

The Postbiotic ReFerm® Reduces Activated Hepatic Stellate Cells in Advanced Alcohol-Related Liver Disease (GALA-POSTBIO): A randomized controlled phase 2 trial

Corresponding Author: Professor Aleksander Krag

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this article, Hansen et al. evaluated the effect of the postbiotic ReFerm on fibrogenesis in patients with advanced ALD. Unfortunately, the primary endpoint of this trial was not met. However, the number of patients is small and the data are very preliminary.

As regards other analyses, especially as regards gut microbiota and intestinal permeability, Figure 5 is very nice, but I think that presented data should be at least confirmed in animal models prior to hypothesize the mechanism of action.

Another pitfall is that patients with stage 2-3-4 fibrosis were enrolled. This may have affected the results because once cirrhosis is established, regression of fibrosis may be more limited.

Minor comments:

-Introduction lines 121-128 is quite elementary, I suggest a revision

-Line 163-165: compensated liver disease should not comprise Child B (at least Child B8-9). However I think that the authors may rewrite the sentence as they almost included Child A patients (only 1 patient in the ReFerm arm was not Child A... but some patients were not classified as cirrhotics at liver histology!)

-Grammatical errors should be corrected (e.g. line 125 "induce", line 127 "halter" etc)

Reviewer #2

(Remarks to the Author)

The authors have admirably performed a randomized phase 2 trial of a postbiotic (ReFerm) vs standard nutritional support to determine if hepatic stellate cell activation was improved in the intervention cohort. Patient were compensated ACLD patients with current or former drinking and were designated as ALD patients with no other cause of liver disease. Though the primary outcome was not different between groups, the results are intriguing and the authors should be commended for pushing the field forward to find new treatments for ALD, which has few, if any, treatments besides alcohol abstinence.

Cohort:

--the cohort is predominantly male and older, which is less generalizable to the population with ALD, which is skewing younger and more female

--did the exclusion criteria also exclude those who would be designated as MASLD or MetALD?

--the manuscript seems to indicate that Childs Score C patients were enrolled but this seems to be a mistake perhaps?

Protocol/Intervention:

--the protocol reviewed appeared slightly different from the manuscript. The product in the protocol (Profermin) is different from ReFerm, which is reported as the intervention product in the manuscript. Can the authors clarify? Are these the same product but just under a different name? They appear to be described as the same thing but with a different name.

--the control cohort is getting a nutritional supplement which in and of itself may improve the outcomes (as nutrition is a key modulator of many features of ALD). One wonders if perhaps comparing to patients eating a standard Danish diet might have been more helpful and pragmatic.

Outcomes/Measures:

--for the primary outcome, can the authors please indicate why 10% or greater absolute reduction in α SMA expression (marker for hSC activation) was chosen as the threshold? Is there some data that suggests that a 10% reduction is significant or would be clinically meaningful? There is some discussion of this on page 12 lines 227-236 but it was a bit confusing and didn't really explain why 10% would be a meaningful threshold. And was this reduction level influenced by power calculations?

--was PETH measured throughout the trial or just at baseline?

Results:

--50% of each group was abstinent 1 week prior to inclusion. Does this mean that the other 50% were actively drinking? ----it appears that, by PP and ITT analysis, the primary outcome was not met but that this may have been related to non-compliance. Compliance was reported as 92% for the intervention cohort and 99% for the control cohort. How was compliance assessed? Also, this seems like a quite high compliance rate overall so it may be more difficult to say that noncompliance was the reason for failure.

--alcohol intake is a very significant confounder for the outcomes assessed. It appears that the median of the group was 32 g/day (about 2-3 drinks per day), but there were some that were abstinent. How did results differ between an abstinent cohort vs an actively drinking cohort? Either way, that data would be interesting to note as an intervention that helped improve liver function with or without drinking would be quite beneficial.

--ReFerm group had higher baseline TE scores (27 vs 20 kPa). ELF and Fibrosis scores were more comparable so one wonders why this difference is present.

Reviewer #3

(Remarks to the Author)

Thank you for the opportunity to revise this interesting work comparing two randomized arms, ReFerm and Fresubim. The manuscript is well written and many outcomes are assessed. Please find below my comments:

- Abstract: authors report an 8.3% reduction in α -SMA in patients under ReFerm compared to Fresubin. I suggest to report in the abstract the changes within the two arms instead of difference only, in order to show magnitude and direction of changes.
- Page 9: authors report that compliance was assessed but it is not clear how it was assessed in terms of type of variable or variables collected.
- A general comment is about adjustment. As stated at page 12, results were reported without adjustment unless specified. It would be helpful to more clearly report why no possible confounders were taken into account. Additionally, at page 15 compliance adjustment was done for α -SMA mean reduction but it is not clear if adjustment was assessed also for the primary endpoint ($\geq 10\%$ absolute reduction).
- Page 14: the authors state that variables in Table 1 were comparable but no balance measures such as standardized mean difference are reported. Why?
- Figure S2A: please report in the legend what is shown in the figures. Median(IQR)?
- How was normality assessed?
- I am a bit confused with Figure S3. What is the difference between the two figures?
- For variables collected at several time points, did the authors consider to perform statistical models for longitudinal analysis?
- Page 19, line 349: the authors state that a reduction in α -SMA was strongly correlated with I-FABP and NT-3. Do they really refer to correlation coefficients?

Version 1:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

The authors have addressed my suggestions or justified their choices and they updated the manuscript accordingly.

Reviewer #4

(Remarks to the Author)

The revised version of the article add more details to a delicate and interesting topic: reversal of liver fibrosis in ALD. The authors sensibly tried to upgrade and update the article.

Some point needs to be clarified:

- The data on the animal model are of interest. Indeed, they are limited. I think this point " limited data " should be pointed out by authors also in the Discussion section, subsection Limits of the study. Furgether, larger data are needed to confirm proof of concept data.
- the population in study is variegated and the representation of fibrosis is not " normal ". This may have affected significance

of results.

I would suggest to highlight the point in the limitations. Moreover, I think that when the cohorts are enlarged, the effect on advanced fibrosis could be milder. Could the authors comment on this ?

- the Fresubin control arm: can it have biased results ? has it just nutritional impact ? can the increased and ameliorated protein pool 8if any) affect intestinal permeability parameters ? other measured variables ? Please, comment on this.

- an English language revision and, final, text cleaning is warranted.

- is 10% effect tight for Referm (reduction in aSMA expression) ?

- Hot point: the active drinking is a bias of the study and, also, a real-life data: how can this affect and, possibly, counteract the Referm effects ? Please, comment on this.

- Thus, non-compliance has affected results' significance (namely, primary end-point) ? I sit the only explanation ? Please, comment on that.

Version 2:

Reviewer comments:

Reviewer #4

(Remarks to the Author)

The authors have sufficiently addressed the queries and hot topics risen from article reviewing.

We look forward a cleaned version (during the editing process of the manuscript).

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer #1:

In this article, Hansen et al. evaluated the effect of the postbiotic ReFerm on fibrogenesis in patients with advanced ALD.

Unfortunately, the primary endpoint of this trial was not met. However, the number of patients is small and the data are very preliminary.

As regards other analyses, especially as regards gut microbiota and intestinal permeability, Figure 5 is very nice, but I think that presented data should be at least confirmed in animal models prior to hypothesize the mechanism of action.

Reply:

We acknowledge that the animal model can strengthen the human data and thereby support the proposed mechanism of action.

In a pilot set of experiments using 11 mice, we tried to confirm pathophysiological changes and the effects of ReFerm®. Indeed, we were lucky enough to have the authorities approval in place for the experiments included in the manuscript.

For an analysis beyond a pilot study, we would need a much higher number of animals, which would take a long time for the animal experimentation application, approval process as well as performing the models in groups of more than 12 animals per group. However, due to the short time for the revisions and especially because our study is evaluating the treatment in humans, we hope that the reviewer would agree that this pilot animal study is sufficient as a proof of concept in animals.

