UNIVERSITY OF THESSALY

SCHOOL OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE AND BIOMEDICAL INFORMATICS

Meta-analysis of gastrointestinal microbiome data in relation to the consumption of natural and artificial sweeteners

Thesis submitted by:

Lydia Panagiota Kalomiris

Bachelor of Science (B.Sc.) in Computer Science and Biomedical informatics

DIPLOMA THESIS

Supervisor

Georgia Braliou,

Associate professor

UNIVERSITY OF THESSALY

SCHOOL OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE AND BIOMEDICAL INFORMATICS

Meta-analysis of gastrointestinal microbiome data in relation to the consumption of natural and artificial sweeteners

Lydia Panagiota Kalomiris

DIPLOMA THESIS

Supervisor

Georgia Braliou,

Associate professor

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις [1], που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί ***χωρίς να τα περικλείω σε εισαγωγικά*** *και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάσθηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.*

Δέχομαι ότι η αυτολεξεί ***παράθεση χωρίς εισαγωγικά****, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.*

Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια

4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία: …../…/20…

Ο – Η Δηλ.

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

Meta-analysis of gastrointestinal microbiome data in relation to the consumption of natural and artificial sweeteners

Lydia Panagiota Kalomiris

Committee: **Greta Braliou**, Associate Professor, Department of Computer Science and Biomedical Informatics, University of Thessaly (supervisor)

Panagiota Kontou, Assistant Professor, Department of Mathematics, University of Thessaly

Pantelis Bagos, Professor, Department of Computer Science and Biomedical Informatics, University of Thessaly

Μετα-ανάλυση δεδομένων στομαχικού μικροβιώματος σε σχέση με την κατανάλωση φυσικών και τεχνητών γλυκαντικών

Λυδία Παναγιώτα Καλομοίρη

Επιτροπή: **Γεωργία Μπράλιου**, (επιβλέπουσα) Associate Professor, Department of Computer Science and Biomedical Informatics, University of Thessaly

Παναγιώτα Κοντού, Assistant Professor, Department of Mathematics, University of Thessaly

Παντελής Μπάγκος, Professor, Department of Computer

Science and Biomedical Informatics, University of Thessaly

Table of Contents

ACKNOWLEDGEMENTS

This thesis was conducted as part of the final stage of my undergraduate studies in the Department of Biomedicine at the University of Thessaly, from January 2024 to June 2025. It marks the first time I have undertaken an independent scientific research project of this scale, and it has been an invaluable learning experience.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Georgia Braliou, for her guidance, patience, and continuous support throughout this process. Her insightful feedback, passion for research, and genuine interest in the project were instrumental in shaping the direction and quality of this work.

I would also like to thank the members of my thesis committee, Prof. Bagos Pantelis and Prof. Panagiota Kontou, at the Department of Computer Science and Biomedical Informatics and Department of Mathematics for their valuable comments and constructive criticism, which helped refine both the scientific and methodological aspects of this thesis.

A special thanks goes to Dr. Argyro Kalogeropoulou, Postdoctoral Researcher at the University of Thessaly, whose consistent support, availability, and valuable guidance contributed significantly to the development of this thesis. Her unwavering encouragement, collaborative spirit, and exemplary work ethic made working with her both inspiring and intellectually enriching.

I am sincerely thankful to my family for their unwavering emotional and financial support throughout my academic journey. Their belief in me made it possible to pursue my goals with determination and confidence.

Finally, I would like to thank my friends for their constant encouragement, patience, and motivation during both the highs and lows of this academic endeavor. Their presence made this experience more manageable and far more enjoyable.

ABSTRACT

INTRODUCTION: The gut microbiome plays a critical role in host health, and its disruption, commonly referred to as dysbiosis, is a hallmark of inflammatory bowel diseases (IBD) such as colitis. Inulin, a natural prebiotic, has been widely studied for its ability to promote beneficial bacteria and restore microbial balance. This study aims to evaluate the effect of inulin on the gut microbiome in murine models of DSS-induced colitis through a meta-analytical approach.

MATERIALS AND METHODS: A literature search was conducted using the PubMed database, yielding 1,144 articles. After applying inclusion and exclusion criteria, 7 studies were deemed suitable for meta-analysis—6 providing data on alpha diversity indices (*Shannon* and *Chao1*) and 5 on the relative abundance of major bacterial phyla (*Firmicutes*, *Bacteroidota*, and *Proteobacteria*). From each study, we extracted information on the model organism, sequencing method, means and standard deviations of alpha diversity indices, as well as relative abundance percentages and corresponding standard errors.

RESULTS: We performed multivariate meta-analysis using standardized mean differences (SMDs) and weighted mean subgroup meta-analysis based on experimental conditions: healthy control, DSS-induced colitis, and DSS-induced colitis with inulin supplementation. Results showed that in the case + inulin condition, both alpha diversity indices and relative abundance values were restored toward levels observed in the healthy control group. The diseased condition consistently exhibited the lowest values, indicating severe dysbiosis. An exception was noted for *Proteobacteria*, which increased progressively across the three conditions, suggesting inulin did not reverse its expansion.

CONCLUSION: In conclusion, the findings indicate that inulin has a beneficial effect on the dysbiotic gut microbiome of DSS-induced colitis in mice. It significantly contributes to the recovery of alpha diversity and enhances the relative abundance of beneficial bacterial phyla such as *Firmicutes* and *Bacteroidota*, supporting its potential as a therapeutic dietary intervention for gut health.

Keywords: *Gut microbiome, prebiotics, inulin, colitis, IBD, meta-analysis, alpha diversity, relative abundance, dysbiosis, bacterial taxa restoration.*

ΠΕΡΙΛΗΨΗ

Η εντερική μικροχλωρίδα διαδραματίζει καθοριστικό ρόλο στην υγεία του ξενιστή, ενώ η διαταραχή της, γνωστή ως δυσβίωση, αποτελεί χαρακτηριστικό γνώρισμα φλεγμονωδών παθήσεων του εντέρου (IBD), όπως η κολίτιδα. Η ινουλίνη, ένα φυσικό πρεβιοτικό, έχει μελετηθεί εκτενώς για την ικανότητά της να προάγει την ανάπτυξη ευεργετικών βακτηρίων και να αποκαθιστά τη μικροβιακή ισορροπία. Η παρούσα μελέτη αποσκοπεί στην αξιολόγηση της επίδρασης της ινουλίνης στη μικροχλωρίδα του εντέρου σε μοντέλα κολίτιδας με χρήση DSS σε ποντίκια, μέσω μετα-ανάλυσης.

Διενεργήθηκε συστηματική αναζήτηση βιβλιογραφίας στη βάση PubMed, από την οποία εντοπίστηκαν 1.144 άρθρα. Μετά από αξιολόγηση κριτηρίων ένταξης και αποκλεισμού, επιλέχθηκαν 7 μελέτες κατάλληλες για μετα-ανάλυση—6 με στοιχεία για δείκτες άλφα ποικιλότητας (Shannon, Chao1) και 5 με δεδομένα σχετικής αφθονίας βασικών μικροβιακών φυλών (Firmicutes, Bacteroidota, Proteobacteria). Από κάθε μελέτη συλλέχθηκαν πληροφορίες σχετικά με το πειραματικό μοντέλο, τη μεθοδολογία αλληλούχισης, το alpha diversity, το relative abundance καθως και τους μέσους όρους, τις τυπικές αποκλίσεις και τα σφάλματα για τις μελετώμενες μεταβλητές.

Πραγματοποιήθηκε πολυμεταβλητή μετα-ανάλυση με χρήση τυποποιημένων μέσων διαφορών (SMD), καθώς και αναλυση σταθμισμένων όρων με υποομαδική μετα-ανάλυση ανά πειραματική κατάσταση: υγιή ομάδα ελέγχου, ομάδα με DSS-επαγόμενη κολίτιδα και ομάδα με κολίτιδα + inulin. Τα αποτελέσματα έδειξαν ότι στην κατάσταση κολίτιδα + ινουλίνη, οι τιμές τόσο των δεικτών ποικιλότητας (alpha diversity) όσο και της σχετικής αφθονίας (relative abundance) αποκαταστάθηκαν σε επίπεδα παρόμοια με αυτά της υγιούς ομάδας. Η ομάδα με κολίτιδα εμφάνιζε σταθερά τις χαμηλότερες τιμές, επιβεβαιώνοντας την ύπαρξη σοβαρής δυσβίωσης. Μοναδική εξαίρεση αποτέλεσε τα Proteobacteria, των οποίων η αφθονία αυξήθηκε προοδευτικά από την υγιή προς την ομάδα με inulin, υποδηλώνοντας ότι η ινουλίνη δεν αναστέλλει τη συγκεκριμένη αύξηση.

Συμπερασματικά, τα ευρήματα υποδεικνύουν ότι η ινουλίνη έχει θετική επίδραση στην αποκατάσταση της εντερικής μικροχλωρίδας σε ποντίκια με DSS-επαγόμενη κολίτιδα. Συμβάλλει σημαντικά στην ανάκτηση της μικροβιακής ποικιλότητας και στην αύξηση των ευεργετικών φυλών Firmicutes και Bacteroidota, επιβεβαιώνοντας την πιθανή χρήση της ως διατροφική παρέμβαση για την εντερική υγεία.

Λέξεις-κλειδιά: *Εντερική μικροχλωρίδα, πρεβιοτικά, ινουλίνη, κολίτιδα, φλεγμονώδεις νόσοι του εντέρου (IBD), μετα-ανάλυση, άλφα ποικιλότητα, σχετική αφθονία, δυσβίωση, αποκατάσταση βακτηριακών ταξινομικών μονάδων.*

1. INTRODUCTION

1.1 Probiotics

The word "probiotic" was most likely created by Ferdinand Vergin in 1954 in his paper "Anti- und Probiotika," which contrasted the negative effects of antibiotics and other antibacterial agents on the intestinal microbiota with the positive effects ("probiotika") of some helpful microorganisms. Probiotics were later defined by Lilly and Stillwell in 1965 as microorganisms that promote the growth of other microbes.

According to the most recent definition, which was developed in 2002 by experts from working groups of the Food and Agriculture Organization of the United Nations (FAO) and the World Health Organization (WHO), probiotics are “live strains of strictly selected microorganisms which, when administered in adequate amounts, confer a health benefit on the host” [1]. In 2013, the International Scientific Association for Probiotics and Prebiotics (ISAPP) upheld the concept [6].

In other words, probiotics are bacteria that promote the natural balance of the gut microbiota and, when given in appropriate amounts, benefit the host's health. Genera Lactobacillus and Bifidobacterium, Saccharomyces, Enterococcus, Streptococcus, Pediococcus, Leuconostoc, and Bacillus are the probiotic strains that have been shown to provide health advantages. The primary probiotic categories are outlined in Figure 1, which separates them into conventional strains such lactic acid bacteria, Bifidobacterium species, and Saccharomyces as well as next-generation candidates like Akkermansia and Christensenella, which are becoming more well-known for their possible health advantages. These bacteria have been shown to modulate the immune system, enhance adhesion to the intestinal mucosa and consequent inhibition of pathogen adhesion, and strengthen the epithelial barrier, and prevent colonization by pathogens. Clinical studies suggest that probiotics may help reduce the symptoms of gastrointestinal conditions like inflammatory bowel disease (IBD) and irritable bowel syndrome (IBS) by regulating inflammation and reestablishing microbial balance. [1]

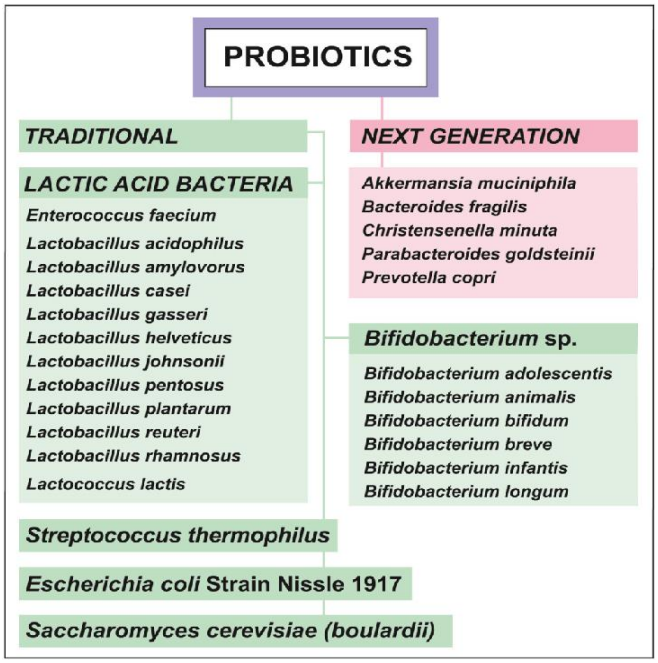


Figure 1 *- Some probiotics classified [24]*

1.2 Prebiotics

Prebiotics are non-digestible food ingredients that are being consumed by probiotics and selectively stimulate the growth and/or activity of those beneficial microorganisms in the gut improving host health in the process. The concept of prebiotics was originally proposed by Glenn R. Gibson and Marcel Roberfroid in 1995.[1]

Unlike probiotics, which are live organisms, prebiotics serve as substrates for resident beneficial bacteria. Figure 2 presents the classification of prebiotics into carbohydrates (including oligosaccharides, disaccharides, and polysaccharides), phytochemicals (notably polyphenols), and other compounds such as proteins, organic acids, and bacterial metabolites, highlighting the diversity of compounds that support beneficial gut microbes.

The most common prebiotics include fructooligosaccharides (FOS), galactooligosaccharides (GOS), and inulin-type fructans (ITFs). Other known types of prebiotics are xylooligosaccharides(XOS), chitooligosaccharides (COS), lactulose, resistant starch (RS), and polyphenols [1]. They contribute to increased production of short-chain fatty acids (SCFAs) such as butyrate, acetate, and propionate, which have anti-inflammatory effects and support gut health. Importantly, prebiotics have been associated with enhanced mineral absorption, improved glycemic control, and immune regulation [6].

The International Scientific Association for Probiotics and Prebiotics (ISAPP) reinterpreted prebiotics in 2016 as a substrate that host microbes selectively use to provide health benefits [1].Lastly, prebiotics were defined by FAO/WHO experts in 2007 as a nonviable food component that benefits the host's health by modulating the microbiota.Prebiotics can be used as a supplement to probiotics or as a substitute for them. Nonetheless, various prebiotics will help with the development of gut microflora.[6]

The introduction of probiotics, prebiotics, or the combination of the two (synbiotics), into a diet is favourable for the intestinal microbiota. They may be consumed in the form of raw vegetables and fruits, fermented pickles, or dairy products. Another source may be pharmaceutical formulas and functional food. [6]

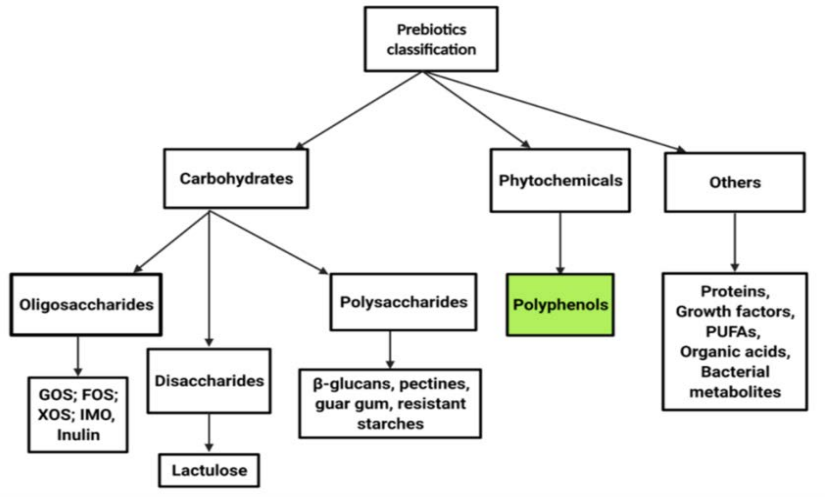


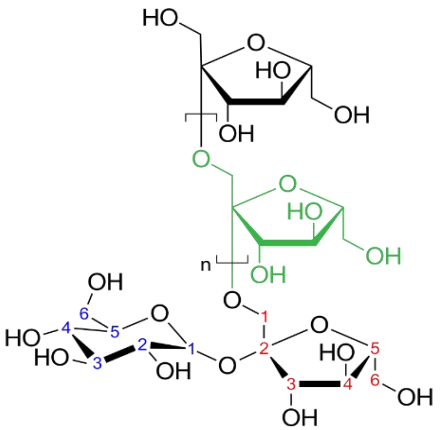
Figure 2 *- Classification of prebiotics [19]*

1.3 Inulin

More than 36,000 plant species contain inulin, a soluble dietary fiber, as a reserve polysaccharide. As seen in Figure 4, Jerusalem artichoke tubers and chicory roots are frequently utilized as raw materials for inulin production in the food sector. Other main sources of inulin include onion, garlic, barley, and dahlia [4]. Inulin is a naturally occurring fructan composed of β(2→1)-linked fructose units with a terminal glucose moiety as illustrated in *Figure 3*. The term "inulin" refers to any β- (2, 1) linear fructan; nevertheless, inulin-type fructans must have β-(2,1) links, which gives inulin its distinct structural and physiological characteristics and makes it resistant to enzymatic hydrolysis by small intestinal digesting enzymes and human saliva [5]. Figure 5 illustrates how different fructans, including inulin, display structural diversity based on the location of fructose addition. These structural variations affect fermentation behavior as well as physiological function. One factor contributing to inulin's distinct metabolic destiny in the colon is its linear structure, as opposed to branching or alternatively connected fructans (Fig.6) [1]. As a dietary fiber, inulin resists digestion in the upper gastrointestinal tract (small intestine) and is consumed by specific gut bacteria in the large intestine. Almost 90% of inulin travels to the colon, where the bacteria digest it there. [1]. This selective fermentation promotes the growth of beneficial bacteria such as *Bifidobacteria* and *Lactobacilli*, contributing to an improved gut microbial profile and having a remarkable impact on gut microbiota regulation [4]. Inulin fermentation results in the production of SCFAs, particularly butyrate, which nourishes colonic epithelial cells and exerts anti-inflammatory properties. Research has shown that inulin supplementation can reduce intestinal permeability and enhance mucosal immunity in animal models with colitis[4] [5].

