participants were included based on the diagnosis of BV according to Amsel criteria, which resulted by definition in 100% prevalence of BV before intervention according to these criteria. In contrast the prevalence at the starting point was somewhat lower (81.3 in the verum arm and 64.7% in the placebo arm).
in case of assuming Nugent criteria for the presence of BV which reduced the probability for finding a statistically significant difference after intervention.

According to the shift from abnormal vaginal microbial pattern to one with higher lactobacilli counts, as indicated by Nugent score, the vaginal pH decreased from 5.32±0.62, which is outside the normal range (3.8 to 4.4) in adult women (Hauth et al., 2003; Kiss et al., 2004; Krauss-Silva et al., 2014; Martius and Hoyme, 2006; Platz-Christensen et al., 1993) to 4.19±0.50 in the verum group, whereas it remained on average above the normal range with 4.57±0.74 in the placebo group.

The design compares the outcomes after standard therapy in a group taking yoghurt containing the pertinent strains with that in a group taking a milk product with identical nutrient content, taste and appearance. So far the differences in outcomes found between the groups can be interpreted as caused by the yoghurt containing the pertinent strains. Since we did not investigate in this pilot trial which strains colonised the vagina in which proportion of the women, we do not yet know which strains were most effective in this mixture. We also cannot exclude that *S. thermophilus* or *L. delbrueckii* ssp. bulgaricus strains colonised the vagina of some women. The epidemiological data on the composition of vaginal microbiota (Antonio et al., 2001; Arouetcheva et al., 2001; Fredricks, 2011; Fredricks et al., 2005, 2007; Hyman et al., 2005; Ling et al., 2010; Oakley et al., 2008; Ravel et al., 2011; Srinivasan et al., 2012; Verhelst et al., 2004), however, make this appear rather unlikely.

Probiotic lactobacilli by definition are supposed to survive gastrointestinal transit after oral administration resulting in an increase of the number of lactobacilli in general and the number of the orally administered strains in the colon and faeces (Schrezenmeir et al., 2001). Gut and faeces, respectively, share some *Lactobacillus* strains with the vagina (El Aila et al., 2009, 2011; Petricevic et al., 2012) and serve as reservoir for some strains of lactobacilli that can colonise the vagina through the close proximity of the rectum and anus, respectively, and the vagina. This colonisation is evidently based on mechanisms known as contact infections by pathogens (Lincoln et al., 1970; Orskov and Orskov, 1985; Turck et al., 1962; Vahlne, 1945) and was demonstrated in intervention trials in healthy women for lactobacilli, too, at the genus, species and strain level (Anukam et al., 2006; Reid and Bocking, 2003; Reid et al., 2001; Strus et al., 2012; Vasquez et al., 2005). Particular properties of some species and strains may favour their physiological growth in the vaginal environment and the colonisation of the vagina by the rectal and faecal route, respectively (Antonio et al., 2005; Meyn et al., 2002; Swidsinski et al., 2010a).

The similar values of *verum* and placebo group in gastrointestinal symptoms, particularly in incidence, duration and severity of AAD, in defecation frequency and stool form, in total symptom GSRS score, and in number of adverse events show that the yoghurt containing the pertinent strains was well tolerated. The fermentation study and stability testing showed that the strains used in this study can be included in the fermentation process for yoghurt production and that they are adequately stable for applying them in yoghurt drinks.

5. Conclusions

The results indicate that the administration of a yoghurt containing *L. crispatus* LbV 88 (DSM 22566), *L. gasseri* LbV 150N (DSM 22583), *L. jensenii* LbV 116 (DSM 22567) and *L. rhamnosus* LbV96 (DSM 22560) in addition to antibiotic treatment may improve recovery rate and symptoms of BV and is well tolerated.

Supplementary material

Supplementary material can be found online at https://doi.org/10.3920/BM2017.0018.

Appendix S1. Probiotic product stability tests.

Appendix S2. Recovery rate of bacterial vaginosis for PP and ITT population.

Appendix S3. Amsel score, symptoms and vaginal pH for PP and ITT population.

Appendix S4. Nugent score for PP and ITT population.

Appendix S5. Dietary intake.

Acknowledgements

We thank Tara Dezhahang for their excellent support and for technical assistance.

References


Reid, G., Bruce, A.W., Fraser, N., Heinemann, C., Owen, J. and Henning, B., 2001. Oral probiotics can resolve urogenital infections. FEMS Immunology and Medical Microbiology 30: 49-52.