We have added the following text, figures, and tables to the results, discussion, and methods sections respectively:

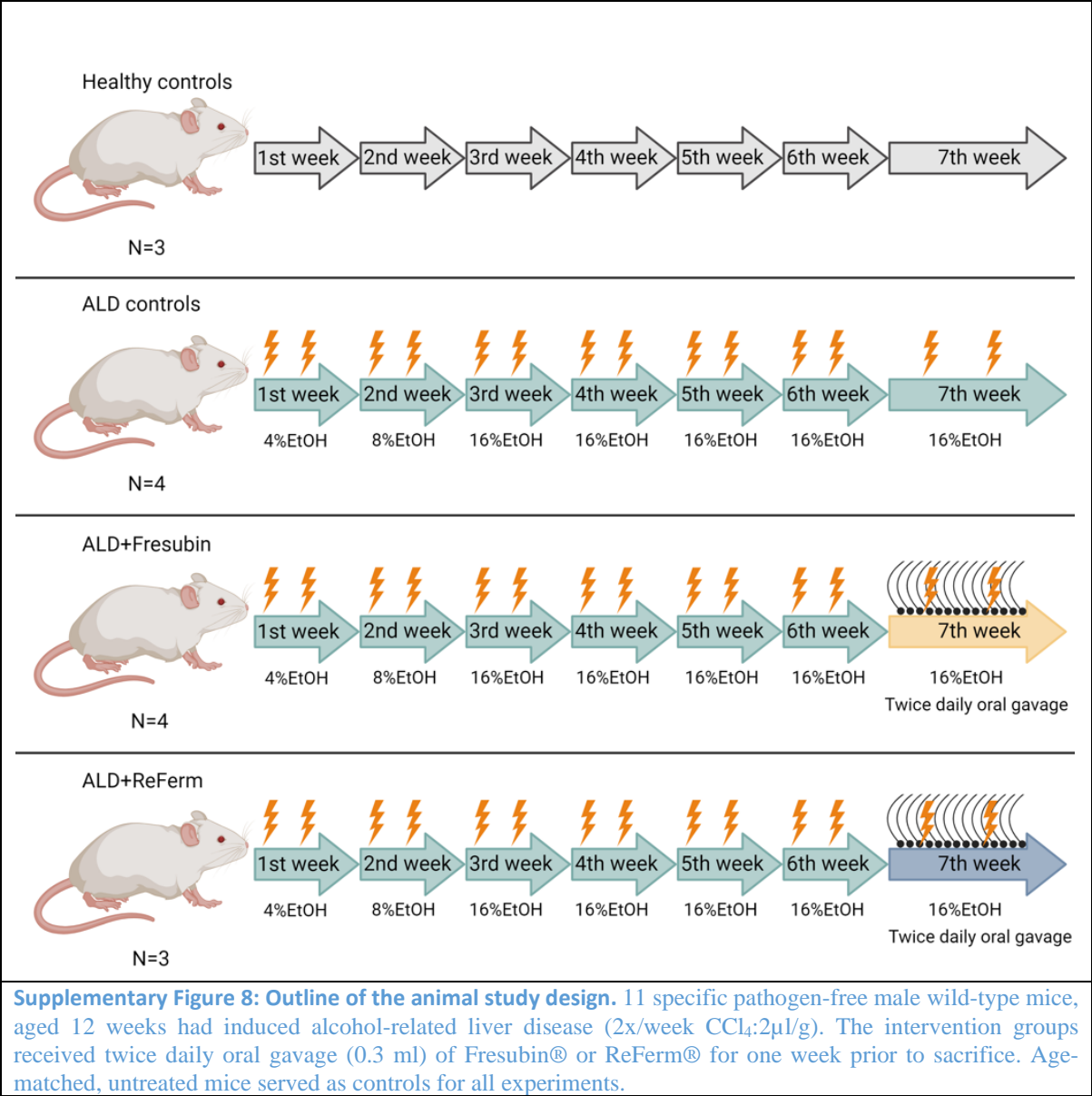
Results

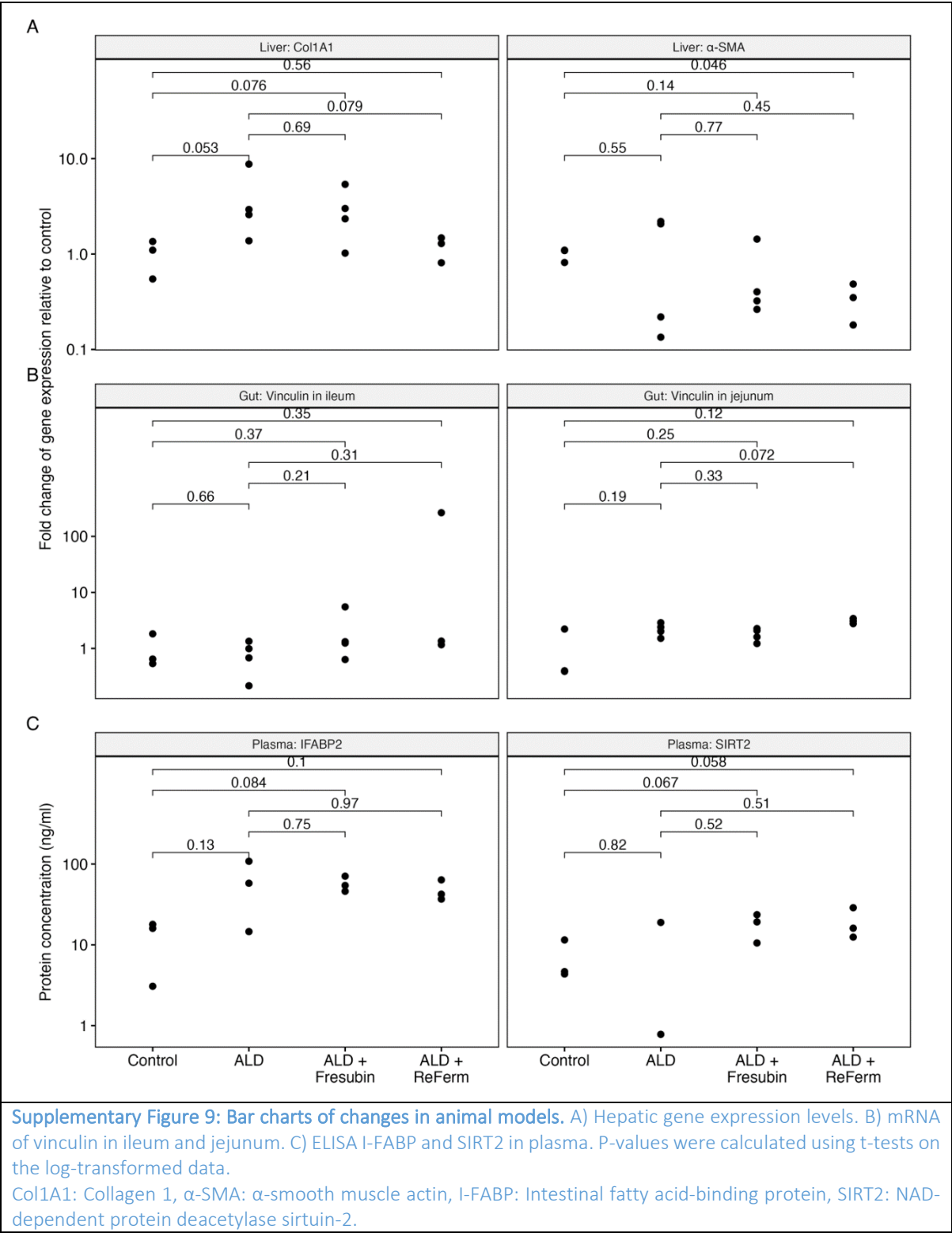
Page 11, line 224

“Validation of key findings in animal models

In order to explore the proposed mode of action in an animal model, ReFerm® and Fresubin® were applied in an animal model of ALD (Figure S8). In this model, 11 specific pathogen-free male wild-type mice, aged 12 weeks had induced ALD. The intervention groups received twice daily oral gavage (0.3 ml) of Fresubin® (N=4) or ReFerm® (N=3) for one week prior to sacrifice, while an ALD control group remained untreated (N=4). Age-matched, untreated mice served as controls for all experiments (N=3). On gene expression level, induction of ALD led to an increase in liver Collagen 1 (Col1A1, $p = 0.053$) that was ameliorated by ReFerm® treatment ($p = 0.079$ for ALD with and without ReFerm®). Compared to untreated controls, ReFerm® treatment also led to a decrease in α -SMA ($p = 0.046$). These changes suggest less activation of HSCs and decreased collagen production (Figure S9A). In the gut, an increase of gut barrier integrity was demonstrated through elevated

mRNA levels of vinculin in jejunum ($p = 0.07$ for ALD with and without ReFerm[®], Figure S9B). I-FABP and SIRT2 protein levels were increased in the ALD group compared to control mice. However, no significant changes could be observed for I-FABP and SIRT2 protein levels (Figure S9C). In conclusion, this small-scale exploration on the effects of ReFerm[®] in an animal model shows a similar response to treatment in humans and mice.”