In addition, inulin also exhibits excellent health benefits in regulating blood lipogenesis, plasma triacylglycerol concentrations and weight loss and blood sugar, and in relieving constipation, reducing the risk of gastrointestinal disease, lowering the risk of colon cancer, enhancing mineral absorption (e.g. calcium, magnesium, iron) and relieving depression [1][4].

Inulinase has the ability to hydrolyze oligofructose from inulin into chains ranging in length from two to ten. Because of this, inulin's sugar chain is longer than oligofructose's, which causes fermentation and gas production to proceed more slowly. Inulin has been widely employed as a texture adjuster, fat and sugar alternative, prebiotic, and in the creation of functional meals.[5].



Inulin's chemical form/composition: $C_{6n}H_{10n+2}O_{5n+1}$

Figure 3 - *Inulin's structural formula (fructan). Colors are used to draw attention to additional characteristics, such as carbon numbers and repetitive fructosyl units. [source: https://en.wikipedia.org/wiki/Inulin]*

Figure 4 - *Inulin's origin products from right to left: Asparagus, Jerusalem artichoke and chicory roots*

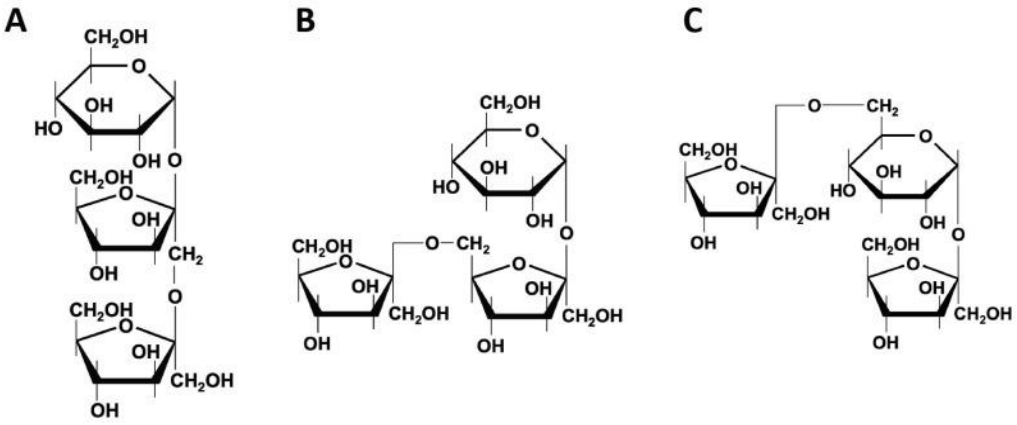


Figure 5 - *Chemical structures of fructans. Differences in fructans structures arise based on the position of fructose addition. [source: doi: 10.3390/ijms25094834]*

Types of fructans.

Fructans	Linkage	Kestose Type
Inulin Fructooligosaccharides	β(2→1)	1-kestose
Neo-inulin	β(2→1)	6G-kestose
Levan	β(2→6)	6-kestose
Mixed levan (graminan)	β(2→1) β(2→6)	1-kestose 6-kestose
Neo-levan	β(2→1) β(2→6)	6G-kestose

Figure 6 - *Types of fructans. [source: doi: 10.3390/ijms25094834]*

1.4 Microbiome

The gut microbiome is a highly diverse and dynamic community of trillions of microorganisms found in the gastrointestinal tract, including mostly bacteria. This complex ecosystem is made up of about 1000 different types of microorganisms that live in the gut and is involved in critical host functions, such as digestion, vitamin synthesis, immune modulation, and protection against pathogens. Most of the microorganisms in our microbiome are beneficial. The main function of beneficial bacteria is to balance out harmful ones in order to maintain gut health. They also provide additional health advantages [12].

However, occasionally the balance between beneficial and harmful bacteria is disrupted. This is where probiotics are useful. Live yeasts and bacteria make up probiotics. To reestablish the relationship between beneficial and harmful bacteria, they increase the number of good bacteria in your stomach [3]. The gut microbiota located in the gastrointestinal tract is essential to both health and illness. A balanced microbiome is essential for homeostasis, while disturbances in microbial composition—referred to as *dysbiosis*—have been linked to various diseases including obesity, diabetes, and inflammatory bowel disease. The dominant bacterial phyla in a healthy human gut includes *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, and *Proteobacteria* [12].

A steady core of microorganisms and a high taxonomic diversity and microbial gene richness are characteristics of a healthy microbiota.

Together with the gut microbiota, gut immune, neuronal, and endocrine cells constitute a complex gut ecosystem that supports the host's homeostatic balance through host-microbe crossover. Additionally, the gut microbiota plays a crucial role in the fermentation of indigestible substrates such as dietary fiber. The development of specialized microbes that can produce short-chain fatty acids (SCFAs) is encouraged by this fermentation. Acetate, propionate, and butyrate are the three SCFAs that are most important for human health [1].

The human gut microbiota develops during infancy for a variety of reasons. One of the primary factors influencing gut microbiota throughout life is thought to be diet. In addition to defending against infections, intestinal bacteria are essential for preserving immunological and metabolic equilibrium. Numerous inflammatory illnesses and infections have been linked to dysbiosis or altered gut bacterial composition [2].

1.5 IBD disorders and colitis

Inflammatory bowel disease **(*IBD)*** is a set of related chronic disorders that mostly affect the gastrointestinal system, and it is a persistent, recurring condition that has been linked to an increased risk of colon cancer in later life [1]. It includes two main types of diseases; *Ulcerative Colitis (UC)* which affects only the colon and rectum, and *Crohn's Disease* that can affect any part of the GI tract (mouth to anus), often involving deeper layers of the bowel wall. ***Colitis*** is a symptom of IBD, particularly UC, but not all colitis is IBD (Fig.7). So, all ulcerative colitis is IBD, and it causes colitis, but not all colitis is IBD (Table 1).

Both diseases are associated with altered gut microbiota composition, often involving a decrease in beneficial bacteria such as *Firmicutes and Bacteroidota* and an overrepresentation of *Proteobacteria and Actinobacteria* [1]. These reductions are linked to lower SCFA levels in patients' stools. Prebiotics have been found in numerous studies to be beneficial for IBD. In animal models, when taken orally, inulin lessens the severity of dextran sodium sulfate (DSS)-induced colitis. Prebiotics may have had positive anti-inflammatory effects in studies comparing mice with and without DSS by modifying the gut microbiota, fortifying the intestinal barrier, and blocking IL-6/STAT3 signaling [1]. Dysbiosis may exacerbate the disease by increasing intestinal permeability and activating immune responses. Notably, studies have shown that dietary interventions targeting the microbiota—such as prebiotic supplementation—can modulate inflammation and support remission in IBD patients [1].

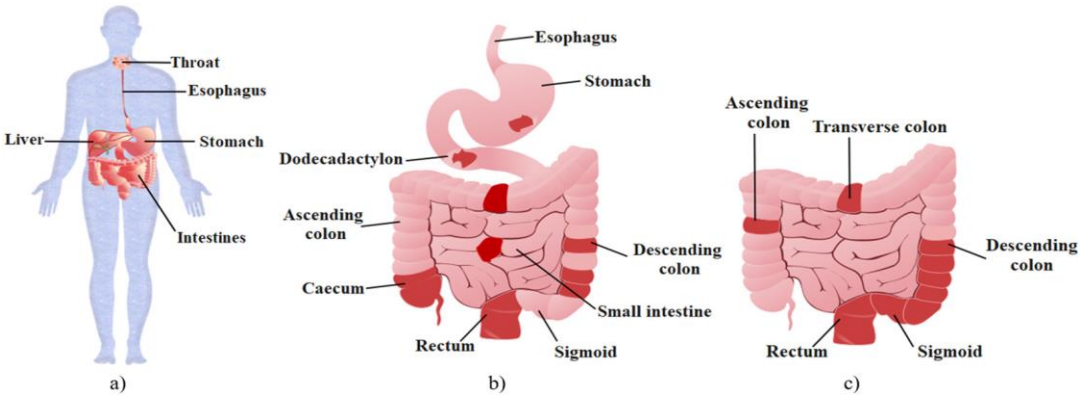


Figure 7 - *a) Human digestive system b) Whole digestive tract lesions in Crohn's disease c) Ulcerative colitis lesions are primarily found in the rectum and sigmoid, but they can also spread to the descending colon or even the entire organ [16].*

Despite the widely accepted theory that the illness results from intricate interplay between immunological, environmental, microbial, and genetic factors, little is known about the etiology and pathophysiology of IBD [11].

Furthermore, the term "colitis" describes swelling or inflammation of the colon (large intestine). It may appear as a chronic problem or as a transient, self-resolving injury. Colitis belongs to the larger group of illnesses of the digestive system [20].

Reducing the inflammation that causes symptoms is the aim of treatment for inflammatory bowel disease. In the best circumstances, this could result in a reduction in the risk of complications as well as long-term remission and symptom alleviation. Typically, IBD is treated with medication or surgery, so currently, there is no radical cure for IBD [21].

Term	***Scope***	***Chronic (Yes/No)***	***Cause***	***Example types***
Colitis	*Any colon inflammation*	*Not always*	*Infection, ischemia, etc.*	*Infectious, ischemic, etc.*
IBD	*Chronic GI inflammation*	*Yes*	*Autoimmune*	*Ulcerative **colitis**, Crohn's*

***Table 1** - Differences between colitis and IBD.*

1.6 Metagenomics

The study of the structure and function of whole nucleotide sequences, or genomes, straight from an environmental sample, such as stool, that have been separated and examined from every creature (usually bacteria) in a bulk sample is known as metagenomics. A particular community of microorganisms, found on human skin, in soil, or in a water sample, can be studied using metagenomics. It enables scientists to examine the makeup and roles of microbial communities in their native habitat [22].

This approach enables comprehensive characterization of the gut microbiome, including both abundant and rare microbial taxa. High-throughput sequencing technologies such as targeted amplicon (e.g., 16S- or 18S rRNA, internal transcribed spacer), shotgun metagenomic, metatranscriptomic, and genomic sequencing are widely applied in gut microbiome research and allow researchers to explore taxonomic composition, functional potential, and diversity within microbial communities [1].

1.6.1 Alpha Diversity

Ecologist R.H. Whittaker coined the term ***"alpha diversity" ("α-diversity")*** to describe the range of species present in a particular environment. Within a single microbial community, it quantifies species richness (diversity, number of species) and evenness (distribution across species). By looking at the number of species and their distribution, it sheds light on the general stability and complexity of microbial ecosystems. Alpha diversity enables us to comprehend the intricate relationships in ecosystems and the ways in which environmental factors impact them by examining the interactions and geographical distribution of organisms. A healthier gut microbiota is generally linked to a higher alpha diversity. Alpha diversity is frequently reduced in diseases like colitis and inflammatory bowel disease (IBD), which is indicative of a decline in microbial complexity and ecological resilience [10]. The Observed Species count, Simpson's index, and the Shannon diversity index are commonly used indexes to evaluate alpha diversity [8] [9].

In this study we are focusing solely on the **Shannon** and **Chao1** alpha diversity indices (Table2).

Metric	***Description***	***Strengths***
Chao1	*Counts the number of species present in a sample (species richness)*	*Simple and intuitive*
Shannon	*Considers abundance and evenness*	*Reflects overall community structure*

Table 2*- The roles of each alpha diversity index [23].*

1.6.1.1 Shannon Alpha Diversity Index (Shannon Entropy)

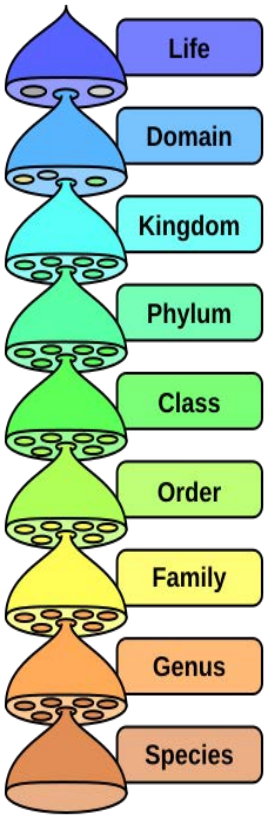
One mathematical metric for measuring species variety within a community is the Shannon index, sometimes referred to as the Shannon–Wiener or Shannon–Weaver index. It takes into account species evenness, which measures the consistency of individual distribution among those species, as well as species richness, which measures the number of unique species present. In ecology and microbiome research, this index is frequently used to assess within-sample biodiversity, also known as alpha diversity. A higher total diversity is indicated by a higher Shannon index, which rises with species richness and more uniform distribution. It is sensitive to rare species, although not as sensitive as other diversity indices [9][23].

1.6.1.2 Chao1 Alpha Diversity Index

Observed and unobserved (undetected) species in a community are both taken into consideration by Chao1, a non-parametric species richness estimator. It works especially effectively when a large number of species are uncommon, which makes it ideal for ecological surveys with insufficient sample and microbiome research. Chao1 only assesses richness; it ignores species evenness, in contrast to other diversity indices. Greater hidden biodiversity within the sampled environment is indicated by a higher Chao1 value. [9][23].

1.6.2 Relative Abundance

Relative abundance is the percentage of a particular microbe in the total microbial population. In other words, it illustrates the percentage of that microorganism in relation to the overall microbial count rather than the actual quantity of bacteria. Usually, the total of the relative abundances is 100% (or 1). It is used to compare the presence and dominance of different microbial groups across various samples or conditions [9].



Relative abundance refers to the proportion of individual taxa within a microbial community, typically reported at various taxonomic levels such as phylum, genus, or species (Fig.8). In gut microbiome studies, alterations in the relative abundance of key bacterial phyla—such as a reduction in *Firmicutes* and an increase in *Proteobacteria [11]* —are often indicative of dysbiosis. Such microbial imbalances can significantly influence host metabolism and immune responses and are frequently associated with inflammatory conditions, including colitis.

In this study, we focus on bacteria, specifically at the phylum level of taxonomic classification (Fig.9). In biological taxonomy, a phylum (plural: phyla) is a major rank below kingdom and above class. Within the domain of bacteria, *bacterial phyla* are vast evolutionary lineages that are mostly categorized by genetic links, especially 16S rRNA gene sequences, as well as physiological and ecological traits. This categorization offers a structure for arranging the enormous variety of bacteria into more advanced groups. Bacteroidota, Firmicutes, Proteobacteria, and Actinobacteria are notable examples of bacterial phyla.[17][18].

Figure 8 - *"The hierarchy of biological classification's eight major taxonomic ranks"* [17].

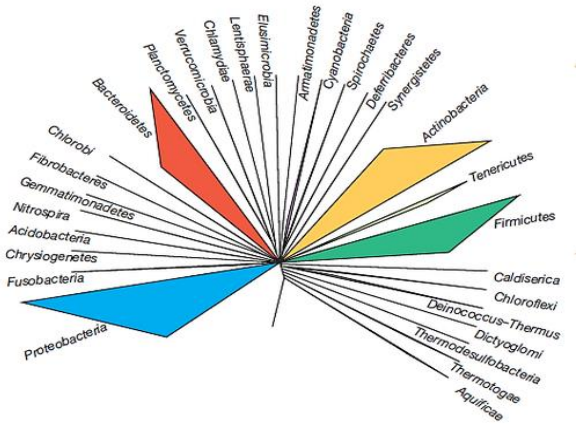


Figure 9 - *Bacteria Phyla [25*]

1.7 Meta-analysis

A meta-analysis refers to an “analysis of analyses” and the term “meta” is derived from Greek, meaning “after” or “beyond”. Although meta-analyses did not start to be frequently published in the medical literature until the late 1970s, they have since proliferated and are growing at an exponential rate[14].

A statistical technique called meta-analysis has grown in popularity as a means of resolving conflicts in clinical research by combining the results of several studies on a certain subject. Meta-analysis provides an objective and quantitative method of summarizing results by combining data from previous studies. This improves statistical power and precision and is especially helpful when individual studies have small sample sizes or little statistical significance. Beyond calculating total effects, meta-analysis can also look at study variability (heterogeneity) and find subgroups associated with particular parameters, frequently exposing trends that guide further research.

When carried out correctly, a meta-analysis is regarded as a high degree of proof in medical research and is frequently seen as definitive. It provides a more methodical and efficient way to synthesize study results than depending only on expert’s opinions. Meta-analysis is becoming a popular and growing research methodology as a result.[7].

A meta-analysis's findings could provide a more accurate assessment of the impact of a treatment or risk factor for illness, among other outcomes, than any particular research that contributed to the pooled analysis. Analyzing the variation or heterogeneity in research findings is another crucial consequence. One of the advantages of meta-analysis is that it provides a comprehensive, quantitative study of a sizable, frequently intricate, and perhaps seemingly contradictory body of literature [14].

You have to follow a methodical procedure when gathering data for a meta-analysis. Establish your research question first, then look for relevant research studies. Next, extract and arrange the data from these studies, making sure that you have a well-defined process for handling missing data and coding. Lastly, evaluate the data and present your conclusions, taking into account any potential biases and heterogeneity [15].

We conducted a meta-analysis using alpha diversity indices and relative abundance data, which are examined in detail in the following chapters. Specifically, we employed the Chao1 and Shannon diversity indices to assess alpha diversity and primarily focused on bacterial taxa at the phylum level.

1.8 Aim of the Study

The primary aim of this study is to investigate whether the dietary supplementation of the prebiotic inulin can modulate the gut microbiota of murine models suffering from colitis, promoting a microbial shift toward a eubiotic, or healthy-like, state.

We focus specifically on evaluating bacterial community composition at the phylum level, utilizing data extracted from multiple animal studies in which colitis was chemically induced (e.g., via DSS or TNBS) and inulin was administered as a dietary intervention.

Through meta-analysis of published metagenomic data, we seek to identify whether inulin supplementation consistently restores microbiota diversity and relative abundances—particularly of phyla such as *Firmicutes*, *Bacteroidetes*, and *Proteobacteria*—toward profiles observed in non-diseased, control animals.

This work also aims to assess the robustness and reproducibility of inulin's microbial effects across different studies and experimental designs.

Previous studies have shown that inulin supplementation can enhance short-chain fatty acid production, reduce gut permeability, and partially revert dysbiosis in colitic mice, suggesting a modulatory role on the gut ecosystem [5]. However, the consistency of these effects across experimental conditions remains unclear. To that end, here, we performed a meta-analysis to elucidate the role of inulin in re-establishing gut microbial balance under inflammatory conditions. Our goal is to investigate the role of inulin as a potential diet-therapeutic strategy for microbiome-targeted interventions in IBD.

Ultimately, our goal is to elucidate whether inulin holds therapeutic potential in re-establishing gut microbial balance under inflammatory conditions, thus supporting its role as a dietary strategy for microbiome-targeted interventions in IBD.

2. MATERIALS AND METHODS

2.1 Search strategy

Our research followed the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* guidelines, also known as PRISMA, which outlines essential reporting items to promote transparency and completeness in systematic reviews and meta-analyses [26]. To achieve the most relevant results when searching published studies, we crafted a clear, well-structured research question with appropriate content. In the following sections, our strategy is presented in detail.

2.1.1 Creation of the question

Our main interest was to investigate the relationship between gut microbiome and inulin under pathological conditions like colitis and IBD. To investigate the effect of the systematic administration of inulin, a known prebiotic, in gut microbiome under pathological conditions, we formed a clear question which included the microbiome of interest, the pathological phenotype, the defined intervention as well as the specific model organism for which we collected data.

2.1.2 Selection and Inclusion Criteria of reports

To ensure the appropriate selection of studies for this research, clearly defined inclusion criteria were established, regarding the methodology of the study and the quality of the presented data. This research aims to evaluate the effects of the prebiotic inulin on alpha diversity and relative abundance indices in animal models before and after the induction of colitis. Therefore, eligible studies were required to include data on both alpha diversity and relative abundance to ensure compatibility with the scope and objectives of the analysis. Only studies published in English were considered, in order to minimize the risk of information loss during translation, which could impact data extraction and the interpretation of results.

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines were applied for detailed documentation of the inclusion criteria (Figure 10).

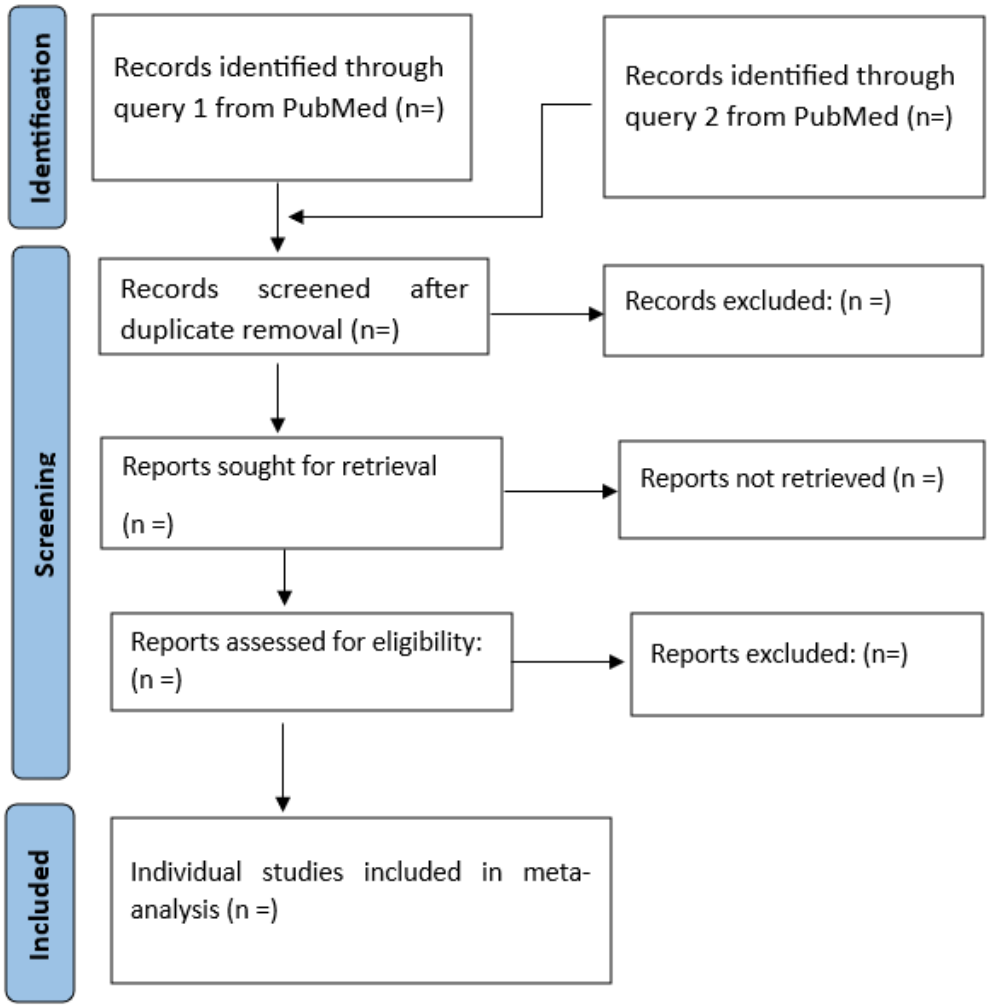


***Figure 10**- PRISMA-2020 flow diagram showing the pipeline for systematic meta-analysis.*

2.1.3 Literature Search Strategy

A structured and comprehensive search strategy was implemented to identify relevant literature. A well-established large scientific database called **PubMed** was utilized for this purpose. This database hosts a vast range of peer-reviewed articles across various scientific disciplines.

Search queries were constructed using combinations of relevant keywords and Boolean operators (AND, OR, NOT) to filter results effectively. The selection of databases and search terms was tailored to the biomedical field, specifically focusing on microbiome-related studies involving inulin and colitis in animal models.

2.1.4 Data Extraction and Management

Eligible studies were systematically recorded in a Microsoft Excel spreadsheet, where each row represented a unique study and each column captured a distinct variable. Key information included the organism used, next-generation sequencing (NGS) technique, PubMed ID (PMID), and additional metadata relevant to study design.

After refining the focus of the research exclusively to colitis, only those studies that provided both alpha diversity indices and relative abundance values were retained for analysis. Non-eligible articles were excluded at this stage. The final dataset included studies with detailed methodological descriptions, such as inulin dosage, mouse model type, colitis induction method, and even the country of origin of each study. These variables are critical for assessing the consistency and comparability of the experimental conditions across studies.

2.2 Statistical Analysis

Statistical analysis was performed on data presented in a predefined Microsoft Excel spreadsheet by applying the STATA software

2.2.1 STATA package

The statistical analysis for this meta-analysis was conducted using the STATA software package. We specifically used the version STATA SE 13.1.

Stata was originally created by the Computing Resource Center in California, with its first release in 1985. Since then, a new version has been launched every two years [28]. Stata is a comprehensive statistical software package that supports a wide range of data science tasks. It offers tools for data management, graphical representation, statistical modeling, and automated reporting, making it suitable for conducting complex analyses such as meta-analyses [29].

2.2.2 Main STATA Commands used in our analysis

In this section, the commands and the special characters applied in STATA during our analysis will be presented.

Commands	Command explanation
generate (gen)	Create or change contents of variable [36]
display (di)	Display strings and values of scalar expressions [33]
lincom	Computes point estimates, standard errors, test statistics (t or z statistics), significance levels (p-value), and confidence intervals for linear combinations of coefficients after any Stata estimation command, including survey estimation. Results can optionally be displayed as odds ratios, hazard ratios, incidence-rate ratios, or relative-risk ratios and are also displayed in the usual table format used for displaying estimation results. Standard error calculations are based on the "delta method", an approximation appropriate in large samples.[30][34]
encode	Encode string into numeric and vice versa. encode creates a new variable named newvar based on the string variable varname (condition_num), creating, adding to, or just using (as necessary) the value label newvar or, if specified, name [35].
mvmeta	Multivariate random-effects meta-analysis and meta-regression [31]. Shown below in detail.
metan	Performs meta-analysis of aggregate (summary) data[32]. Shown below in detail.

Table 3 *- Table showing the STATA commands used throughout the meta-analysis. [source: STATA help section]*

Special character	Special character explanation
= =	Equal to
=	Equal sign for assignment
+	(Plus Sign): Used for addition
-	(Minus Sign): Used for subtraction
*	(Asterisk): Used for multiplication
/	(Slash): Used for division
^	(Caret): Often used for exponentiation

Table 4 *- Table showing the special characters used in the STATA commands.*

2.2.3 STATA commands for meta-analysis

The main STATA commands used in our meta-analysis are the following two:

(I) mvmeta -- *Multivariate random-effects meta-analysis and meta-regression [31]*

Syntax: **mvmeta b V [xvars] [if] [in]**

"The **mvmeta command** conducts multivariate random-effects meta-analysis (White, 2009) and multivariate random-effects meta-regression (White, 2011) on a dataset including point estimates, variances, and (optionally) covariances" [31].

As mentioned earlier, the standard meta-analysis combines estimates of one parameter over several studies. Multivariate meta-analysis is an extension that can combine estimates of several related parameters [31].

In our analysis, we calculated two SMDs per study: one comparing healthy controls with diseased mice (b3 = x0 – x1 / sp), and another comparing inulin-treated mice with diseased mice (b4 = x2 – x1 / sp), where sp is the pooled standard deviation across the three groups. The difference between these two SMDs (d = b4 – b3) represents the extent to which inulin reverses the microbial changes induced by colitis towards the healthy control condition

To perform this analysis, we first computed the pooled standard deviation (sp) and then calculated the effect sizes b3 and b4 for each study. We also derived the associated variances and covariance of these estimates: V33 (variance of b3), V44 (variance of b4), and V34 (covariance between b3 and b4). These components form the variance–covariance matrix required for the mvmeta command. The covariance term V34 is particularly important, as both b3 and b4 share the diseased group as a reference, introducing correlation between these comparisons. Including this term allows the multivariate model to account for the non-independence of the estimates within each study [31].

The multivariate meta-analysis was conducted using the following syntax in Stata SE 13.1:

mvmeta b V, vars(b3 b4)

Here, b is the vector containing the effect sizes (b3 and b4), and V is the corresponding variance–covariance matrix consisting of V33, V44, and V34. This model estimates the overall effect sizes across studies and accounts for within- and between-study variance [31]. To evaluate whether inulin reversed the microbial changes caused by DSS-induced colitis, we computed the difference d = b4 – b3 and its standard error using:

di_b[b4] -_b[b3]
di sqrt (V33 + V44 - 2*V34)

To test for statistical significance and obtain confidence intervals, we used the lincom command:

lincom _b[b4] - _b[b3]

This approach provides an estimate, standard error, p-value, and confidence interval for the difference in effect sizes, allowing us to interpret whether inulin brings the microbiome closer to the healthy condition [34]. A statistically non-significant or small value of d would suggest that the inulin-treated group is similar to the healthy control, indicating a potential restorative effect of inulin.

(II) metan -- *Perform meta-analysis of aggregate (summary) data [32]*

Syntax: **metan** varlist [if] [in] [, precalculated effect estimates options]

Depending on the data's structure, the varlist may have two, three, four, or six numerical variables.
More precisely, any of the following syntaxes can be used with metan command: Generic (pre-calculated) effect sizes and associated standard errors are meta-analyzed, and participant numbers can be provided using npts(varname). Effect sizes, such as log odds-ratios instead of odds ratios, must be based on a normal distribution [32].

Syntax: ***metan ES seES*** *[if] [in] [, npts(varname) model_spec options_main]*

Generic (pre-calculated) effect sizes and their 95% CIs are meta-analyzed, and participant numbers can be provided using npts (varname). Assuming symmetrical confidence intervals, the standard error is calculated as (UCI - LCI) / 2*1.96. Therefore, the pooled result won't be accurate unless the confidence limits are based on a normal distribution and have a 95% level [32].

In order to conduct our meta-analysis, we have to use the metan STATA command. By using the prior command to conduct the research, we were able to combine and analyze our data, illustrate our conclusions using forest plots, and improve our precision. We use the random, by(), and label() commands to access the precalculated effect estimates choices. Heterogeneity is drawn from the Cohran model, while “random” defines a random effects model using the DerSimonian & Laird approach. We can better analyze the treatment effect between subgroups by separating them into subgroups with the use of by(). We identified the data by the name of the author and the published year by using the label() command, and more specifically, label([namevar=namevar], [yearvar=yearvar]).

metan data stde if index=="Shannon ", by(condition_num) random label(namevar=author, yearvar=year

metan data stde if index=="Chao ", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Firmicutes", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Bacteroidota", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Proteobacteria", by(condition_num) random label(namevar=author, yearvar=year)

In this command:

data refers to the effect size variable (e.g., Shannon index values).
stde represents the corresponding standard error for each effect size.
index == "Shannon" restricts the analysis to a specific diversity measure, in this example in

the Shannon diversity index.
by(condition_num) specifies that the meta-analysis should be conducted separately for each condition (e.g., control, disease, disease + inulin), allowing for subgroup-specific summary estimates.
random invokes a random-effects model, which accounts for between-study heterogeneity and provides a more conservative estimate compared to fixed-effect models.
label(namevar=author, yearvar=year) ensures that each study is labeled in the forest plot with its corresponding author and publication year, aiding interpretation.