Discussion (underlined added):

“On top of this the reduction of gut barrier impairment marker I-FABP was associated with an increase in levels of NT-3 that is considered to promote hepatocyte proliferation leading to liver regeneration^{1,2} and increased levels of SIRT2 potentially preventive for ethanol-induced liver injury.³ These results suggest a potential for hepatic recovery when arresting disease progression. This trend was confirmed in the proof-of-concept animal model, where we also identified signals indicating enhanced gut barrier function, along with reduced α -SMA in the liver and combined decreased systemic inflammation.”

Method:

Page 24 line 472

“Animal experimentation

*A total of 11 specific pathogen-free male wild-type (WT, C57Bl6/J) mice, aged 12 weeks, were used in this study. The mice were obtained from Charles River Laboratories Research Model and Services, Sulzfeld, Germany. They were housed at 22°C with a 12-hour light/dark cycle in individually ventilated cages. ALD was induced by administering intraperitoneal CCl₄ injections twice weekly for 7 weeks. In addition to CCl₄, the mice received phenobarbital (0.33 g/l) in their drinking water to stimulate cytochrome P-450 metabolic activity. Ethanol was added to their drinking water (4% in week 1, 8% in week 2, and 16% until euthanasia). Water and chow were available ad libitum, and further details about the diet are provided in **Table S5**. The intervention groups received twice daily oral gavage (0.3 ml) of Fresubin® or ReFerm® for one week prior to sacrifice. Age-matched, untreated mice (N=3) served as controls for all experiments. The experimental design is visualized in **Figure S8**. Before euthanasia, the mice were anesthetized with an intraperitoneal injection of ketamine-xylazine (100 mg ketamine/kg body weight and 10 mg xylazine/kg body weight). At organ harvest, liver and serum samples were collected. Liver samples were snap-frozen and stored at -80°C. Blood samples were allowed to clot for 30 min at room temperature and aliquoted after centrifugation at 2000g for 10 min. Serum aliquots were immediately stored at -80°C. All animals received human care in accordance with the criteria outlined in the EU regulations on animal research (2010/63/EU). All experiments were performed in accordance with the German animal protection and welfare law and the guidelines of the animal care facility at the Hospital of the Goethe University Frankfurt and were approved by the responsible local authorities, the Darmstadt regional council (File reference number: FK/2005).*

Quantitative PCR and Enzyme linked immunosorbent assay (ELISA)

Total RNA was extracted using a standard TRIzol-based protocol (TRIzol Reagent, Ambion, Carlsbad, CA, USA). cDNA synthesis and quantitative polymerase chain reaction (qPCR) were performed as described previously.⁴ TaqMan gene expression assays (Thermo Fisher Scientific, Waltham, MA, USA) were used for qPCR according to the manufacturer's protocol on a 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA). Each qPCR analysis included duplicate wells, and appropriated control reactions were performed in all samples. The expression of each gene was calculated by the $2^{-\Delta\Delta C_t}$ method by Schmittgen and Livak.⁵ Gene

*amplification results were standardized against 18S rRNA expression in each sample, and expression levels were presented as x-fold changes relative to the corresponding control group. A full list of the gene expression assays used is provided in **Table S6**. Sandwich enzyme-linked immune-sorbent assays for I-FABP (EM1144, FineTest, Wuhan, China) and SIRT2 (A77325, Antibodies.com, Stockholm, Sweden) were performed with murine serum samples. Serum samples were thawed and diluted 1:2 with sample dilution buffer and all buffers and standards have been prepared according to the respective assay protocols. The ELISA protocols were applied in accordance with the manufacturer's specifications. The final step of the protocol is a color change, which is detected at 450nm on a microplate reader. The concentrations of SIRT2 and I-FABP were calculated referring to the standard curve by applying a four parameter logistic curve and multiplication of the dilution ratio.”*

Ingredients	Control (V1534–300)	Diet
Crude protein, %	19.0	
Crude fat, %	3.3	
Crude fibers, %	4.9	
Crude ash, %	6.4	
Starch, %	35.2	
Sugar, %	5.3	
Vitamin A, IU/kg chow	25,000	
Vitamin D ₃ , IU/kg chow	1,500	
Vitamin E, mg/kg chow	125	
Vitamin K ₃ , mg/kg chow	20	
Copper, mg/kg chow	5	
Supplementary Table 5: Animal Chow. Manufacturer: Ssniff Spezialdiäten, Soest, Germany; IU, international units.		

Gene	Assay ID	Species
Col1a1	Mm00801666_g1	Mus musculus
Acta2 (α -SMA)	Mm00725412_s1	Mus musculus
Vinculin	Mm00447745_m1	Mus musculus
Supplementary Table 6: Gene expression assays. Manufacturer: Thermo Fisher Scientific. Col1a1, collagen type 1a1; Acta2 (α -SMA), α -smooth muscle actin.		

Another pitfall is that patients with stage 2-3-4 fibrosis were enrolled. This may have affected the results because once cirrhosis is established, regression of fibrosis may be more limited.

Reply: We agree that there can be biological differences in the ability for fibrosis regression depending on the severity of fibrosis. However, while it may be difficult to achieve regression in the histologically defined fibrosis stage 4, it is considered possible to achieve regression in the thickness of fibrotic septa for both severe fibrosis and compensated cirrhosis.⁶ This is also the reason why we did not choose

histological fibrosis regression as primary endpoint. However, a regression in the overall area of fibrosis would in some cases lead to a lower histological fibrosis score as well. Therefore, we chose to include both groups of patients in the design of the study.

	F2, baseline	F3, baseline	F4, baseline
Patients, n	4	13	26
Improving Kleiner fibrosis stage, n(%)	0(0)	4(30)	5(19)

In total 9 patients regressed at least one fibrosis stage during the study period. We cannot address whether this finding is explained by sampling variability or reflects a true regression of fibrosis. We acknowledge that this may contribute to a heterogeneous effect of the intervention, and we have therefore added that this is a limitation of our study:

Page 17 line 333

"Third, this study design included patients with ALD ranging from significant fibrosis (F2) to compensated cirrhosis (F4), as regression is considered achievable across this spectrum of liver fibrosis.⁶ There may likely be a biological difference in the ability for fibrosis regression depending on the severity of fibrosis, which may have led to a heterogeneous effect of the interventions."

Minor comments:

-Introduction lines 121-128 is quite elementary, I suggest a revision

Reply:

We acknowledge that lines 120-129 are relatively elementary. However, this is a part of the introduction, and we believe it is important to keep the level of detail such that readers can easily understand the background of the study including taking the broad readership of Nature Communications into account including both physicians and basic scientists. Therefore, we have chosen to keep this section. If the editors and reviewers feel strongly against this, we will of course revise the section accordingly.

-Line 163-165: compensated liver disease should not comprise Child B (at least Child B8-9). However I think that the authors may rewrite the sentence as they almost included Child A patients (only 1 patient in the ReFerm arm was not Child A... but some patients were not classified as cirrhotics at liver histology!)

Reply:

We thank the reviewer for addressing the one participant classified as Child B. This led us to reassess the algorithm for calculating the Child score, revealing that the classification of Child B was due to a typographical error (a missing comma) in the INR value. This has now been corrected, and in the revised version, all participants

Manuscript reference number: NCOMMS-24-26561

Title: The Postbiotic ReFerm® Reduces Activated Hepatic Stellate Cells in Alcohol-Related Liver Disease: A randomized controlled trial.

are classified as Child A. We apologies for this and we have corrected this error and updated Table 1 accordingly.

-Grammatical errors should be corrected (e.g. line 125 “induce”, line 127 “halter” etc)

Reply:

Thanks for noticing. Identified errors have been corrected.

Reviewer #2:

The authors have admirably performed a randomized phase 2 trial of a postbiotic (ReFerm) vs standard nutritional support to determine if hepatic stellate cell activation was improved in the intervention cohort. Patient were compensated ACLD patients with current or former drinking and were designated as ALD patients with no other cause of liver disease. Though the primary outcome was not different between groups, the results are intriguing and the authors should be commended for pushing the field forward to find new treatments for ALD, which has few, if any, treatments besides alcohol abstinence.

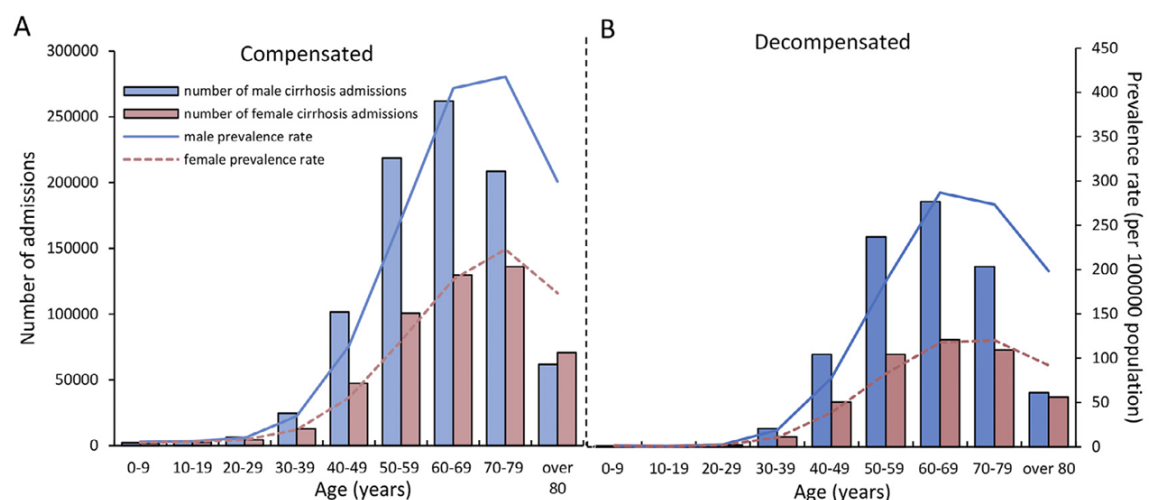
Cohort:

--the cohort is predominantly male and older, which is less generalizable to the population with ALD, which is skewing younger and more female

Reply:

We thank the reviewer for all the relevant comments that have improved our manuscript significantly.

The prevalence of ALD in relation to age and gender varies geographically and among different cultures, but in Europe, the highest prevalence of ALD with cirrhosis is among men in their 60s (see figure below from Gu et al, [Lancet Reg Health Eur](#)).⁷ As the reviewer points out, the cohort in the study mainly consists of men with a median age in the early 60s, which is representative of the group with the highest prevalence.



However, we acknowledge that this may reduce the generalizability of the results to younger patients, including women. We have therefore added that this is a limitation of our study.

Discussion

Page 17, line 341

“Furthermore, this study included mainly men in the 50s and 60s which may limit generalizability of the results to younger patients, including women”

--did the exclusion criteria also exclude those who would be designated as MASLD or MetALD?

Reply:

Thank you for this relevant question. The Steatotic Liver Disease nomenclature change was published in 2023, long after we designed the protocol. At the time of trial planning, NAFLD was an exclusion diagnosis, reserved for patients who did not have a history of excess drinking. Alcohol-related liver disease patients exhibit a high prevalence of cardiometabolic risk factors, which we also permitted at time of inclusion.⁸ In accordance, nearly all participants presented with at least one such cardiometabolic risk factor at inclusion. Approximately half of the trial cohort reported abstinent at the time of inclusion. They could technically be classified as MASLD according to the new SLD nomenclature⁹, which for now does not account for past alcohol use.¹⁰ However, this may change. For example the recent EASL-EASD-EASO guidelines on MASLD suggests to classify those with a history of excess alcohol consumption as MetALD or ALD.¹¹

To make this choice regarding classification transparent for the reader, we have added the following to the methods and discussion sections, respectively:

Method (the underline part is added)

Page 20 Line 391

“We considered excessive alcohol intake as the etiology if there was a history of excessive alcohol consumption, averaging at least 24 grams per day for women and 36 grams per day for men, sustained for a minimum of 5 years without any other known liver disease. Presence of cardiometabolic risk factors at time of inclusion were permitted”

Discussion

Page 17 Line 352

“It should be noted that the design and execution of this study were completed prior to the introduction of the SLD nomenclature.⁹ At inclusion, nearly all participants exhibited at least one cardiometabolic factor, and 50% reported being alcohol abstinent. These individuals could technically be classified as MASLD according to the SLD nomenclature⁹, as the SLD classification does not account for past alcohol use or potential future increases in consumption.¹⁰ However, over 50% of individuals with a history of high alcohol intake, who are classified as MASLD according to the

SLD nomenclature, subsequently increase their consumption to levels consistent with MetALD/ALD.¹⁰ We therefore chose to label this as a study of patients with ALD."

--the manuscript seems to indicate that Childs Score C patients were enrolled but this seems to be a mistake perhaps?

Reply:

This is a mistake. Patients with Child-Pugh C cirrhosis were not included. However, we have searched for this the manuscript and unfortunately cannot identify the section to which the reviewer refers. During our search, we discovered that the way the exclusion criteria are described could potentially be misunderstood, so we have revised this.

Method

Page 20, line 395

Original

"Compensated liver disease was characterized by no hospitalization within 3 months of inclusion, absence of moderate or severe ascites, high-risk varices requiring intervention, Child-Pugh score of C, and a model for end-stage liver disease-Na (MELD-Na) score of less than 15."

Revised

"To ensure inclusion of only participants with compensated liver disease we excluded participants who had been hospitalised within 3 months of inclusion or who had moderate or severe ascites, high-risk varices requiring intervention, Child-Pugh score of C, or a model for end-stage liver disease-Na (MELD-Na) score of more than 15."

Protocol/Intervention:

--the protocol reviewed appeared slightly different from the manuscript. The product in the protocol (Profermin) is different from ReFerm, which is reported as the intervention product in the manuscript. Can the authors clarify? Are these the same product but just under a different name? They appear to be described as the same thing but with a different name.

Reply:

Thank you for the opportunity to clarify. As correctly assumed, the company producing the product changed the name from Profermin® to ReFerm® in June 2022.

We have added the following in the method

Page 19 line 379

“The intervention product in this manuscript is called ReFerm®, while in previous studies^{12,13} and in the protocol, it was referred to as Profermin®. In June 2022, the company changed the name of the intervention product, but the product composition remains the same. Similarly, the bacterial name reported as part of the intervention product is labelled Lactobacillus plantarum DSM 9843, whereas in the protocol it appears as Lactobacillus plantarum 299v® (Brand name owned by Probi AB, Sweden). The product remains the same and the change has only been done for proprietary reasons.”

--the control cohort is getting a nutritional supplement which in and of itself may improve the outcomes (as nutrition is a key modulator of many features of ALD). One wonders if perhaps comparing to patients eating a standard Danish diet might have been more helpful and pragmatic.

Reply:

We acknowledge the challenges associated with food supplement trials¹⁴, and that a pragmatic design with a control group consuming regular Danish diet would have been simpler. However, as the reviewer notes, it is generally accepted that receiving a nutritional supplement in itself appears to improve outcomes, at least in patients with cirrhosis.¹⁵ Therefore, we prioritized minimizing the risk that any positive effect would be due to nutrition support alone. Consequently, we chose to use an active comparator (Fresubin®). We believe this limitation has been sufficiently addressed in the 'limitations' section.

Page 16, line 320 (unrevised)

“This study has limitations. First, we applied a trial design using Fresubin® as an active comparator, which is a product used for general nutritional support in patients with liver disease in Denmark. The rationale for using an active comparator was that many patients with advanced liver disease have compromised nutritional status and consequently nutritional support in general improves outcome.^{15,16} Therefore, it appears plausible that treatment with Fresubin® may have been beneficial for the patients. Consequently, the therapeutic effects of ReFerm® should be interpreted as improvements beyond the baseline improvements attributed to general nutritional support.”

Outcomes/Measures:

--for the primary outcome, can the authors please indicate why 10% or greater absolute reduction in aSMA expression (marker for hSC activation) was chosen as the threshold? Is there some data that suggests that a 10% reduction is significant or would be clinically meaningful? There is some discussion of this on page 12 lines 227-

236 but it was a bit confusing and didn't really explain why 10% would be a meaningful threshold. And was this reduction level influenced by power calculations?

Reply:

Thank you for pointing out that the rationale for selecting the threshold for the primary endpoint was not clearly presented. We acknowledge that the use of both absolute and relative values in the original manuscript may have been confusing, so we have clarified this in the revised version. The primary endpoint was chosen based on a randomized controlled trial¹⁷ which investigated the effects of candesartan (an angiotensin II receptor antagonist) for patients with compensated ALD. In that study, α -SMA expression was measured to quantify hepatic stellate cell activation, and a significant absolute mean change of -5% ($\pm 7\%$) (from 28.7% (± 10.5) to 23.9% (± 10.3)) in α -SMA was observed, equivalent to a relative decrease of 17%. Considering the standard deviation, we determined that an absolute change of 10% would be clinically relevant.