Syntax: ***metan x0 n0 v0 x1 n1 randomi nostandard label(namevar=author,yearvar=year)***

In this command:

x0, x1: Indicate the mean values for the two groups being compared in each study (e.g., control and treatment groups).

n0, n1: Represent the sample sizes corresponding to each group.

randomi: Specifies the use of a random-effects model, which accounts for between-study heterogeneity and provides a more conservative and generalizable estimate compared to fixed-effects models.

nostandard: Prevents the command from standardizing the effect sizes, which is appropriate when using raw mean differences rather than standardized mean differences (e.g., when all outcomes are measured on the same scale).

label(namevar=author, yearvar=year): Ensures that each study is labeled in the forest plot by the corresponding author's name and year of publication, facilitating easier interpretation and traceability of the results.

2.2.4 Effect size

Effect size is a quantitative measure that expresses the magnitude of the difference or relationship between two groups or conditions. In meta-analysis, it enables the comparison and synthesis of results across studies that may use different scales or metrics [37].

Required variables: Means (or proportions) for each group, Standard deviations or standard errors, sample sizes per group

Type of data: Continuous data (e.g., alpha diversity indices), Binary data (e.g., presence/absence or relative abundance)

2.2.4.1 Fixed effect model

As the name suggests, the fixed-effect model implies that all of the studies in the meta-analysis have a single real effect size. Since there is only one actual effect, the singular (effect) is utilized [20].

In the *fixed-effect model,* all observed differences reflect sampling error. Therefore, any variance would come from within each study. The pooled estimate is calculated as a weighted average, where the weight assigned to each study is the inverse of that study's variance. Larger studies have much more weight than smaller studies in the fixed-effect models [37].

2.2.4.2 Random effects model

The random-effects model assumes that because of the heterogeneity (differences) between studies, the actual effect may differ [45].

The effect size may vary from study to study due to variations in participant mixtures. The effect estimates of every study would have a normal distribution if an infinite amount of research could be conducted. The mean or average effect would be the pooled estimate. The phrase "random effects" comes from the assumption that the effect sizes in the conducted research are a random sample of all potential impact sizes. Because there are different types of actual impacts, researchers use the plural (effects) [37].

In the *random-effects model,* 2 sources contribute to the variance, one arising from within each study and one that occurs between studies. Both need to be taken into account. While larger studies still have larger weights, in the random effects models, smaller studies have relatively greater weight than in fixed-effect models [37][44].

2.2.5 Confidence Interval (CI)

A confidence interval provides a range of values that likely contain the true effect size with a certain level of confidence, commonly 95%.

Required variables: Effect size, Standard error (SE) of the effect size.

Confidence intervals help interpret the precision and uncertainty of the estimated effect [46].

$$CI = ES \pm Z \cdot SE$$

where:

ES = estimated effect size

SE = standard error

Z = Z-score corresponding to the confidence level (e.g., 1.96 for 95%)

2.2.6 P-value

The probability of obtaining results at least as extreme as those observed, assuming the null hypothesis is true. It is used to test statistical significance.

Required variables: Effect size, Standard error or variance, sample size

$$Z = \frac{ES}{SE}$$

Then, the Z-score is compared to the standard normal distribution to compute the corresponding p-value.

The purpose of the p-value statistic is to assess whether the observed effect is statistically significant and unlikely to have occurred by random chance [47].

2.2.7 Mean difference meta-analysis (Random Effect non standardized)

Mean difference (MD) meta-analysis is a commonly used approach when combining continuous outcomes measured on the **same scale across studies**. It estimates the average absolute difference in means between two groups (e.g., treatment vs. control). Unlike standardized mean difference (SMD), which adjusts for differences in scale or measurement units and will be discussed later, the mean difference retains the original units of measurement, making the results more interpretable when scales are consistent.
This analysis was performed under a **random-effects model**, which assumes that the true effect size may vary between studies due to clinical or methodological differences. The random-effects model incorporates between-study heterogeneity and provides more conservative confidence intervals compared to fixed-effect models [53].

Required variables:

- **x**: Mean of each group (e.g., treatment and control)
- **sd**: Standard deviation of each group
- **n**: Sample size of each group

The analysis does **not standardize** the effect sizes (nostandard option in Stata), as all included studies used the same outcome metric (e.g., alpha diversity index measured uniformly across studies). Therefore, raw mean differences were appropriate.

$$Mean\ Difference = \overline{X1} - \overline{X2}$$

where:

x1: mean of the 1st group/condition

x2: mean of the 2nd group/condition

2.2.8 Multivariate meta-analysis (Random Effect)

Variables required for this analysis: sd,n,x,b3,b4,d,V33,V44 and V34.

Multivariate meta-analysis extends the traditional approach by modeling multiple correlated outcomes simultaneously. It incorporates both within- and between-study variation (heterogeneity), allowing for more flexible and realistic modeling.

Required variables: Multiple effect sizes per study, Variance (or standard errors) for each effect.

Alpha diversity data: Used SMD due to differences in scale and lack of normalization. We performed multivariate meta-analysis using SMD on the alpha diversity data because we wanted to see the effect of our 3 conditions (healthy, case, case+inulin) combined.

Relative abundance data: Used mean difference because the data were already normalized (same unit because they were percentages). We also used the **Bernoulli formula** to estimate SE from relative abundance proportions considering the prevalence as a Bernoulli variable.

2.2.8.1 Standardized Mean Difference (SMD)

We performed multivariate meta-analysis using SMD (Standardized Mean Difference) on the alpha diversity data. We used SMD in meta-analysis to compare effects from studies that measure the same outcome but use different scales or units. By standardizing the mean difference, we can express the effect size in units of standard deviation, making it easier to compare across studies. We also used the mean difference and the Bernoulli formula in order to calculate the SE of the relative abundance values of the different bacterial phyla.

This approach is especially useful in our analysis, as the included studies involved murine models of DSS-induced colitis with varying inulin dosages and presentation formats. Without standardization, raw differences in bacterial diversity would not be directly comparable due to differences in data scaling, experimental design, and reporting methodology.

By applying SMD, we were able to mitigate the bias introduced by these differences and derive a more consistent and interpretable estimate of effect size across the studies. This method also aligns with recommendations from established meta-analytic guidelines, such as those outlined in the Cochrane Handbook, which advocates for the use of SMD when outcomes are conceptually similar but measured using diverse methods. Ultimately, the use of SMD enabled us to pool the data more effectively and draw more reliable conclusions about the impact of inulin on gut microbiota composition in the context of colitis-induced dysbiosis [38][39][43].

Mathematical formula (Cohen's d):

$$SMD = \frac{\overline{X1} - \overline{X2}}{SDpooled}$$

$$SDpooled = \sqrt{\frac{[(n0-1)\cdot SD0^2]+[(n1-1)\cdot SD1^2]+[(n2-1)\cdot SD2^2]}{(n0+n1+n2+n3-3)}}$$

where:

x1_x2_= means of the two groups.

2.2.8.2 Standard Deviation (SD)

Measures the dispersion or spread of individual values around the mean [41].

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(xi-\overline{x})^2}{n-1}}$$

2.2.8.3 BERNOULLI Formula

The Bernoulli distribution is a type of discrete probability distribution that describes a random variable which can only take two values: 1, occurring with probability **p**, and 0, occurring with probability q = 1 – p. In simpler terms, it's used to represent the outcomes of an experiment with a binary (yes-or-no) result. These outcomes can be seen as Boolean values—such as success/failure, true/false, or on/off—where the probability of success (1) is **p**, and the probability of failure (0) is **q** [42].

Bernoulli-Based SE for Proportion Data

For relative abundance data (expressed as proportions), we assumed a Bernoulli distribution, where outcomes are binary (presence/absence) would help us interpret our data in a better way.

We need variables like p for the percentage, 1-p=q, n (sample size).

We used the p variable to store the relative abundance percentages after we divided them by 100.

Standard Error for a Proportion (Bernoulli SE):

$$SE = \sqrt{\frac{p \cdot q}{n}} = \sqrt{\frac{p \cdot (1-p)}{n}} = \frac{SD}{\sqrt{n}}$$

where:

- p= proportion (relative abundance value)
- n = sample size
- q = 1-p

The purpose is to estimate the uncertainty associated with a proportion, which is then used in meta-analytic models [48].

2.2.8.4 Standard error (SE)

A metric that shows how accurate sample data is in the population. The standard deviation is used to determine accuracy. SE is equal to SD divided by the sample's square root, according to the relationship between the two.

The more representative the sample, the lower the SE [40]. The following formula is applied where:

σ is the population standard deviation, while n is the sample size.

$$SE = \frac{\sigma}{\sqrt{n}}$$

2.2.9 Weighted Mean

The weighted mean aggregates effect sizes across studies while accounting for the precision of each estimate. Studies with smaller variance contribute more heavily.

In a meta-analysis, a weighted average of the effect estimates from each of the individual studies is used to compute a pooled estimate. Each study is given a weight according to the inverse of the total error variance, or 1/variance. Studies with bigger sample sizes are typically accorded more weight [37].

$$\overline{X}weighted = \frac{\sum wi \cdot Xi}{\sum wi}$$

where:

Xi = effect size of study i

Wi=1/Vari = inverse variance weight

The main purpose is to produce a pooled estimate that reflects both effect magnitude and study quality (through variance).

The weighted nature of the subgroup meta-analysis ensured that studies with larger sample sizes and more precise estimates contributed more to the overall summary statistics, enhancing the reliability of the findings. This methodological strength adds robustness to the observed trends and supports the conclusion that inulin has a selective but measurable effect on the recovery of gut microbial diversity and composition following colitis-induced dysbiosis [37].

2.2.9.1 Subgroup Meta-analysis

Variables required for this analysis: data=x=mean on alpha diversity data or percentage of relative abundance, stde=SE

Subgroup meta-analysis involves calculating weighted means separately for predefined subgroups.

The main purpose is to explore heterogeneity and to identify whether effect sizes differ significantly across subgroups.

This approach allows us to analyze effect sizes across different experimental conditions separately—namely, healthy controls, DSS-induced colitis (diseased), and colitis with inulin supplementation. By grouping the analysis by condition, we were able to explore how alpha diversity indices and relative abundance of bacterial taxa varied across each state of microbiome health.

The subgroup meta-analysis enabled us to compare the weighted mean effect sizes of microbial diversity and abundance under different experimental treatments. This approach is particularly useful when heterogeneity exists between subgroups—such as variation in inulin dosage, disease severity, or microbial response—allowing for the detection of patterns that might be obscured in an overall pooled estimate.

This approach provides greater insight into how the gut microbiome shifts in response to disease and dietary intervention, and how closely the intervention group approximates the healthy state [49].

Required variables: Effect size, Variance, grouping variable (condition)

Pipeline: i) Partition data by subgroup ii) Calculate subgroup-specific weighted means iii) Compare subgroup estimates using Q-tests or interaction models.

2.3 STATA Commands used for Alpha diversity data

2.3.1 Multivariate standardized mean difference meta-analysis (SMD) Variance Covariance - SHANNON

```
gen sp=sqrt((((n0-1)*(sd0^2))+ ((n1-1)*(sd1^2))+ ((n2-1)*(sd2^2)))/(n0+n1+n2-3)))

gen b3=(x0-x1)/sp

gen b4=(x2-x1)/sp

gen V33= (1/n0) + (1/n1) + ((b3^2)/(2*(n1+n0+n2)))

gen V44= (1/n2) + (1/n1) + ((b4^2)/(2*(n1+n0+n2)))

 gen V34= (1/n1) +((b3*b4)/(2*(n1+n0+n2)))

mvmeta b V, vars(b3 b4)

**calculation of d

di _b[b4] - _b[b3]

**calculation of standard error of d - Equation 5.2

di sqrt( V33+ V44- 2* V34)

**automatic calculation of confidence intervals for d

lincom _b[b4] - _b[b3]

 ( 1) - [Overall_mean]b3 + [Overall_mean]b4 = 0
```

2.3.2 Multivariate standardized mean difference meta-analysis (SMD) Variance Covariance - CHAO

```
gen sp=sqrt((((n0-1)*(sd0^2))+ ((n1-1)*(sd1^2))+ ((n2-1)*(sd2^2)))/(n0+n1+n2-3))

gen b3=(x0-x1)/sp

gen b4=(x2-x1)/sp

gen V33= (1/n0) + (1/n1) + ((b3^2)/(2*(n1+n0+n2)))

gen V44= (1/n2) + (1/n1) + ((b4^2)/(2*(n1+n0+n2)))

gen V34= (1/n1) +((b3*b4)/(2*(n1+n0+n2)))

mvmeta b V, vars(b3 b4)

**calculation of standard error of d - Equation 5.2

di sqrt( V33+ V44- 2* V34)

**automatic calculation of confidence intervals for d

lincom _b[b4] - _b[b3]

( 1) - [Overall_mean]b3 + [Overall_mean]b4 = 0
```

2.3.3 Weighted mean subgroup meta-analysis command for alpha diversity indices

encode condition, gen(condition_num)

metan data stde if index=="Shannon ", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if index=="Chao ", by(condition_num) random label(namevar=author, yearvar=year)

2.3.4 Mean difference meta-analysis command for alpha diversity indices

metan n0 x0 sd0 n1 x1 sd1, randomi nostandard label(namevar=author,yearvar=year)

metan n0 x0 sd0 n2 x2 sd2, randomi nostandard label(namevar=author,yearvar=year)

metan n1 x1 sd1 n2 x2 sd2, randomi nostandard label(namevar=author,yearvar=year)

2.4 STATA Commands used for Relative abundance data

2.4.1 Multivariate mean difference meta-analysis Variance Covariance (using the Bernoulli formula)

gen sp=sqrt((((n0-1)*(sd0^2))+ ((n1-1)*(sd1^2))+ ((n2-1)*(sd2^2)))/(n0+n1+n2-3))

gen b3=(x0-**x1**)/sp

gen b4=(x2-**x1**)/sp

gen V33= (1/n0) + (1/n1) + ((b3^2)/(2*(n1+n0+n2)))

gen V44= (1/n2) + (1/n1) + ((b4^2)/(2*(n1+n0+n2)))

gen V34= (1/n1) +((b3*b4)/(2*(n1+n0+n2)))

mvmeta b V, vars(b3 b4)

**calculation of d

di _b[b4] - _b[b3]

**calculation of standard error of d - Equation 5.2

di sqrt(V33+ V44- 2* V34)

**automatic calculation of confidence intervals for d

lincom _b[b4] - _b[b3]

2.4.2 Weighted mean subgroup meta-analysis command for relative abundance

encode condition, gen(condition_num)

metan data stde if taxa=="Firmicutes", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Bacteroidota", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Proteobacteria", by(condition_num) random label(namevar=author, yearvar=year)

3. RESULTS

3.1 Search strategy and selection criteria

A comprehensive literature search was performed in PubMed database with the following search terms: "inulin microbiome" and "INULIN AND (GUT MICROBIOTA' OR 'GUT MICROBIOME')". We focused on studies published from 2018 onwards. Search with the first set of search terms was performed on January 2024 and returned 935 articles, whereas search with the second set of terms, performed during the same period, resulted in 755 articles. Duplicate articles were removed before the screening process. A 3rd search was conducted 18th of March 2025 with the first search term, which retrieved 209 articles.

Eventually, all 1144 articles with their PubMed IDs were collected in an Excel spreadsheet. Initially, we excluded 104 reports identified as bibliographic reviews, systematic reviews, meta-analyses and commentaries. We then screened 1040 articles to identify the disease model and the respective model organism as well as the application of NGS in metagenomics analysis. The initial inclusion criteria were (1) the disease model of colitis/IBD; (2) the use of the mouse model; (3) the metagenomics analysis based on high-throughput sequencing of the 16s genetic locus; (4) the articles had to be published from 2018 onwards; (5) only inulin not probiotics or a prebiotic-probiotic mix and (6) no drugs prior or after the induction of the disease; We used the YES/NO column to explain if the articles were eligible or not while keeping in mind the inclusion criteria.