In support of this, a 2021 publication by Sanyal et al. examined cirrhosis regression in two large NASH trials, focusing on histological fibrosis assessments, including changes in α -SMA.¹⁸ In patients with cirrhosis regression, the mean absolute change in α -SMA was -5.9% (95% CI: -7.2 to -4.5), while those without regression showed an increase of 0.6 (95% CI: 0.0 to 1.2), $p < 0.0001$. Based on these findings, we continue to believe that an absolute decrease of 10% in α -SMA-positive cells is a relevant clinical threshold.

We have now revised the text, to make it clearer that the power calculation is based on an expected absolute reduction of at least 10%.

Methods (sample size calculation):

Page 25 line 514

Original:

"Here, a mean reduction of 17% in α -SMA expression was reported in the intervention group.¹⁷ Anticipating a clinically relevant difference in the proportion of responders ($\geq 10\%$ absolute reduction in α -SMA) from 15% for the treated with Fresubin® and 60% for the treatment with ReFerm®. Accounting for an expected drop-out rate of 20%, α of 5%, and a power of 80%, 40 patients were needed in the study. "

Revised

"Here, an absolute reduction of 5% ($\pm 7\%$) in α -SMA expression was reported in the intervention group, equivalent to a relative decrease of 17% as described in the protocol.¹⁷ Due to the standard deviation of $\pm 7\%$, we decided that to consider an absolute reduction in α -SMA of $\geq 10\%$ as a clinically relevant response. Anticipating a clinically relevant difference in the proportion of responders ($\geq 10\%$ absolute reduction in α -SMA) we anticipated that 15% of participants treated with Fresubin® would achieve a clinically relevant α -SMA response, while 60% of participants treated with

ReFerm® would achieve the response. Accounting for an expected drop-out rate of 20%, α of 5%, and a power of 80%, 40 patients were needed in the study.

Discussion

Page 14, line 274

“An analysis of two RCTs involving 1,135 patients with MASH and cirrhosis showed that α -SMA decreased by -6% in those who experienced cirrhosis regression, compared to an increase of +0.6% in those who did not experience cirrhosis regression.¹⁸ Conversely, an increase in α -SMA from baseline was associated with liver-related events (HR 1.18, 95% CI 1.05 to 1.32).”

--was PETH measured throughout the trial or just at baseline?

Reply:

This is a relevant question. Since the submission of the original version of this manuscript, we have also measured PETH in whole blood at baseline, 4 weeks, and 24 weeks corresponding to the time points when whole blood was available. We have incorporated these analyses into the results section.

Page 6, line 130

Original

“During the trial, the alcohol intake was comparable between the groups with a median alcohol intake of 32 (IQR 24-39) gram/day (Figure S2A).”

Revised (underlined added)

“During the trial, the self-reported alcohol intake was comparable between the groups with a median alcohol intake of 32 (IQR 24-39) gram/day (Figure S2A). PEth measured at baseline, 4 weeks, and 24 weeks were also comparable between groups (Table SX).”

	Total	ReFerm®	Fresubin®	p-value
PEth, baseline	0.20 (0.00-1.00)	0.16 (0.00-0.89)	0.20 (0.00-1.28)	0.57
PEth, 4 weeks	0.11 (0.00-1.03)	0.23 (0.00-1.03)	0.11 (0.00-1.21)	0.68
PEth, 24 weeks	0.45 (0.00-1.46)	0.31 (0.02-1.46)	0.70 (0.00-1.51)	0.88
Supplementary Table 1: Comparing median (IQR) PEth values between the intervention groups through the trial. P-values are derived using Wilcoxon rank-sum test.				

Results:

--50% of each group was abstinent 1 week prior to inclusion. Does this mean that the other 50% were actively drinking?

Reply:

Yes, 50% reported to be actively drinking at time of inclusion. Among the participant reporting abstinence 1 week prior to inclusion, 1 (2.4%) reported relapse to active alcohol intake during the trial. We have clarified and added these data to the results:

Page 6, line 125

"In the per protocol population the participants reporting abstinence 1 week prior to inclusion, 1 (2%) reported alcohol intake during the trial."

---it appears that, by PP and ITT analysis, the primary outcome was not met but that this may have been related to non-compliance. Compliance was reported as 92% for the intervention cohort and 99% for the control cohort. How was compliance assessed? Also, this seems like a quite high compliance rate overall so it may be more difficult to say that noncompliance was the reason for failure.

Reply:

We acknowledge that the initial reporting on compliance may not have been completely clear. Compliance was monitored during four in-hospital visits by assessing the lids of consumed products. Additionally, nine follow-up phone calls were made between visits to gather self-reported compliance data. During each compliance assessment, study personnel also evaluated the presence of any side effects. In cases where patients demonstrated low compliance due to side effects or personal reasons, they either chose to withdraw from the study or were advised to do so by the study team. The compliance data reported reflects only those patients who completed the study. Accordingly, we have revised the text for clarity as follows:

Method:

Page 21, line 417

Original

"Compliance was assessed at four in-hospital visits and further nine phone calls (Figure S1)."

Revised (underlined added)

"Compliance was monitored during four in-hospital visits through self-reporting compliance and by counting the lids of consumed products. If there was a discrepancy between the reported consumption and the counted lids, patients were asked to explain. In the absence of a satisfactory explanation, the counted lids were used as

the measure of compliance. Additionally, nine follow-up phone calls were made between visits to gather self-reported compliance data. During each compliance assessment, study personnel also evaluated the presence of any side effects. If patients frequently forgot to consume the product, they were offered additional reminders via phone calls or text messages. If patients experienced difficulties with side effects resulting in decreased compliance, the study coordinator considered discontinuing the patient's participation (Figure S1)."

--alcohol intake is a very significant confounder for the outcomes assessed. It appears that the median of the group was 32 g/day (about 2-3 drinks per day), but there were some that were abstinent. How did results differ between an abstinent cohort vs an actively drinking cohort? Either way, that data would be interesting to note as an intervention that helped improve liver function with or without drinking would be quite beneficial.

Reply:

We acknowledge that abstinence could likely affect the results of the study in line with several previous studies.¹⁹ In response to this comment, we conducted a subgroup analysis, specifically comparing alcohol abstinence between the treatment groups using both self-reported alcohol consumption and PEth levels. Our findings indicated no difference in abstinence between the two groups:

We have added following in the discussion:

Page 15, line 295

"We conducted a subgroup analysis to assess the impact of alcohol abstinence. Among the participants who completed the study, 17 self-reported maintaining abstinence throughout the study: 8 in the ReFerm® group and 9 in the Fresubin® group. Of these 17 participants, 11 had PEth measurements consistently below 0.05 µmol/L at baseline, 4 weeks, and after 24 weeks of treatment, with 4 in the ReFerm® group and 7 in the Fresubin® group. There was a higher proportion of participants with both low and high PEth values achieving the primary endpoint in the group treated with ReFerm® compared to Fresubin®, although this difference was not statistically significant (Table S8). It is well established that alcohol consumption impacts the prognosis of ALD.¹⁹ This subgroup analysis suggests that the effect of ReFerm® on α-SMA was not mediated by lower alcohol intake."

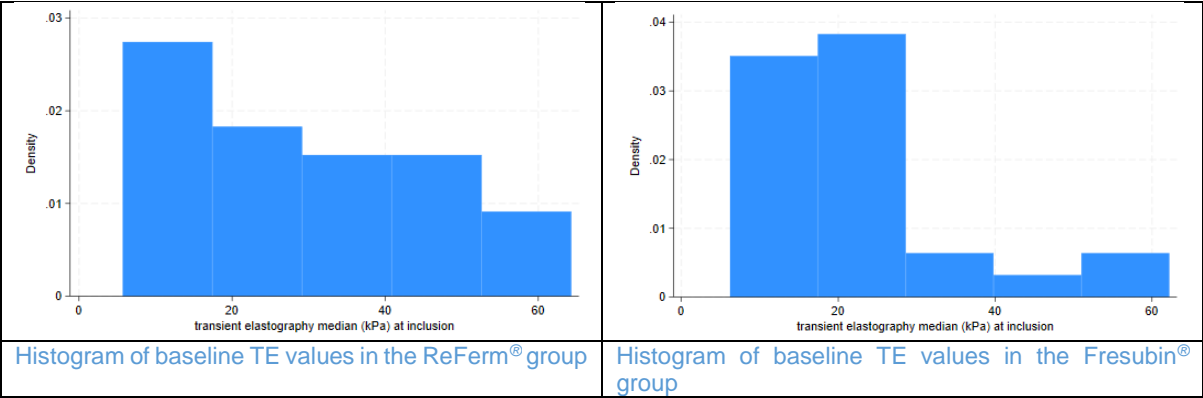
	ReFerm®		Fresubin®	
	Responders n=8	Non-responders n=13	Responders n=4	Non-responders n=15
PEth <0.05, n	2	2	2	5
PEth >0.05, n	6	11	2	10
Supplementary Table S8: PEth measurements in responders vs. non-responders. Subgroup analysis of number of participants achieving the primary endpoint (≥10%				

reduction in α -SMA expression) according to PEth values at baseline as indicators for alcohol abstinence at time of inclusion. “Responders” are defined as patients who meet the primary endpoint of.

--ReFerm group had higher baseline TE scores (27 vs 20 kPa). ELF and Fibrosis scores were more comparable so one wonders why this difference is present.

Reply:

We recognize the observed differences in TE measurements at baseline when compared to other fibrosis assessment modalities. However, histology, α -SMA and all other non-invasive tests were balanced. Additionally, the interquartile range of TE (27.5 with IQR 16.5-46.2 kPa, versus 20.2 with IQR 13.4-27.0 kPa, $p= 0.098$) indicates that the difference is especially carried by more patients with high measurements.



TE is known to exhibit a high variability, exceeding 30%, especially when investigating participants with advanced liver disease.²⁰ Furthermore, for TE values above 20-25 kPa, TE no longer holds prognostic information as development of liver related complications is driven by factors outside mere fibrosis (dysimmunity, coagulation defects, portal hypertension, subclinical inflammation). Therefore, and especially given comparability of the other non-invasive fibrosis and liver function tests, we believe that the randomization procedure achieved comparable intervention groups.

We have addressed the point of a difference in TE measurements in the limitations section of our analysis.

Page 17, line 336

“Although patients were randomly allocated, there was an observed higher baseline liver stiffness measurement in the ReFerm® group compared to the Fresubin® group. This discrepancy may influence the interpretation of liver stiffness changes over time. However, histology and other non-invasive tests, such as the ELF score and FIB-4, were more consistent between the groups at baseline, suggesting that the overall liver fibrosis stage was comparable.”

Reviewer #3:

Thank you for the opportunity to revise this interesting work comparing two randomized arms, ReFerm and Fresubim. The manuscript is well written and many outcomes are assessed. Please find below my comments:

- Abstract: authors report an 8.3% reduction in α -SMA in patients under ReFerm compared to Fresubim. I suggest to report in the abstract the changes within the two arms instead of difference only, in order to show magnitude and direction of changes.

Reply:

We appreciate the suggestion. However, due to the format restrictions for abstracts in Nature Communications, which do not allow the inclusion of specific data, we were unable to implement this change. As a result, we have generally revised the abstract to align with the format restrictions

- Page 9: authors report that compliance was assessed but It is not clear how it was assessed in terms of type of variable or variables collected.

Reply:

We recognize that the reporting of compliance was not entirely clear as also pointed out by Reviewer 2. Consequently, we have clarified the text in Method as follows:

Page 21, Line 417

Original

"Compliance was assessed at four in-hospital visits and further nine phone calls (Figure S1)."

Revised (underlined added)

"Compliance was monitored during four in-hospital visits through self-reporting compliance and by counting the lids of consumed products. If there was a discrepancy between the reported consumption and the counted lids, patients were asked to explain. In the absence of a satisfactory explanation, the counted lids were used as the measure of compliance. Additionally, nine follow-up phone calls were made between visits to gather self-reported compliance data. During each compliance assessment, study personnel also evaluated the presence of any side effects. If patients frequently forgot to consume the product, they were offered additional reminders via phone calls or text messages. If patients experienced difficulties with side effects resulting in decreased compliance, the study coordinator considered discontinuing the patient's participation (Figure S1)."

- A general comment is about adjustment. As stated at page 12, results were reported without adjustment unless specified. It would be helpful to more clearly report why no possible confounders were taken into account. Additionally, at page 15 compliance

adjustment was done for α -SMA mean reduction but It is not clear if adjustment was assessed also for the primary endpoint ($\geq 10\%$ absolute reduction).

Reply:

We acknowledge that adjusting for potential baseline confounders is generally recommended.²¹ However, such confounders should be selected prospectively and outlined in the study protocol.²¹ In this study, the most important feature is alcohol abstinence at baseline, which was also used for stratification. Since the protocol and statistical analysis plan prospectively specified that the primary analyses would be conducted without adjustment, and because there was balance between the two groups on this parameter (Line 126: 'Abstinence from alcohol one week prior to inclusion was reported by 14 of 28 (50%) in each group'), no further adjustment was made. We have specified this in the method section:

Method section

Page 26, line 538

“Adjusting for baseline confounders is generally recommended, but these should be prospectively specified in the statistical analysis plan.²¹ In the statistical analysis plan of our study it was specified that primary analyses would be done unadjusted. However, alcohol abstinence at baseline must be considered the most important potential confounder, which is why stratification was based on this parameter. At inclusion there was balance between the groups, no further adjustments were made.”

- Page 14: the authors state that variables in Table 1 were comparable but no balance measures such as standardized mean difference are reported. Why?

Reply:

We are not entirely certain of what is requested here. The standardized mean difference is typically calculated in connection with effect estimates, while Table 1 is an overview of baseline characteristics. If the suggestion is to perform a statistical calculation to support the statement that baseline characteristics were comparable between the groups, we would like to point out that It is in general not recommended to perform significance testing of baseline differences in randomized controlled trials, because the randomization itself should ensure that any differences between groups are due to chance rather than systematic bias.^{22,23} Performing significance tests on baseline variables can lead to misinterpretations and may undermine the validity of the randomization process.²² For this reason, we have decided not to make changes to the manuscript.

- Figure S2A: please report in the legend what is shown in the figures. Median(IQR)?

Reply:

Thank you for pointing this out. We have now added the missing details (underlined):

“Supplementary Figure 2: Changes in alcohol intake and BMI. (A) Changes in self-reported alcohol intake during the treatment period (Median, IQR). (B) Changes in BMI during the treatment period (Mean, SD).”

- How was normality assessed?

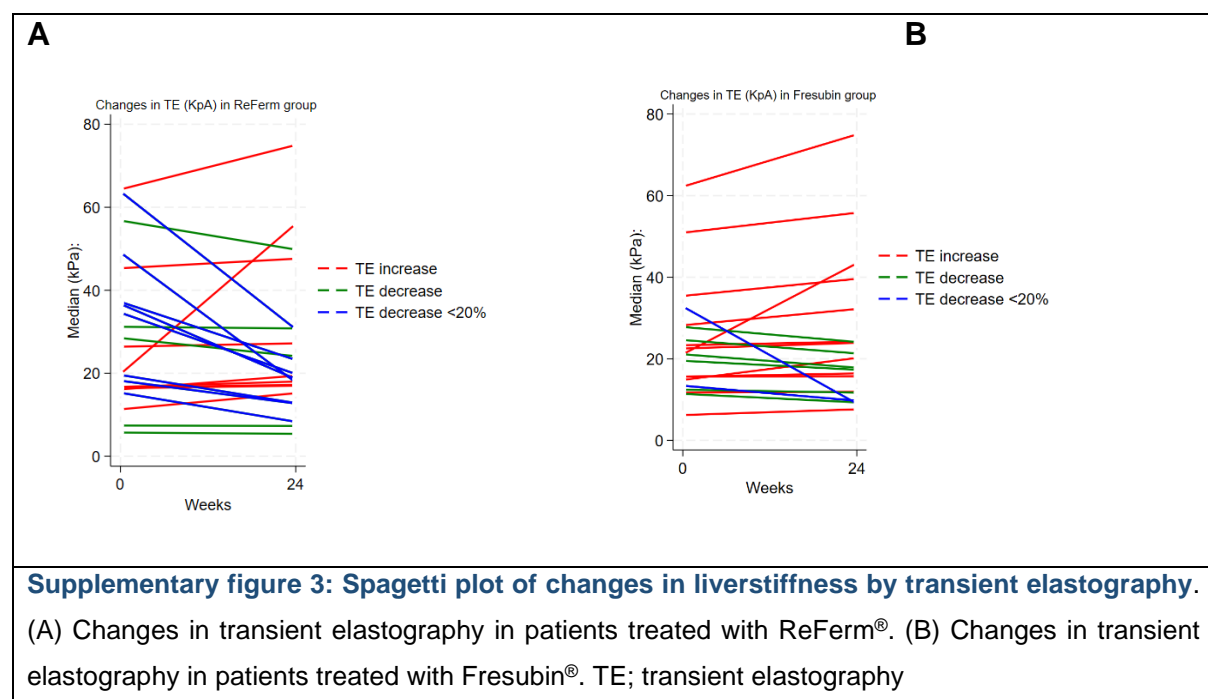
Reply:

We assessed normality using both a histogram and a Q-Q plot. In cases of uncertainties or discrepancies, the Shapiro-Wilk test was used. The choice of subsequent statistical methods for analyses was based on the results of this test.