The screening revealed 18 articles focused on colitis, NGS, the mouse model and had no prior intervention of drugs or probiotics. From the 18 articles that were assessed for eligibility we excluded 11 articles for reasons such as (7) studies did not include both of the following microbial community metrics: alpha diversity (Chao or Shannon index) and relative abundance.

At last, we ended up with 7 articles in total that were eligible for the meta-analysis and more specifically we performed our meta-analysis on alpha diversity and relative abundance data separately from which we used 6 out of 7 studies for the alpha diversity and 5 out of 7 for the relative abundance.

Figure 11 shows the PRISMA guidelines that we followed in order to select the most appropriate articles for our meta-analysis and the number of reports included in each step of the process.

For the Alpha diversity analysis, we categorized the data by the two most common alpha diversity indices we gather from all our data, named Chao1 and Shannon.

For the Relative abundance analysis, we focused on the Bacterial phyla (Phylum), because we found the most data unlike other bacterial classifications (Class, Order, Genus, etc.), and categorized the data by the 3 most abundant Bacterial Phylum called Firmicutes, Bacteroidota and Proteobacteria.

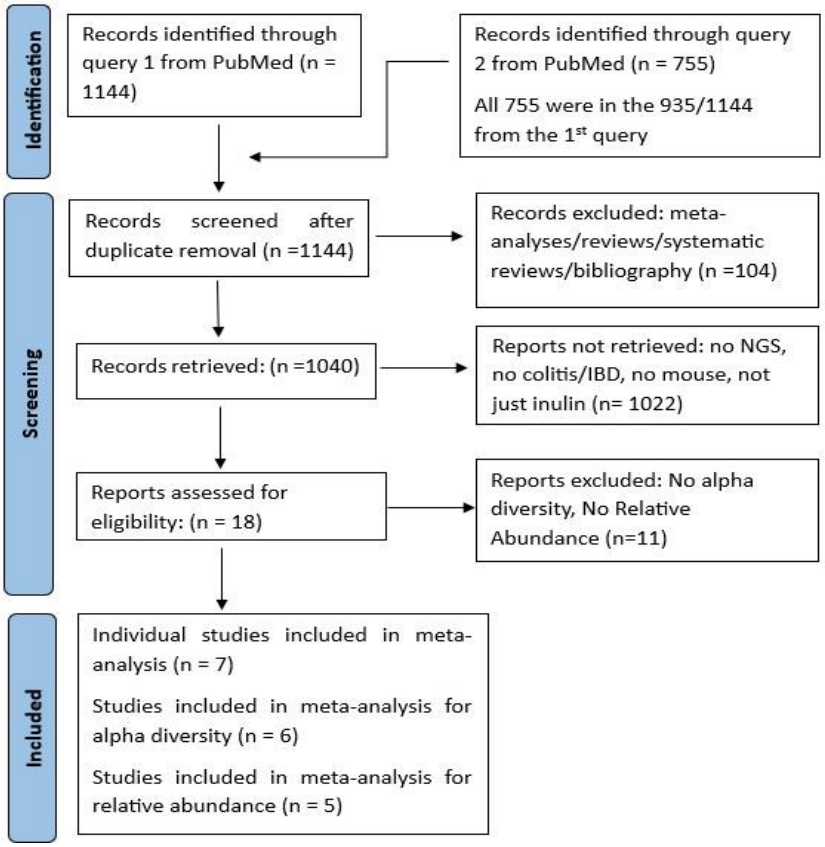


Figure 11- *PRISMA-2020 Flow diagram depicting the number of studies in all stages of the PRISMA pipeline for alpha diversity and relative abundance data.*

3.2 Meta-analysis of alpha diversity

Alpha diversity refers to the variety of species within a specific ecological unit or single sample, measuring both species' richness and evenness. It provides insight into the structural complexity and dynamics of a biological community by considering species count, abundance, and distribution within a localized habitat. Alpha diversity offers valuable information about the health of an ecosystem; a high-level diversity suggests a resilient system with strong adaptability to environmental shifts, whereas low diversity can indicate ecological fragility. In this study we are focusing solely on the **Shannon** and **Chao1** alpha diversity indices, which describe species evenness and richness respectively [8][9]

3.2.1 Subgroup meta-analysis of alpha diversity indices

We conducted a weighted mean subgroup meta-analysis, with the help of STATA SE 13.1 package, on the alpha diversity data (see Table 4) while using the following commands:

metan data stde if index=="Shannon ", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if index=="Chao ", by(condition_num) random label(namevar=author, yearvar=year)

Table 4 categorizes the alpha diversity data by index and condition (healthy, IBD, IBD + inulin). For each of the two indices, we included variables such as sample size (number of animals per group), data (mean), sd (standard deviation), and se (standard error) to provide a comprehensive summary of the data before we performed the analysis.

PMID	author	year	index	condition	data	stde	sd	sample size
38024237	Zhang Z	2023	Shannon	healthy	3.56	0.10	0.2	4
37659809	Li Z	2023	Shannon	healthy	4.13	0.14	0.35	6
37659809	Li Z	2023	Shannon	healthy	4.13	0.14	0.35	6
35415348	Qiao H	2022	Shannon	healthy	3.99	0.24	0.68	8
30589960	Cai Y	2019	Shannon	healthy	6.99	0.07	0.18	6
29232851	Chen Q.	2017	Shannon	IBD	5.85	0.10	0.35	13
38024237	Zhang Z	2023	Shannon	IBD	2.56	0.14	0.28	4
37659809	Li Z	2023	Shannon	IBD	3.71	0.21	0.50	6
37659809	Li Z	2023	Shannon	IBD	3.71	0.21	0.50	6
35415348	Qiao H	2022	Shannon	IBD	2.85	0.16	0.44	8
30589960	Cai Y	2019	Shannon	IBD and inulin	6.76	0.05	0.13	6
29232851	Chen Q.	2017	Shannon	IBD and inulin	6	0.19	0.6	10
38024237	Zhang Z	2023	Shannon	IBD and inulin	2.77	0.13	0.28	5
37659809	Li Z	2023	Shannon	IBD and inulin	3.81	0.29	0.64	5
37659809	Li Z	2023	Shannon	IBD and inulin	3.63	0.30	0.72	6
35415348	Qiao H	2022	Shannon	healthy	3.59	0.16	0.46	8
30589960	Cai Y	2019	Shannon	healthy	7.15	0.25	0.62	6
29232851	Chen Q.	2017	Shannon	healthy	5.97	0.16	0.54	12
38024237	Zhang Z	2023	Chao	healthy	305	16.42	36.71	5
37659809	Li Z	2023	Chao	healthy	5014.39	251.80	616.77	6
37659809	Li Z	2023	Chao	healthy	5014.39	251.80	616.77	6
32813896	Tao C.	2021	Chao	healthy	414.93	4.48	7.75	3
30589960	Cai Y	2019	Chao	healthy	900.61	13.80	33.81	6
38024237	Zhang Z	2023	Chao	IBD	214.03	13.02	29.11	5
37659809	Li Z	2023	Chao	IBD	4697.84	712.23	1744.6	6
37659809	Li Z	2023	Chao	IBD	4697.84	712.23	1744.6	6
32813896	Tao C.	2021	Chao	IBD	58.20	1.492	2.58	3
30589960	Cai Y	2019	Chao	IBD	815.76	46.71	114.42	6
38024237	Zhang Z	2023	Chao	IBD and inulin	251.32	11.20	22.39	4

37659809	Li Z	2023	Chao	IBD and inulin	4978.41	496.40	1109.99	5
37659809	Li Z	2023	Chao	IBD and inulin	4769.78	280.58	687.26	6
32813896	Tao C.	2021	Chao	IBD and inulin	47.76	5.97	10.34	3
30589960	Cai Y	2019	Chao	IBD and inulin	880.3	15.57	38.14	6

Table 4 - *The Alpha diversity data categorized by index and by condition for their subgroup analysis.*

As we can see in the forest plots, the overall result of each condition's subgroup analysis was statistically significant and favored the condition on the right (Fig. 12 and Fig13).

The forest plots display clear stratification across conditions for each alpha diversity index. For both indices, inulin supplementation in colitis-induced mice moves the value of alpha diversity closer to that of the Healthy group, suggesting a potential restorative effect. All results across conditions were statistically significant ($p < 0.05$), supporting the validity of the observed microbial shifts.

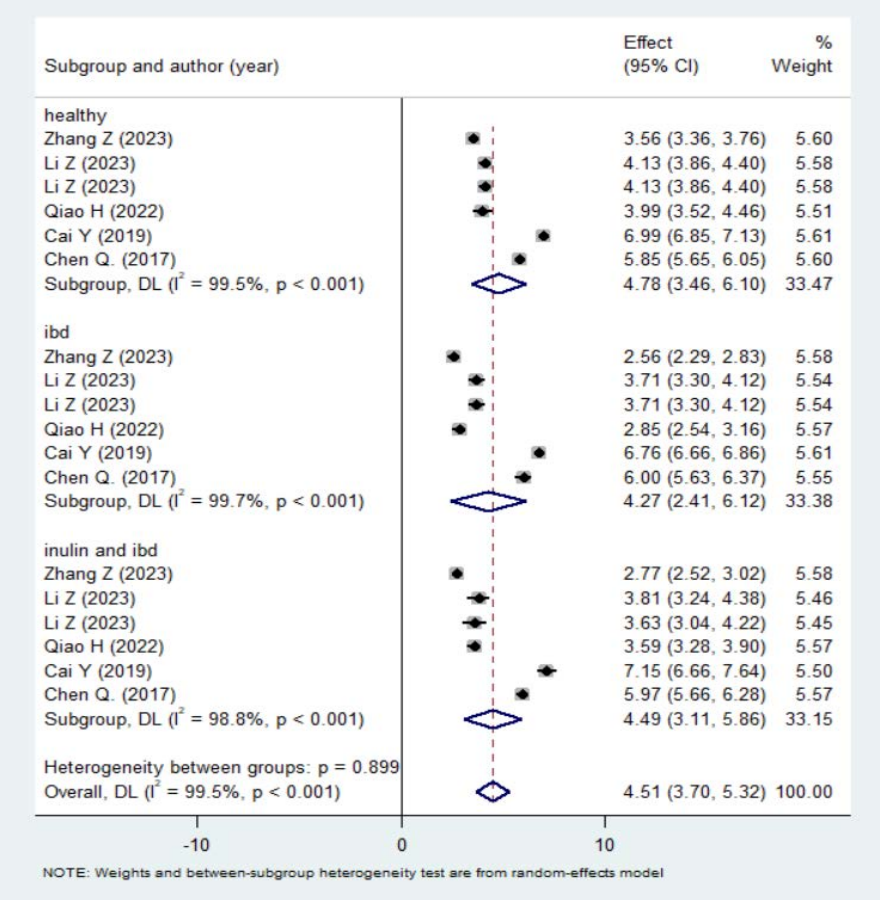


Figure 12 – *Shannon index weighted mean stratification forest plot. IBD or colitis refers to the second condition (case).*

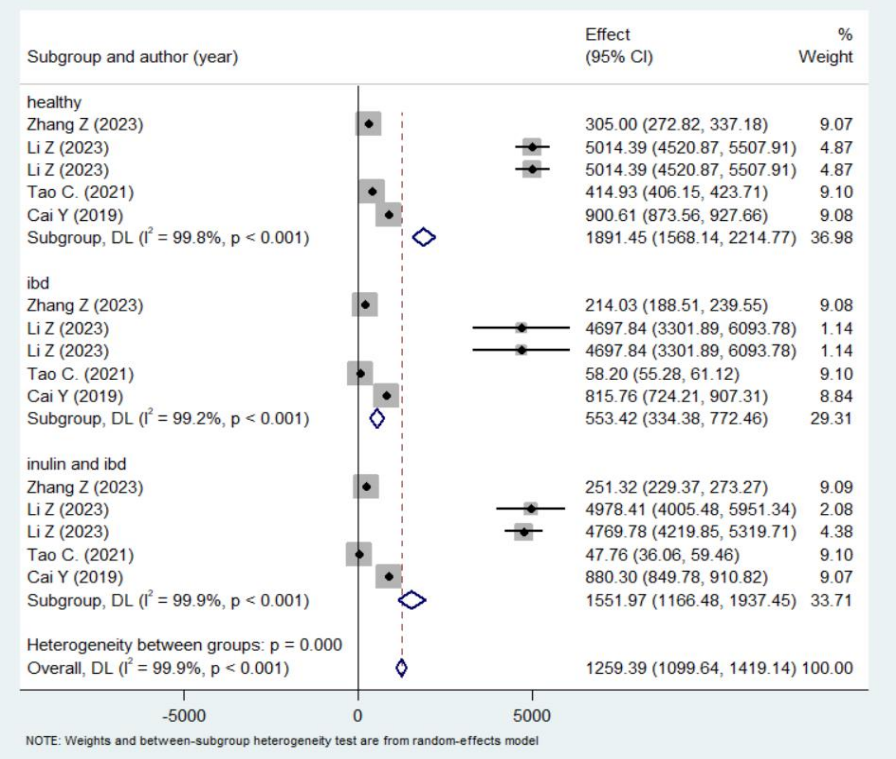


Figure 13 – *Chao index weighted mean stratification forest plot.*

The mean weight for alpha diversity and associated p-values for each condition were evaluated for Shannon (Fig. 14) and Chao1 (Fig.15). For Shannon, the weighted alpha diversity mean was 4.78 for the healthy group, 4.27 for the Colitis group, and 4.49 for the Colitis + Inulin group. All comparisons yielded statistically significant results with p-values of 0.000, indicating a significant difference among groups. In the case of Chao, the mean weight was 1891,455 (Healthy), 553,42 (Colitis), and 1552 (Colitis + Inulin), with corresponding p-values of 0.000. These values suggest a significant reduction in both Chao and Shannon indices in the Colitis condition, which appears to be reversed with Inulin supplementation. These differences were statistically significant in all studies (p = 0.000 for Healthy and Colitis; p = 0.000 for Colitis + Inulin), indicating a consistent trend across multiple datasets. The values of the alpha diversity in the 3rd condition (colitis + inulin) are closer to the healthy state meaning that inulin is trying to reverse the dysbiosis induced by colitis.

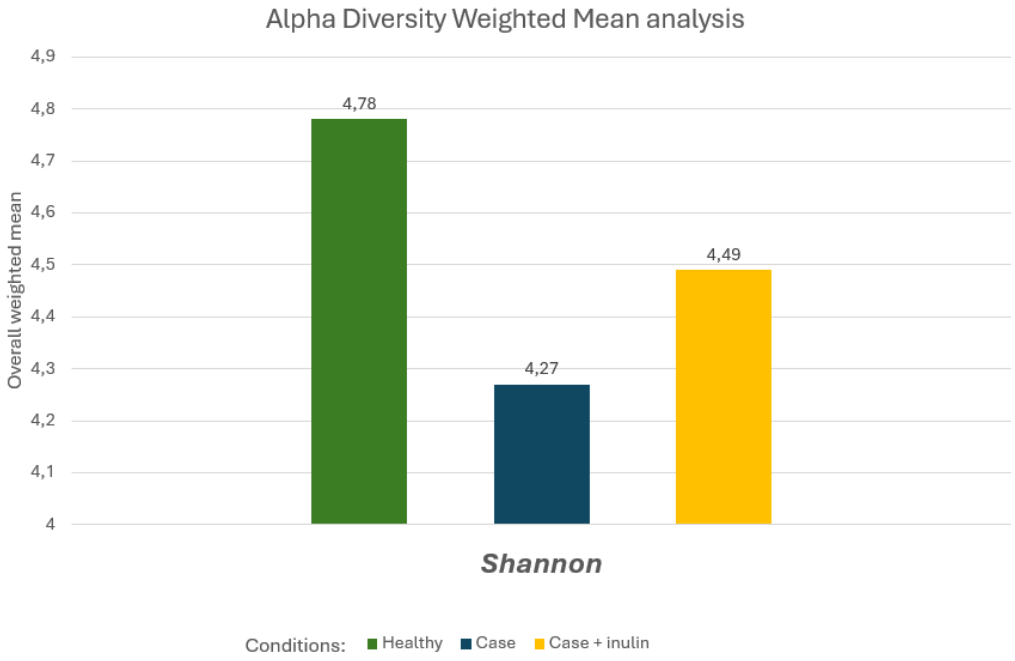


Figure 14 - *Bar Plot showing the overall mean weight of each condition with their p-value, after the stratification command has been conducted on the Shannon index data for each condition.*

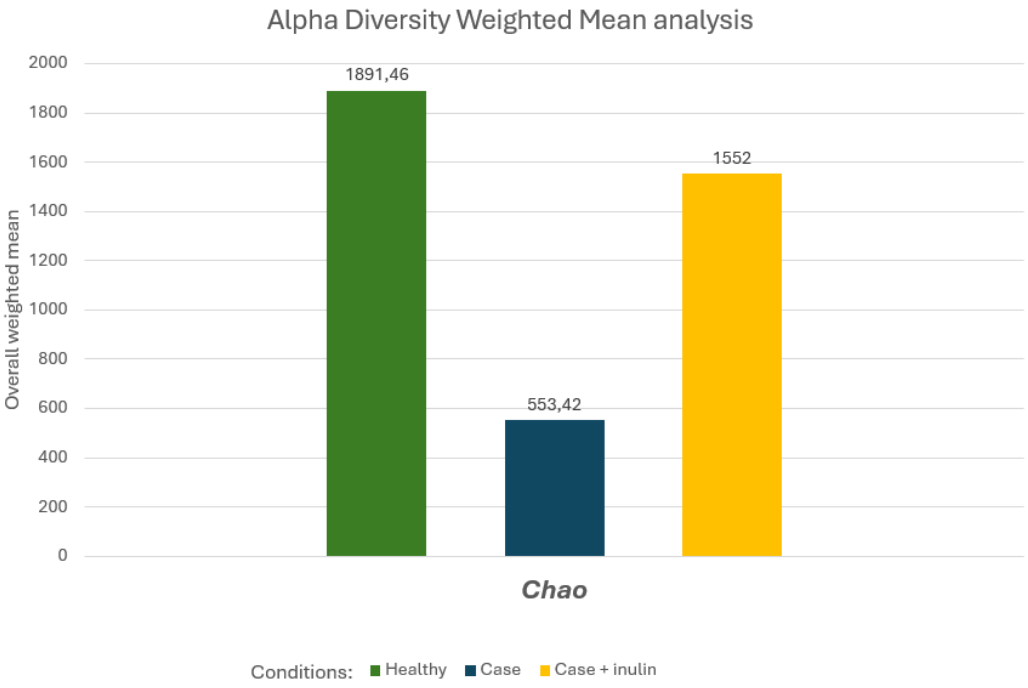


Figure 15 - *Bar Plot showing the overall mean weight of each condition with their p-value, after the stratification command has been conducted on the Chao index data for each condition.*

3.2.2 Mean difference meta-analysis of alpha diversity indices

We conducted a mean difference meta-analysis (SMD), with the help of STATA SE 13.1 package, on the alpha diversity data (see Table 13 on the Appendix chapter) while using the following commands:

metan n0 x0 sd0 n1 x1 sd1, randomi ***nostandard*** *label(namevar=author,yearvar=year)*
metan n0 x0 sd0 n2 x2 sd2, randomi ***nostandard*** *label(namevar=author,yearvar=year)*
metan n1 x1 sd1 n2 x2 sd2, randomi ***nostandard*** *label(namevar=author,yearvar=year)*

Shannon Diversity Index:

Table 5 presents the results of the meta-analysis. The estimated difference for the comparison between healthy and colitis conditions, denoted as Healthy vs IBD, was 0.493, with a statistically significant p-value of 0.009. The comparison between the IBD + inulin condition and IBD alone, denoted as IBD vs IBD+inulin, was -0.262, with a significant p-value of 0.040. The difference for the second comparison between the healthy and the treated, denoted as Healthy vs IBD + Inulin, was 0.289, with a non-statistically significant p-value of 0.125.

	Shannon				
	Effect	**p-value**	**LCI 95%**	**UCI 95%**	**studies**
Healthy vs IBD (0 vs 1)	0.493	0.009	0.126	0.861	6
Healthy vs IBD + Inulin (0 vs 2)	0.289	0.125	-0.080	0.658	6
IBD vs IBD + Inulin (1 vs 2)	-0.262	0.040	-0.512	-0.012	6

Table 5 *- Table showing the results of the mean difference meta-analysis from the Shannon alpha diversity index. Variables are respectively equal to:* ***(healthy-ibd), (healthy-ibd+inulin), and (ibd-ibd+inulin).***

Chao1 Richness Index:

Table 6 shows the effect for the healthy vs. IBD comparison (*Healthy vs IBD*) was 185.019, with a p-value of 0.082, suggesting a non-significant difference in species richness. The effect for the Healthy vs. IBD + inulin comparison was 150.470, with a p-value of 0.128. The difference interpreted as *IBD vs IBD + inulin* was -16.845, with a p-value of 0.391.

	Chao				
	Effect	**p-value**	**LCI 95%**	**UCI 95%**	**studies**
Healthy vs IBD (0 vs 1)	185.019	0.082	-23.809	393.847	5
Healthy vs IBD + Inulin (0 vs 2)	150.470	0.128	-88.788	389.727	5
IBD vs IBD + Inulin (1 vs 2)	-16.845	0.391	-55.317	21.626	5

Table 6 - *Table showing the results of the mean difference meta-analysis from the Chao alpha diversity index. Variables are respectively equal to:* ***(healthy-ibd), (healthy-ibd+inulin), and (ibd-ibd+inulin).***

The results of the meta-analyses for both the Shannon and Chao1 diversity indices are presented in Figures 16 and 17. When the value of the second comparison (healthy vs. IBD + inulin) approaches zero (i.e., $d \rightarrow 0$), it suggests that the IBD + inulin condition is becoming more similar to the healthy condition, which may indicate microbial restoration following inulin supplementation. Based on the results from the Shannon alpha diversity index, we observe that the diversity in healthy individuals is greater than in both IBD and IBD + inulin groups [healthy > IBD and healthy > (IBD + inulin)], while the diversity in the IBD + inulin group is higher than in the IBD group [IBD < (IBD + inulin)]. This explains the negative effect size observed for the IBD vs. IBD + inulin comparison. Notably, both the first comparison (healthy vs. IBD) and the third comparison (IBD vs. IBD + inulin) yielded statistically significant p-values, whereas the second comparison (healthy vs. IBD + inulin) did not reach statistical significance ($p > 0.05$). This pattern suggests that inulin supplementation has a meaningful impact on microbial diversity, as it results in a significant difference between the IBD and IBD + inulin groups, while reducing the difference between the healthy and IBD + inulin groups.

In contrast, the meta-analysis of the Chao1 alpha diversity index did not reveal any statistically significant differences across the comparisons, indicating that inulin supplementation did not have a measurable impact on species richness in this case.

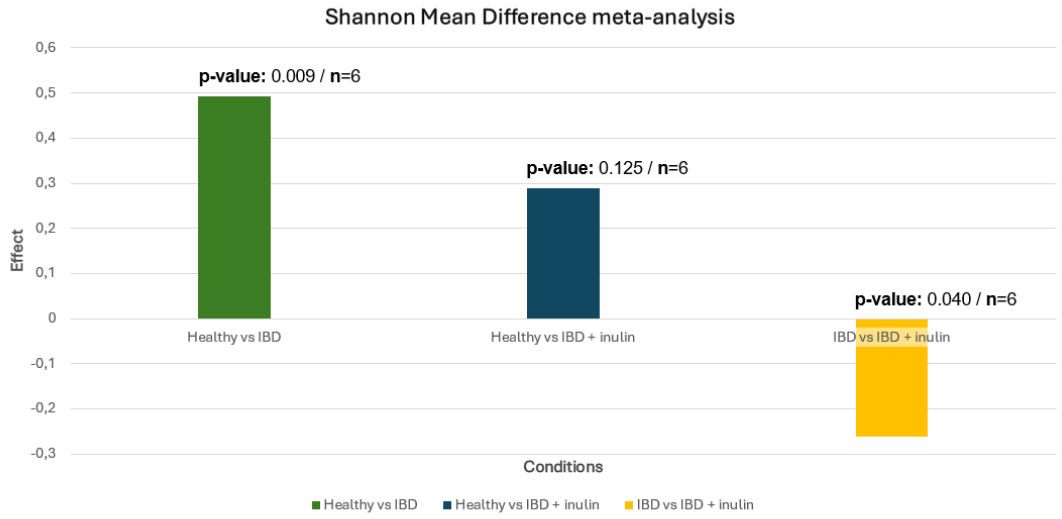


Figure 16 *- Bar plot illustrating the results of the mean difference meta-analysis of Shannon diversity indexes across three groups: healthy vs. IBD (green), healthy vs. IBD with inulin supplementation (blue), and IBD vs. IBD with inulin supplementation (yellow). P-values and sample sizes from the included studies are indicated above each bar.*

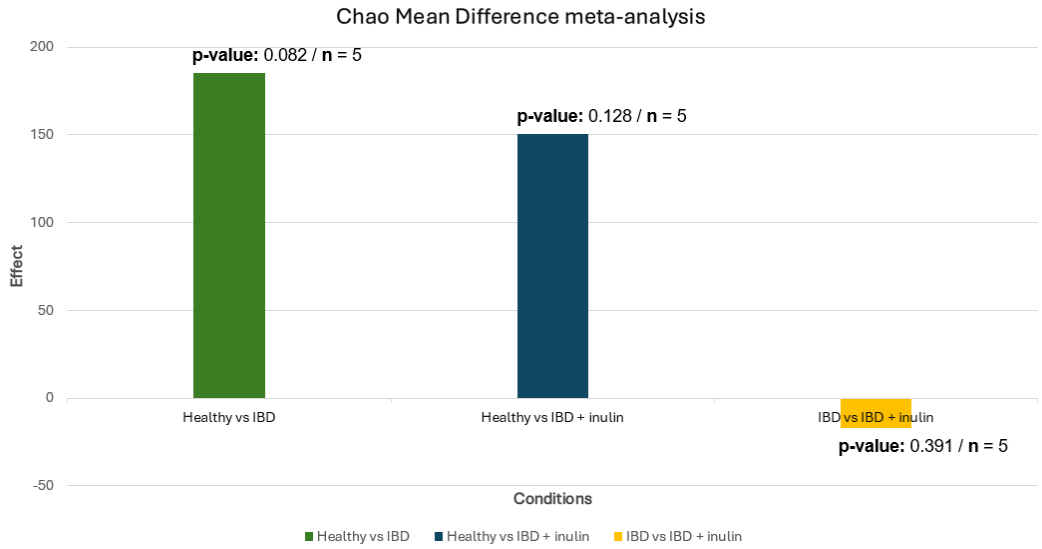


Figure 17 *- Bar plot illustrating the results of the mean difference meta-analysis of Chao diversity indexes across three groups: healthy vs. IBD (green), healthy vs. IBD with inulin supplementation (blue), and IBD vs. IBD with inulin supplementation (yellow). P-values and sample sizes from the included studies are indicated above each bar.*

3.2.3 Multivariate standardized mean difference meta-analysis of alpha diversity indices

We conducted a multivariate standardized mean difference meta-analysis (SMD), with the help of STATA SE 13.1 package, on the alpha diversity data (see Table 13 on the Appendix chapter) while using the following commands:

```
mvmeta b V, vars(b3 b4)
di _b[b4] - _b[b3]
lincom _b[b4] - _b[b3]
```

Shannon Diversity Index:

Table 7 presents the results of the meta-analysis. The estimated coefficient for the comparison between healthy and colitis conditions, denoted as b_3, was 1.111, with a statistically significant p-value of 0.033. The coefficient for the comparison between the colitis + inulin condition and colitis alone, denoted as b_4, was 0.503, with a non-significant p-value of 0.267. The difference between these two comparisons, $d = b_4 - b_3$, was -0.608, with a p-value of 0.130.

	Shannon				
	Coef.	**P>|z|**	**LCI 95%**	**UCI 95%**	**studies**
b3 = *(healthy-case)*	1.111	0.033	0.091	2.131	6
b4 =*(case+inulin) – case)*	0.503	0.267	-0.021	1.026	6
d= *b4-b3 = (case+inulin) - healthy*	-0.608	0.130	-1.395	0.179	6

Table 7 *- Table showing the results of the multivariate meta-analysis from the Shannon alpha diversity index. Variables b3, b4 and d are respectively equal to:* ***b3 = (healthy-case), b4=(case+inulin) – case) and d= b4-b3.***

Chao1 Richness Index:

Table 8 shows the coefficient for the healthy vs. colitis comparison (b_3) was 1.029, with a p-value of 0.054, suggesting a marginally non-significant difference in species richness. The coefficient for the colitis + inulin vs. colitis comparison (b_4) was 0.480, with a p-value of 0.137. The calculated difference $d = b_4 - b_3$ (interpreted as inulin - healthy) was -0.548, with a p-value of 0.198.

	Chao				
	Coef.	**P>|z|**	**LCI 95%**	**UCI 95%**	**studies**
b3 = *(healthy-case)*	1.029	0.054	-0.02	2.073	5
b4 =*(case+inulin) – case)*	0.480	0.137	-0.15	1.113	5
d= *b4-b3 = (case+inulin) - healthy*	-0.548	0.198	-1.38	0.286	5

Table 8 - *Table showing the results of the multivariate meta-analysis from the Chao alpha diversity index. Variables b3, b4 and d are respectively equal to:* ***b3 = (healthy-case), b4=(case+inulin) – case) and d= b4-b3.***

The findings from the meta-analyses for the Shannon and Chao1 indices are also presented in Figure 18. As the value of d is approximately equal to zero (d~0) or tends to zero (d → 0), the closer condition 3 is to condition 1, which would indicate microbe restoration after inulin administration. As we can see from our results b3 > 0 and b4 > 0, meaning that the (healthy > colitis) and [(colitis + inulin) > colitis] but [healthy > (colitis + inulin)], that's why we have a negative value of d. No value had a significant p-value lower than 0,05 except for b3 from Shannon. These findings indicate that inulin has a significant impact because the d value (Inulin-Healthy) is small and statistically not significant in both alpha diversity indices.

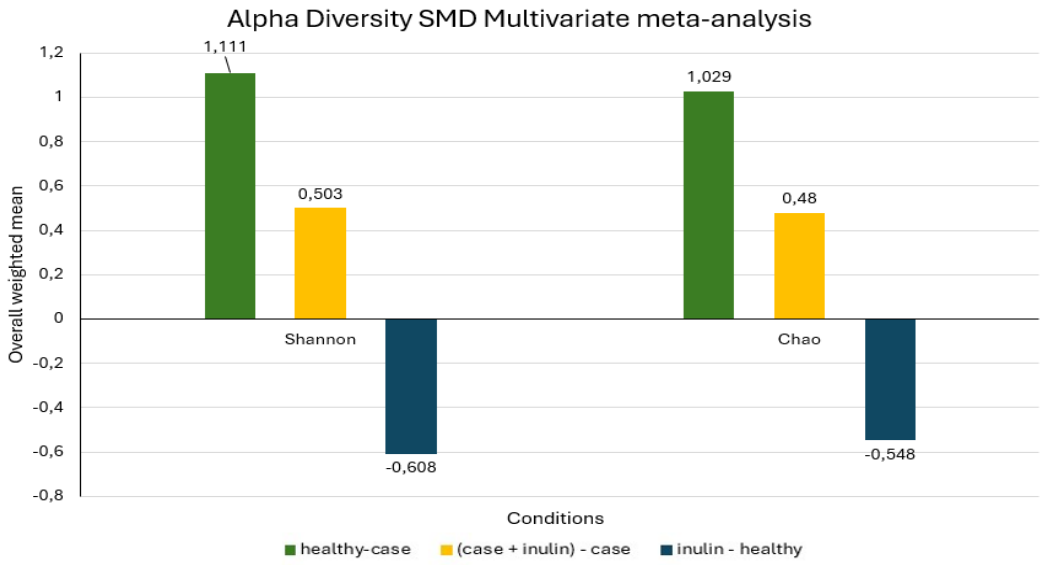


Figure 18 - *Bar Plot showing SMD values from the multivariate meta-analysis of both Shannon and Chao indexes. Variables b3, b4 and d are respectively equal to:* ***(healthy-case), (case+inulin) – case) and (case+inulin)-healthy.***

3.3 Meta-analysis of the microbiota composition in the phylum level

Relative abundance refers to the percentage of a specific type of organism compared to the total number of organisms present in each given area [52]. In this study we are focusing on bacteria and more specifically on the taxonomical unit of ***Phylum.*** Bacterial phyla are the main groups that make up the domain Bacteria [50]. In microbiology, the classification into phyla helps organize the immense diversity of bacteria into broader, more manageable categories. Examples of bacterial phyla include **Proteobacteria, Firmicutes,** Actinobacteria, and **Bacteroidota** [51].