- I am a bit confused with Figure S3. What is the difference between the two figures?

Reply:

Thank you for bringing this error to our attention. We have now corrected the figure for the Fresubin® group.



- For variables collected at several time points, did the authors consider to perform statistical models for longitudinal analysis?

Reply:

We acknowledge, that this could be a way to improve the power of the statistical models. However, in this dataset, only the microbiome data (and now also PEth) has more than two time points, which is necessary to apply statistics for longitudinal models. Regarding the microbiome data, we chose to present the changes over time for each time point separately to keep it the interpretation as simple as possible.

- Page 19, line 349: the authors state that a reduction in α -SMA was strongly correlated with I-FABP and NT-3. Do they really refer to correlation coefficients?

Reply:

We appreciate the reviewer's feedback and suggestion for improvement. We have revised the results section accordingly and included the specific rho to prevent any further misunderstandings. Furthermore, we have included a section in the discussion.

Results

Page 11, Line 209

Old version:

"The association between changes in the primary outcome (α -SMA) and the selected significant features detected from the omics analyses was explored in a combined analysis of α -SMA, Lactobacillus plantarum, and cytokines I-FABP, SIRT2, and NT-3 (Figure 4D). This analysis showed that a reduction of α -SMA was strongly correlated with a reduced marker for impaired gut barrier (I-FABP) and an increased marker for hepatic regeneration (NT-3). Furthermore, the analysis showed that high levels of L. plantarum, as a marker for treatment with ReFerm®, were strongly associated with reductions in I-FABP and α -SMA together with an increase of NT-3."

New version:

"The association between changes in the primary outcome (α -SMA) and the selected significant features detected from the omics analyses were explored in a combined analysis of α -SMA, Lactobacillus plantarum DSM 9843, and cytokines I-FABP, SIRT2, and NT-3 (Figure 4C). This analysis showed that a reduction of α -SMA was correlated with a reduction in I-FABP (rho: 0.27) and an increase of SIRT2 (rho: 0.34), L. plantarum (rho: 0.23), and hepatic regeneration (NT-3 rho: 0.14). Furthermore, the analysis showed that high levels of L. plantarum, as a marker for treatment with

ReFerm®, were correlated with reduction in I-FABP (ρ : -0.17) together with an increase of NT-3 (ρ 0.48) and SIRT2 (ρ : 0.37)."

Discussion (added):

Page 15, line 286

"IL-17A is important for neutrophil recruitment and augmentation of antibacterial responses to pathogenic bacteria and has been found to increase in patients with hepatic encephalopathy treated with Rifaximin.^{24,25} In the ReFerm® group, we observed stability in IL-17A levels, whereas there was a decrease in the Fresubin® group. Leptin is known to promote inflammation.²⁶ In the ReFerm® group, leptin levels remained stable, while the Fresubin® group experienced an increase. These results suggest a deterioration in inflammation in the Fresubin® group while stable in the ReFerm® group."

References

1. Trinh VQ, Lee TF, Lemoine S, et al. Hepatic stellate cells maintain liver homeostasis through paracrine neurotrophin-3 signaling that induces hepatocyte proliferation. *Sci Signal* 2023; **16**(787): eadf6696.
2. Yang Z, Zhang H, Yin M, et al. Neurotrophin3 promotes hepatocellular carcinoma apoptosis through the JNK and P38 MAPK pathways. *Int J Biol Sci* 2022; **18**(15): 5963-77.
3. Zhang Y, Long X, Ruan X, et al. SIRT2-mediated deacetylation and deubiquitination of C/EBP β prevents ethanol-induced liver injury. *Cell Discov* 2021; **7**(1): 93.
4. Brol MJ, Rösch F, Schierwagen R, et al. Combination of CCl₄ with alcoholic and metabolic injuries mimics human liver fibrosis. *Am J Physiol Gastrointest Liver Physiol* 2019; **317**(2): G182-g94.
5. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 2008; **3**(6): 1101-8.
6. Lee MJ. A review of liver fibrosis and cirrhosis regression. *J Pathol Transl Med* 2023; **57**(4): 189-95.
7. Gu W, Hortlik H, Erasmus HP, et al. Trends and the course of liver cirrhosis and its complications in Germany: Nationwide population-based study (2005 to 2018). *Lancet Reg Health Eur* 2022; **12**: 100240.
8. Israelsen M, Juel HB, Detlefsen S, et al. Metabolic and Genetic Risk Factors Are the Strongest Predictors of Severity of Alcohol-Related Liver Fibrosis. *Clin Gastroenterol Hepatol* 2020.
9. Rinella ME, Lazarus JV, Ratziu V, et al. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Hepatology* 2023; **78**(6): 1966-86.
10. Israelsen M, Torp N, Johansen S, et al. Validation of the new nomenclature of steatotic liver disease in patients with a history of excessive alcohol intake: an analysis of data from a prospective cohort study. *Lancet Gastroenterol Hepatol* 2024; **9**(3): 218-28.
11. EASL-EASD-EASO Clinical Practice Guidelines on the Management of Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD). *Obes Facts* 2024; **17**(4): 374-444.
12. Krag A, Munkholm P, Israelsen H, von Ryberg B, Andersen KK, Bendtsen F. Profermin is efficacious in patients with active ulcerative colitis--a randomized controlled trial. *Inflamm Bowel Dis* 2013; **19**(12): 2584-92.
13. Krag A, Israelsen H, von Ryberg B, Andersen KK, Bendtsen F. Safety and efficacy of Profermin® to induce remission in ulcerative colitis. *World J Gastroenterol* 2012; **18**(15): 1773-80.
14. Staudacher HM, Yao CK, Chey WD, Whelan K. Optimal Design of Clinical Trials of Dietary Interventions in Disorders of Gut-Brain Interaction. *Am J Gastroenterol* 2022; **117**(6): 973-84.
15. Fialla AD, Israelsen M, Hamberg O, Krag A, Gluud LL. Nutritional therapy in cirrhosis or alcoholic hepatitis: a systematic review and meta-analysis. *Liver Int* 2015; **35**(9): 2072-8.
16. Bischoff SC, Bernal W, Dasarthy S, et al. ESPEN practical guideline: Clinical nutrition in liver disease. *Clin Nutr* 2020; **39**(12): 3533-62.
17. Kim MY, Cho MY, Baik SK, et al. Beneficial effects of candesartan, an angiotensin-blocking agent, on compensated alcoholic liver fibrosis - a randomized open-label controlled study. *Liver Int* 2012; **32**(6): 977-87.
18. Sanyal AJ, Anstee QM, Trauner M, et al. Cirrhosis regression is associated with improved clinical outcomes in patients with nonalcoholic steatohepatitis. *Hepatology* 2022; **75**(5): 1235-46.

19. Louvet A, Bourcier V, Archambeaud I, et al. Low alcohol consumption influences outcomes in individuals with alcohol-related compensated cirrhosis in a French multicenter cohort. *J Hepatol* 2023; **78**(3): 501-12.
20. Pierce TT, Ozturk A, Sherlock SP, et al. Reproducibility and Repeatability of US Shear-Wave and Transient Elastography in Nonalcoholic Fatty Liver Disease. *Radiology* 2024; **312**(3): e233094.
21. Holmberg MJ, Andersen LW. Adjustment for Baseline Characteristics in Randomized Clinical Trials. *Jama* 2022; **328**(21): 2155-6.
22. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Bmj* 2010; **340**: c869.
23. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Bmj* 2010; **340**: c332.
24. Edwards LA, Nistala K, Mills DC, et al. Delineation of the innate and adaptive T-cell immune outcome in the human host in response to *Campylobacter jejuni* infection. *PLoS One* 2010; **5**(11): e15398.
25. Patel VC, Lee S, McPhail MJW, et al. Rifaximin- α reduces gut-derived inflammation and mucin degradation in cirrhosis and encephalopathy: RIFSYS randomised controlled trial. *J Hepatol* 2022; **76**(2): 332-42.
26. Fischer J, Gutiérrez S, Ganesan R, et al. Leptin signaling impairs macrophage defenses against *Salmonella Typhimurium*. *Proc Natl Acad Sci U S A* 2019; **116**(33): 16551-60.