3.3.1 Subgroup meta-analysis of relative abundance data in Firmicutes, Bacteroidetes and Proteobacteria

We conducted a weighted mean subgroup meta-analysis, with the help of STATA SE 13.1 package, on the relative abundance data (see Table 9) while using the following commands:

metan data stde if taxa=="Firmicutes", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Bacteroidota", by(condition_num) random label(namevar=author, yearvar=year)

metan data stde if taxa=="Proteobacteria", by(condition_num) random label(namevar=author, yearvar=year)

Table 9 categorizes the relative abundance data by taxonomic group (taxa). For each phylum, we included variables such as sample size (number of animals per group), data (relative abundance percentages), sd (standard deviation), and se (standard error) to provide a comprehensive summary of the data before we performed the analysis.

PMID	author	year	taxa	condition	data	stde	sample size	sd
38024237	Zhang Z	2023	Firmicutes	healthy	46.3	0.249	4	0.499
37659809	Li Z	2023	Firmicutes	healthy	27.77	0.183	6	0.448
37659809	Li Z	2023	Firmicutes	healthy	27.77	0.183	6	0.448
35527507	Huang W	2022	Firmicutes	healthy	71.65	0.202	5	0.451
30589960	Cai Y	2019	Firmicutes	healthy	57.5	0.202	6	0.494
38024237	Zhang Z	2023	Firmicutes	case	79.5	0.202	4	0.404
37659809	Li Z	2023	Firmicutes	case	63.67	0.196	6	0.481
37659809	Li Z	2023	Firmicutes	case	63.67	0.196	6	0.481
35527507	Huang W	2022	Firmicutes	case	58.52	0.220	5	0.493
30589960	Cai Y	2019	Firmicutes	case	32.7	0.192	6	0.469
38024237	Zhang Z	2023	Firmicutes	case and inulin	52.5	0.223	5	0.499
37659809	Li Z	2023	Firmicutes	case and inulin	39.21	0.218	5	0.488
37659809	Li Z	2023	Firmicutes	case and inulin	35.2	0.195	6	0.478
35527507	Huang W	2022	Firmicutes	case and inulin	66.61	0.211	5	0.472
30589960	Cai Y	2019	Firmicutes	case and inulin	51.22	0.204	6	0.500
38024237	Zhang Z	2023	Bacteroidota	healthy	34.4	0.238	4	0.475
37659809	Li Z	2023	Bacteroidota	healthy	60.74	0.199	6	0.488
37659809	Li Z	2023	Bacteroidota	healthy	60.74	0.199	6	0.488
35527507	Huang W	2022	Bacteroidota	healthy	22.84	0.188	5	0.420
30589960	Cai Y	2019	Bacteroidota	healthy	36.46	0.196	6	0.481
38024237	Zhang Z	2023	Bacteroidota	case	10.8	0.155	4	0.310
37659809	Li Z	2023	Bacteroidota	case	24.08	0.175	6	0.428
37659809	Li Z	2023	Bacteroidota	case	24.08	0.175	6	0.428
35527507	Huang W	2022	Bacteroidota	case	8.35	0.124	5	0.277
30589960	Cai Y	2019	Bacteroidota	case	58.09	0.201	6	0.493
38024237	Zhang Z	2023	Bacteroidota	case and inulin	43	0.221	5	0.495
37659809	Li Z	2023	Bacteroidota	case and inulin	42.95	0.221	5	0.495
37659809	Li Z	2023	Bacteroidota	case and inulin	49.78	0.204	6	0.500
35527507	Huang W	2022	Bacteroidota	case and inulin	9.83	0.133	5	0.298
30589960	Cai Y	2019	Bacteroidota	case and inulin	41.87	0.201	6	0.493
37659809	Li Z	2023	Proteobacteria	healthy	1.33	0.047	6	0.115
37659809	Li Z	2023	Proteobacteria	healthy	1.33	0.047	6	0.115
30589960	Cai Y	2019	Proteobacteria	healthy	4.5	0.085	6	8.228
37659809	Li Z	2023	Proteobacteria	case	2.17	0.059	6	0.146
37659809	Li Z	2023	Proteobacteria	case	2.17	0.059	6	0.146
30589960	Cai Y	2019	Proteobacteria	case	7.43	0.107	6	0.262
37659809	Li Z	2023	Proteobacteria	case and inulin	8.21	0.123	5	0.274
37659809	Li Z	2023	Proteobacteria	case and inulin	5.83	0.096	6	0.234
30589960	Cai Y	2019	Proteobacteria	case and inulin	4.51	0.085	6	0.208

Table 9 - *The Relative abundance data categorized by taxa and by condition for the subgroup meta-analysis.*

Below are the forest plots (Fi.19, Fig.20 and Fig.21) generated using the ***metan*** command, presenting the relative abundance data grouped by condition for each of the three bacterial phyla under investigation.

The forest plots display clear stratification across conditions for each bacterial phylum. For **Firmicutes** and **Bacteroidota**, inulin supplementation in colitis-induced mice moves the microbial profiles closer to that of the healthy group, suggesting a potential restorative effect. In contrast, **Proteobacteria** showed a continual increase from healthy to colitis and further in colitis + inulin, indicating that inulin did not mitigate this particular dysbiosis. All results across conditions were statistically significant ($p < 0.05$), supporting the validity of the observed microbial shifts.

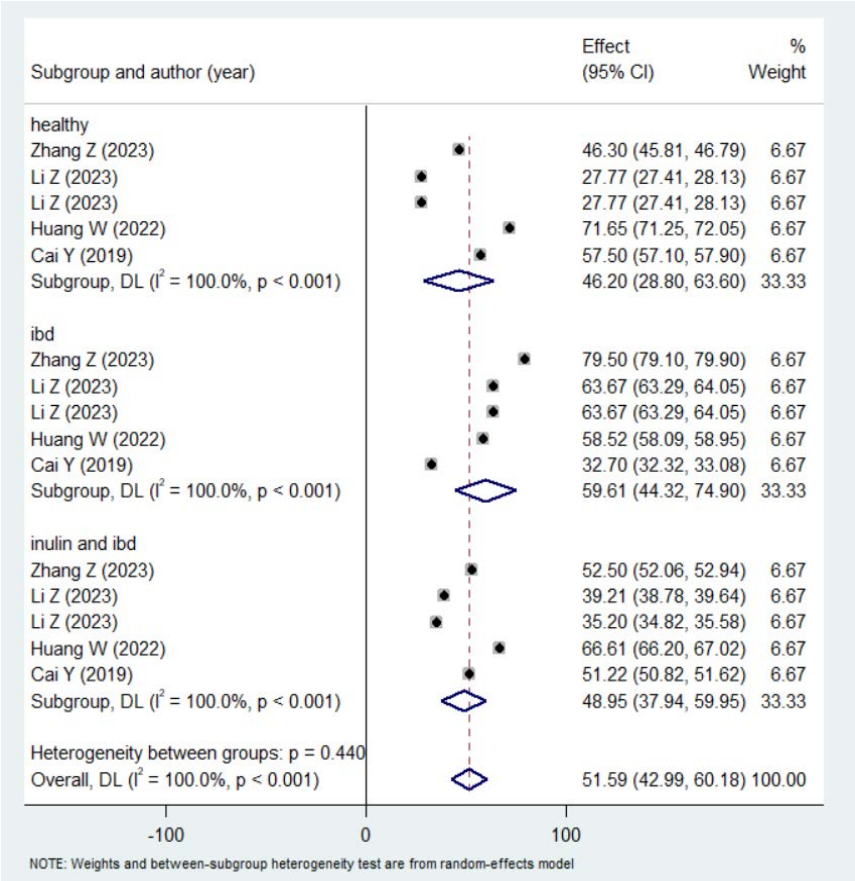


Figure 19 - *Firmicutes metan weighted mean subgroup meta-analysis forest plot.*

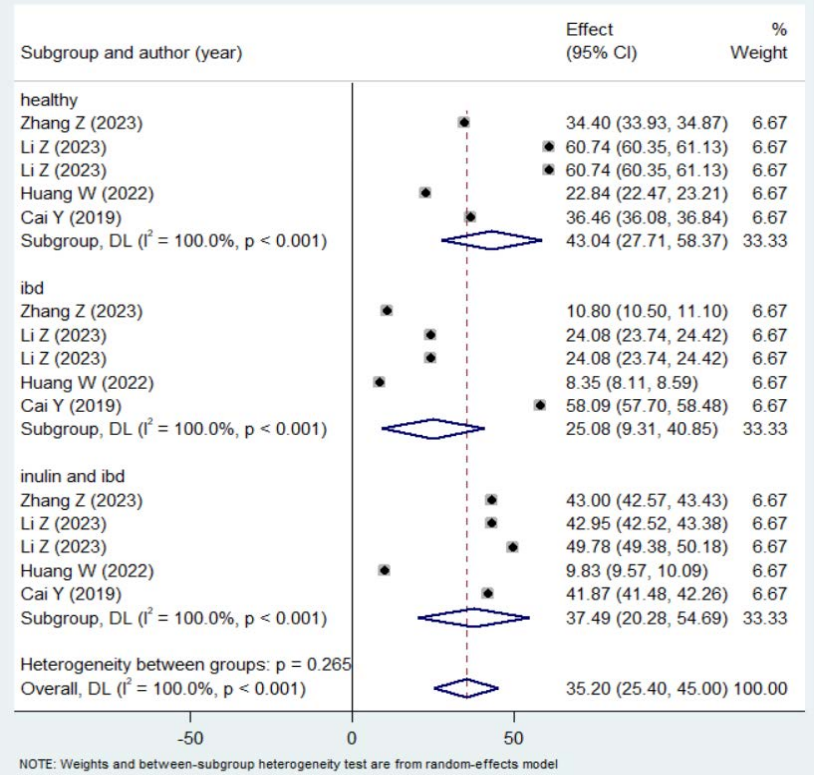


Figure 20 - *Bacteroidota metan weighted mean subgroup meta-analysis forest plot.*

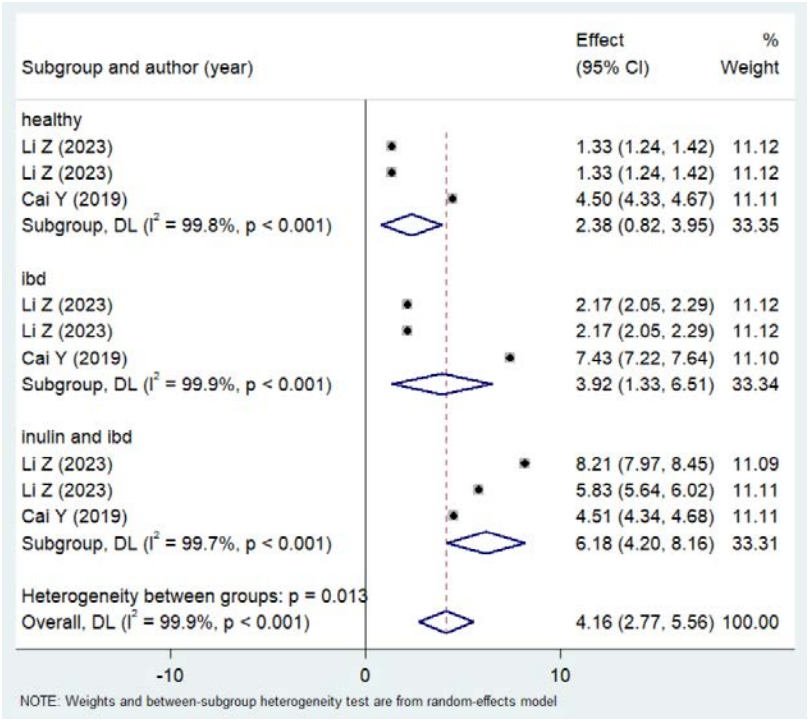


Figure 21 – *Proteobacteria metan weighted mean subgroup meta-analysis forest plot.*

The Huang study was excluded from the Proteobacteria analysis, because we had missing values for the control group. The mean relative abundance and associated p-values for each condition were evaluated for Firmicutes, Bacteroidota, and Proteobacteria. For **Firmicutes**, the mean relative abundances were 46.20 for the Healthy group, 59.61 for the Colitis group, and 49.00 for the Colitis + Inulin group. All comparisons yielded statistically significant results with p-values of 0.000, indicating a significant difference among groups. In the case of **Bacteroidota**, the mean abundances were 36.46 (Healthy), 25.08 (Colitis), and 37.50 (Colitis + Inulin), with corresponding p-values of 0.000, 0.002, and 0.000 (Fig.22). These values suggest a significant reduction in Bacteroidota abundance in the Colitis condition, which appears to be reversed with Inulin supplementation. For **Proteobacteria**, the observed mean abundances were 2.39 in the Healthy group, 3.92 in the Colitis group, and 6.18 in the Colitis + Inulin group. These differences were statistically significant in 3 out of 4 studies ($p = 0.003$ for Healthy and Colitis; $p = 0.000$ for Colitis + Inulin), indicating a consistent trend across multiple datasets.

Similar to our findings on alpha diversity analysis, the values of the relative abundance for the examined Phyla of the 3rd condition (colitis+inulin) are closer to the healthy state compared to the values of the 2nd condition (colitis). These data suggest that administration of inulin reverses the dysbiosis caused by colitis on the Firmicutes and Bacteroidota phylum. However, this is not the case in Proteobacteria, as we can see a rapid increase from the 1st condition to the 2nd and then to the 3rd, meaning that inulin did not have an effect.

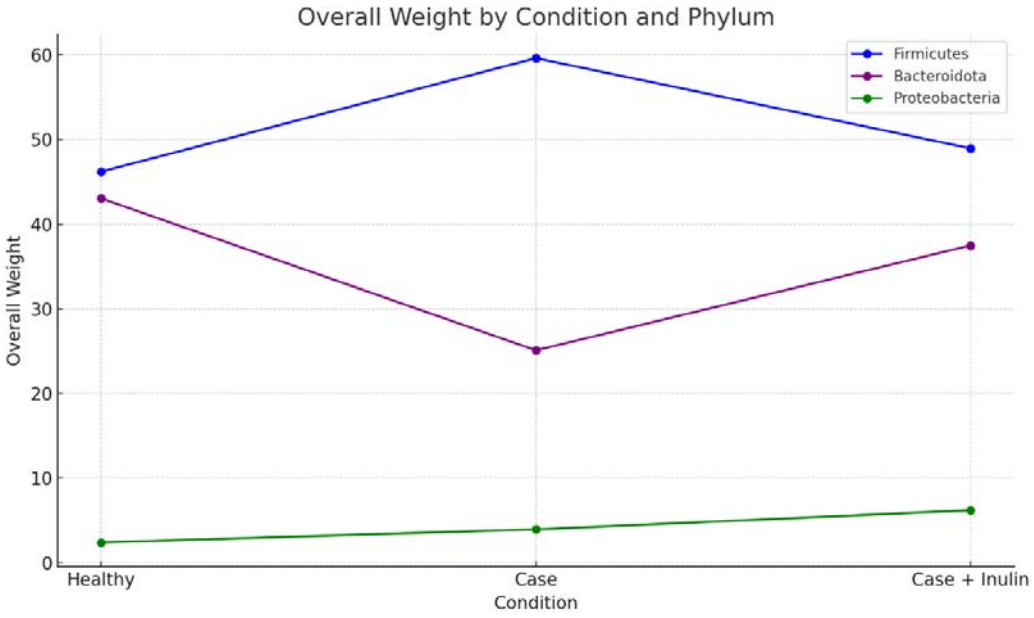


Figure 22 - *Line plot showing the overall mean weight of each, after the subgroup analysis command has been conducted on the Relative Abundance data for each condition*

3.3.2 Multivariate Mean Difference meta-analysis of relative abundance data

We conducted a multivariate mean difference meta-analysis with the help of STATA SE 13.1 package, on the relative abundance data (see Table 14 on the Appendix chapter) while using these commands:

```
mvmeta b V, vars (b3 b4)
di _b[b4] - _b[b3]
lincom _b[b4] - _b[b3]
```

We used the Bernoulli formula in order to calculate first the SD then the SE and then be able to perform the multivariate meta-analysis.