Reply letter, 2nd revision

Reviewer #3 (Remarks to the Author):

The authors have addressed my suggestions or justified their choices and they updated the manuscript accordingly.

Reviewer #4 (Remarks to the Author):

The revised version of the article add more details to a delicate and interesting topic: reversal of liver fibrosis in ALD.

The authors sensibly tried to upgrade and update the article. Some point needs to be clarified:

- The data on the animal model are of interest. Indeed, they are limited. I think this point " limited data " should be pointed out by authors also in the Discussion section, subsection Limits of the study. Furgether, larger data are needed to confirm proof of concept data.

Reply:

We acknowledge the reviewer's concern regarding the small sample size of the animal data. However, it is important to highlight that these data were included at the request of Reviewer #1 as proof of concept to support the findings from the human study. Accordingly, we have added the following to the discussion:

Page 18, line 347

"The small sample size in the animal model may limit the robustness of the findings, and a larger dataset is required could further elucidate the underlying mode of action."

- the population in study is variegated and the representation of fibrosis is not " normal ". This may have affected significance of results. I would suggest to highlight the point in the limitations. Moreover, I think that when the cohorts are enlarged, the effect on advanced fibrosi could be milder. Could the authors comment on this ?

Reply

We acknowledge that the severity of fibrosis in our cohort ranges from significant fibrosis to compensated cirrhosis. The prevalence of milder degrees of fibrosis is likely higher in other cohorts, depending on the population from which the cohort is recruited. Accordingly, we have added the underlined to the discussion.

Page 16, Line 322

"This study design included patients with ALD ranging from significant fibrosis (F2) to compensated cirrhosis (F4)... ...which may have led to a heterogeneous effect of the

interventions and could potentially limit the generalizability to people with less severe liver fibrosis"

- the Fresubin control arm: can it have biased results ? has it just nutritional impact ? can the increased and ameliorated protein pool 8if any) affect intestinal permeability parameters ? other measured variables ? Please, comment on this.

Reply:

This is a relevant observation. There is no immediate indication in the literature that protein has a negative effect on intestinal permeability¹, although the topic is not particularly well studied. What seems more likely is that Fresubin® may have been beneficial for patients with liver disease, as nutritional support appears to improve outcomes.^{2,3} We have discussed this issue under the limitations in this section.

Page 16, Line 310

First, we applied a trial design using Fresubin® as an active comparator, which is a product used for general nutritional support in patients with liver disease in Denmark. The rationale for using an active comparator was that many patients with advanced liver disease have compromised nutritional status and consequently nutritional support in general improves outcome.^{2,3} Therefore, it appears plausible that treatment with Fresubin® may have been beneficial for the patients. Consequently, the therapeutic effects of ReFerm® should be interpreted as improvements beyond the baseline improvements attributed to general nutritional support.

- an English language revision and, final, text cleaning is warranted.

Reply:

Thank you for the suggestion. We have made every effort to ensure clarity at this stage, but further editing and proofreading will, of course, be conducted prior to final publication.

- is 10% effect tight for ReFerm (reduction in αSMA expression) ?

Reply:

It is a relevant question to question what a clinically meaningful reduction of α-SMA expression is. The 10% reduction in α-SMA expression aligns with the predefined primary outcome of the study, which was a ≥10% absolute reduction in the percentage of α-SMA-positive cells from baseline to 24 weeks of treatment. This endpoint was chosen based on the role of α-SMA as a marker of activated hepatic stellate cells (HSCs), which are critical mediators of fibrogenesis and liver collagen accumulation. Thus, a 10% reduction is considered clinically meaningful in the context of this study. Please also see the Method section page 22 line 435

"The primary outcome was a between-group comparison of histological reduction of activated HSCs, defined as ≥10% absolute reduction in the percentage positive for α-SMA expression, from baseline to 24 weeks of treatment. The rationale for choosing

this endpoint was that α -SMA is a marker of activated HSCs. Activated HSCs are the key cellular components in liver collagen accumulation, driving fibrogenesis.⁴⁻⁶

- Hot point: the active drinking is a bias of the study and, also, a real-life data: how can this affect and, possibly, counteract the ReFerm effects ? Please, comment on this.

Reply:

We completely agree that this is a crucial point. Changes in alcohol intake are well-established as significant modifiers of disease progression. To address this, we performed a subgroup analysis using both subjective self-reported data and the objective alcohol biomarker phosphatidylethanol. Unfortunately, this part of the results section was mistakenly included in the discussion section. We sincerely apologise for this oversight, as it should have been presented in the results section. We believe this addresses the reviewer's point and have removed it from the discussion and inserted it into the results section.

Please see page 8, line 151:

"We conducted a subgroup analysis to assess the impact of alcohol abstinence. Among the participants who completed the study, 17 self-reported maintaining abstinence throughout the study: 8 in the ReFerm® group and 9 in the Fresubin® group. Of these 17 participants, 11 had PEth measurements consistently below 0.05 μ mol/L at baseline, 4 weeks, and after 24 weeks of treatment, with 4 in the ReFerm® group and 7 in the Fresubin® group. There was a higher proportion of participants with both low and high PEth values achieving the primary endpoint in the group treated with ReFerm® compared to Fresubin®, although this difference was not statistically significant (Table S8). It is well established that alcohol consumption impacts the prognosis of ALD.⁷ This subgroup analysis suggests that the effect of ReFerm® on α -SMA was not mediated by lower alcohol intake."

- Thus, non-compliance has affected results' significance (namely, primary endpoint) ? Is it the only explanation ? Please, comment on that.

Reply:

We recognise that compliance seems to have significantly influenced the primary endpoint in the intervention group. This likely reflects a positive control, indicating that sufficient intake of the intervention product is necessary for efficacy. However, alternative explanations cannot be excluded. The most plausible alternative is that non-compliance may have been associated with higher alcohol consumption, potentially explaining the lack of effect. Nevertheless, as noted above, our subgroup analyses confirmed that the effect of ReFerm® on α -SMA was not mediated by reduced alcohol intake. To highlight this in the manuscript, we have expanded the assessment as follows:

Page 17, Line 333

"Notably, α -SMA expression showed a significant decrease when adjusting for compliance indicating that sufficient intake of the intervention product is necessary for efficacy. However, alternative explanations cannot be excluded. The most plausible alternative is that non-compliance may have been associated with higher alcohol consumption, potentially explaining the lack of effect. Nevertheless, subgroup analyses confirmed that the effect of ReFerm® on α -SMA was not driven by reduced alcohol intake"

1. Khoshbin K, Camilleri M. Effects of dietary components on intestinal permeability in health and disease. *Am J Physiol Gastrointest Liver Physiol* 2020; **319**(5): G589-g608.
2. Fiolla AD, Israelsen M, Hamberg O, Krag A, Gluud LL. Nutritional therapy in cirrhosis or alcoholic hepatitis: a systematic review and meta-analysis. *Liver Int* 2015; **35**(9): 2072-8.
3. Bischoff SC, Bernal W, Dasarathy S, et al. ESPEN practical guideline: Clinical nutrition in liver disease. *Clin Nutr* 2020; **39**(12): 3533-62.
4. Kisseleva T, Cong M, Paik Y, et al. Myofibroblasts revert to an inactive phenotype during regression of liver fibrosis. *Proc Natl Acad Sci U S A* 2012; **109**(24): 9448-53.
5. Carpino G, Morini S, Ginanni Corradini S, et al. Alpha-SMA expression in hepatic stellate cells and quantitative analysis of hepatic fibrosis in cirrhosis and in recurrent chronic hepatitis after liver transplantation. *Dig Liver Dis* 2005; **37**(5): 349-56.
6. Kim MY, Cho MY, Baik SK, et al. Beneficial effects of candesartan, an angiotensin-blocking agent, on compensated alcoholic liver fibrosis - a randomized open-label controlled study. *Liver Int* 2012; **32**(6): 977-87.
7. Louvet A, Bourcier V, Archambeaud I, et al. Low alcohol consumption influences outcomes in individuals with alcohol-related compensated cirrhosis in a French multicenter cohort. *J Hepatol* 2023; **78**(3): 501-12.