Firmicutes		**Coef.**	**P>|z|**	**LCI 95%**	**UCI 95%**	**studies**
	b3 *=(healthy-case)*	-26.387	0.337	-80.294	27.521	5
	b4 *=(case+inulin) – case)*	-21.037	0.305	-61.254	19.179	5
	d *= b4-b3 =>* *(case+inulin) - healthy*	5.349	0.461	-8.879	19.577	5

Table 10 - *Table showing the results of the multivariate meta-analysis from the Relative abundance data, specifically from the Phylum "Firmicutes". Variables b3, b4 and d are respectively equal to:* ***b3 = (healthy-case), b4=(case+inulin) – case) and d= (case+inulin)-healthy.***

Bacteroidota		**Coef.**	**P>|z|**	**LCI 95%**	**UCI 95%**	**studies**
	b3 *=(healthy-case)*	39.219	0.082	-4.982	83.420	5
	b4 *=(case+inulin) – case)*	25.709	0.163	-10.371	61.789	5
	d *= b4-b3 =>* *(case+inulin) - healthy*	-13.510	0.259	-36.991	9.971	5

Table 11 – *Table showing the results of the multivariate meta-analysis from the Relative abundance data, specifically from the Phylum "Bacteroidota". Variables b3, b4 and d are respectively equal to:* ***b3 = (healthy-case), b4=(case+inulin) – case) and d= (case+inulin)-healthy.***

Proteobacteria		Coef.	P>|z|	LCI 95%	UCI 95%	studies
	b3 *=(healthy-case)*	-3.157	**0.022**	-35.984	8.418	4
	b4 *=(case+inulin) – case)*	16.785	0.089	-2.198	28.532	4
	d *= b4-b3 =>* *(case + inulin)- healthy*	19.942	0.074	5.539	48.361	4

Table 12 – *Table showing the results of the multivariate meta-analysis from the Relative abundance data, specifically from the Phylum "Proteobacteria". Variables b3, b4 and d are respectively equal to:* ***b3 = (healthy-case), b4=(case+inulin) – case) and d= (case+inulin)-healthy.***

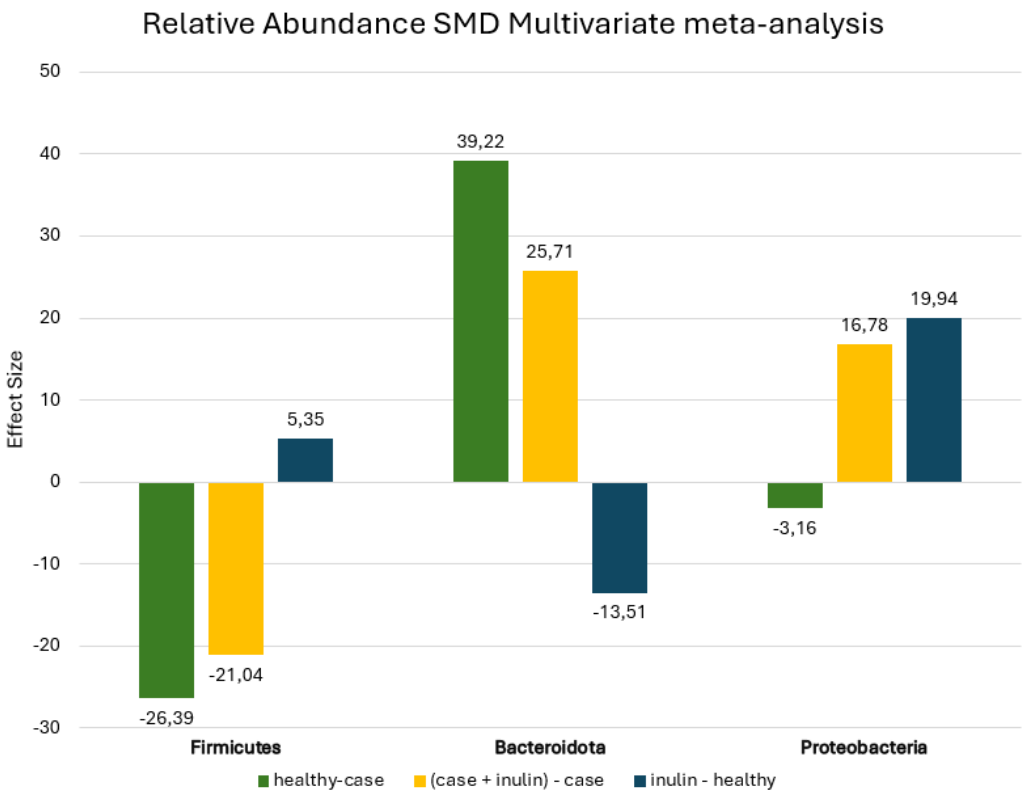


Figure 23 - *Bar Plot showing the results of the multivariate meta-analysis. Variables b3, b4 and d are respectively equal to:* ***(healthy -case), (case+inulin) – case) and d=(case+inulin)-healthy.***

For the phylum Firmicutes, the mean difference comparing the healthy and diseased (colitis) conditions (b3) was –26.39 (p = 0.337), while the mean difference comparing the inulin-treated group to the diseased group (b4) was –21.04 (p = 0.305) (Table 10). The difference between these estimates (d = b4 – b3), reflecting the contrast between inulin and healthy conditions, was 5.35 (p = 0.461). In the case of Bacteroidota, b3 (healthy-case) was 39.22 (p = 0.082) and b4 (case+inulin vs. case) was 25.71 (p = 0.163). The corresponding difference (d = b4 – b3) was –13.51 (p = 0.259) (Table 11). For Proteobacteria, using data the effect sizes was –3.16 for b3 (healthy-case, p = 0.022) and 16.78 for b4 (case+inulin vs. case, p = 0.089), yielding a difference (d = b4 – b3) of 19.94, which reached statistical significance (p = 0.074) (Table 12).

A positive value of *d* (i.e., $d > 0$) indicates that the condition with inulin supplementation is shifting closer to the healthy state, suggesting microbial restoration from the diseased to the healthy profile. We need b3 and b4 to have the same directions, in this case we need both to have negative values so that we can see when the 1st and 3rd condition come close to each other.

As we can see from the results of Firmicutes b3 < 0 and b4 < 0, meaning that the (healthy < colitis) and [(colitis + inulin) < colitis] but [healthy > (colitis + inulin)], that's why we have a positive value of d.

As for the results of Bacteroidota, b3 and b4 are positive, meaning that both healthy and colitis + inulin values are bigger than the colitis values. The d value is negative, which implies that colitis + inulin > healthy.

The Proteobacteria results show a negative b3 (colitis > healthy), which had the only significant p-value of 0.022, a positive b4 (case+inulin > colitis) and a positive d (case+inulin > colitis). These findings suggest that inulin has a significant impact because the d value (Inulin-Healthy) is small and statistically not significant in the case of Firmicutes and Bacteroidota, but not with Proteobacteria (Fig.23).

4. DISCUSSION

Inulin, a well-known prebiotic fiber, was selected as the focal point of this study due to its extensively documented health benefits, particularly in the context of gut microbiome regulation. Prebiotics serve as nutritional substrates for probiotics, beneficial microorganisms residing in the gastrointestinal tract. Several studies have demonstrated that inulin supplementation can alleviate dysbiosis, particularly in conditions such as inflammatory bowel disease (IBD), by promoting the growth of beneficial gut bacteria and supporting the restoration of a healthy microbial ecosystem. Specifically, inulin has been shown to increase the abundance of beneficial bacteria such as *Lactobacillus (classified within the phylum Firmicutes)* and *Bifidobacterium (classified within the phylum Actinobacteria)*, while reducing the prevalence of potentially harmful bacteria like *Desulfovibrio (classified within the phylum Thermodesulfobacteriota, although previously were on Proteobacteria)*. These microbial shifts contribute to the reestablishment of a balanced and resilient gut microbiota [1]. Through this mechanism, inulin facilitates the regeneration of a dysbiotic microbiome, enabling it to return to a state more similar to that of a healthy one.

To evaluate the potential regenerative effects of inulin on a dysbiotic gut microbiome, we conducted a meta-analysis using data from murine studies in which colitis was experimentally induced using dextran sulfate sodium (DSS). Our analysis employed STATA SE 13.1 and focused on two major components of microbial community analysis: alpha diversity and relative abundance. For alpha diversity, we used the Chao1 and Shannon indices, which respectively reflect species' richness and a combination of richness and evenness. In terms of taxonomic composition, we focused on relative abundance at the phylum level, particularly *Firmicutes*, *Bacteroidota*, and *Proteobacteria*, which were consistently reported across the selected studies.

Following the PRISMA guidelines, we conducted a comprehensive literature search on the PubMed database using two specific search terms, which initially yielded 1,144 articles. After a rigorous screening process, we identified 7 studies eligible for inclusion in the meta-analysis. Of these, 6 studies provided alpha diversity data, and 5 included relative abundance data. In cases where a single study included multiple treatment groups with different inulin dosages, we treated them as separate datasets to accurately reflect variability in treatment conditions.

Prior to statistical analysis, we categorized the data into three experimental conditions: healthy controls, DSS-induced colitis (case), and DSS-induced colitis with inulin supplementation (case + inulin). This classification was applied uniformly across all measures—both diversity indices and relative abundance data.

The meta-analysis of alpha diversity using standardized mean difference (SMD) showed that the values for the inulin-treated group moved closer to those of the healthy control group, although the differences were statistically non-significant. Nevertheless, this finding is meaningful, as it suggests a trend toward microbial restoration, implying that inulin may help reverse the loss of microbial diversity caused by DSS-induced colitis.

In our stratified meta-analysis, which grouped results by condition, we observed a consistent pattern: alpha diversity indices decreased in the diseased state and partially recovered with inulin supplementation. Similar trends were seen in the relative abundance data for *Firmicutes* and *Bacteroidota*, both of which showed reductions in the diseased group and partial restoration in the inulin-treated group. In contrast, *Proteobacteria* exhibited a continuous increase from the

healthy to the colitis to the inulin-treated condition, indicating that inulin had no observable effect in reversing the overrepresentation of this phylum.

This outcome aligns with existing literature, which describes IBD and colitis as conditions characterized by a decline in beneficial bacterial phyla such as *Firmicutes* and *Bacteroidota*, alongside an overgrowth of *Proteobacteria, which is associated with disease pathogenesis* [11]. Our results further support this pattern, providing evidence that inulin can contribute to the partial restoration of microbial homeostasis, though its effect appears limited with regard to certain bacterial groups.

In summary, our findings support the hypothesis that inulin acts as a beneficial prebiotic in the context of microbiome dysbiosis. The observed trends toward restoration of alpha diversity and the relative abundance of key beneficial phyla suggest that inulin may facilitate microbiome regeneration and, consequently, contribute to improved host health. However, the effect was not uniform across all taxa, as inulin did not appear to mitigate the rise of *Proteobacteria*, a phylum often linked to inflammation and gut pathology. Future research should further explore these taxon-specific responses and consider dose-dependent effects, as well as long-term outcomes, to better elucidate inulin's therapeutic potential.

5. REFERENCES

[1] Yoo, S., Jung, S. C., Kwak, K., & Kim, J. S. (2024). The Role of Prebiotics in Modulating Gut Microbiota: Implications for Human Health. *International journal of molecular sciences*, *25*(9), 4834. https://doi.org/10.3390/ijms25094834

[2] Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *The Biochemical journal*, *474*(11), 1823–1836. https://doi.org/10.1042/BCJ20160510

[3] Lisa Catanese, ELS, Health Writer (2024).Prebiotics: Understanding their role in gut health

https://www.health.harvard.edu/nutrition/prebiotics-understanding-their-role-in-gut-health

[4] Qin, Y. Q., Wang, L. Y., Yang, X. Y., Xu, Y. J., Fan, G., Fan, Y. G., Ren, J. N., An, Q., & Li, X. (2023). Inulin: properties and health benefits. *Food & function*, *14*(7), 2948–2968. https://doi.org/10.1039/d2fo01096h

[5] Sheng, W., Ji, G., & Zhang, L. (2023). Immunomodulatory effects of inulin and its intestinal metabolites. Frontiers in immunology, 14, 1224092. https://doi.org/10.3389/fimmu.2023.1224092

[6] Markowiak, P., & Śliżewska, K. (2017). Effects of Probiotics, Prebiotics, and Synbiotics on Human Health. *Nutrients*, *9*(9), 1021. https://doi.org/10.3390/nu9091021

[7] Lee Y. H. (2018). An overview of meta-analysis for clinicians. *The Korean journal of internal medicine*, *33*(2), 277–283. https://doi.org/10.3904/kjim.2016.195

[8] Thukral, Ashwani. (2017). A review on measurement of Alpha diversity in biology. Agricultural Research Journal. 54. 1. 10.5958/2395-146X.2017.00001.1.

[9] CD GENOMICS-The Genomics Services company "Understanding Alpha Diversity: What It Is and Why It Matters"

https://www.cd-genomics.com/resource-understanding-alpha-diversity-what-it-is-why-it-matters.html

[10] Williams, C. E., Hammer, T. J., & Williams, C. L. (2024). Diversity alone does not reliably indicate the healthiness of an animal microbiome. The ISME journal, 18(1), wrae133. https://doi.org/10.1093/ismejo/wrae133

[11] Xu, X., Ocansey, D.K.W., Hang, S. *et al.* The gut metagenomics and metabolomics signature in patients with inflammatory bowel disease. *Gut Pathog* 14, 26 (2022). https://doi.org/10.1186/s13099-022-00499-9

[12] Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal*, 474(11), 1823–1836. https://doi.org/10.1042/BCJ20160510

[13] Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, N.J.

[14] Haidich A. B. (2010). Meta-analysis in medical research. *Hippokratia*, *14*(Suppl 1), 29–37.

[15] Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: A practical guide. *Frontiers in Physiology, 10*, Article 203. https://doi.org/10.3389/fphys.2019.00203

[16] Yang, Y., Zhao, P., Li, X., Guo, L., & Gao, W. (2022). The potential roles of natural plant polysaccharides in inflammatory bowel disease: A review. *Carbohydrate Polymers, 277*, 118821. https://doi.org/10.1016/j.carbpol.2021.118821

[17] 'Phylum: Bacterial taxa'. Wikipedia https://en.wikipedia.org/wiki/Phylum

[18] 'Bacterial Phyla'.https://en.wikipedia.org/wiki/Bacterial_phyla

[19] Plămadă, D., & Vodnar, D. (2021). Polyphenols—Gut microbiota interrelationship: A transition to a new generation of prebiotics. *Nutrients, 14*(1), 137. https://doi.org/10.3390/nu14010137

[20] Pardi DS, Cotter TG. Other diseases of the colon. In: Feldman M, Friedman LS, Brandt LJ, eds. Sleisenger and Fordtran's Gastrointestinal and Liver Disease. 11th ed. Philadelphia, PA: Elsevier; 2021:chap 128.

[21] Mayo Clinic. IBD Diagnostics and Treatment.

https://www.mayoclinic.org/diseases-conditions/inflammatory-bowel-disease/diagnosis-treatment/drc-20353320

[22] 'Metagenomics'. NHGRIIH. https://www.genome.gov/genetics-glossary/Metagenomics

[23] CD GENOMICS-The Use and Types of Alpha-Diversity Metrics in Microbial NGS

https://www.cd-genomics.com/microbioseq/the-use-and-types-of-alpha-diversity-metrics-in-microbial-ngs.html

[24] Średnicka, P., Juszczuk-Kubiak, E., Wójcicki, M., Kowalczyk-Akimowicz, M., & Roszko, M. (2021). Probiotics as a biological detoxification tool of food chemical contamination: A review. *Food and Chemical Toxicology, 153*, 112306. https://doi.org/10.1016/j.fct.2021.112306

[25] *Jayashantha*, Eranga. (2015). Actinobacteria- Morphology, Physiology, biochemistry, diversity & Industrial Applications of genus Actinobacteria. 10.13140/RG.2.1.2632.5928.

[26] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.)*, *372*, n71. https://doi.org/10.1136/bmj.n71

[27] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

[28] Stata'. Wikipedia, 10 April 2024. Wikipedia,

https://en.wikipedia.org/w/index.php?title=Stata&oldid=1218255260

[29] STATA. STATA Homepage, 23 April 2025.

https://www.stata.com/

[30] Stata manual, nlcom command, [Available online from: https://www.stata.com/manuals/rnlcom.pdf

[31] STATA SE 13.1 help mvmeta command

[32] STATA SE 13.1 help metan command

[33] STATA SE 13.1 help display command

[34] STATA SE 13.1 help lincom command

[35] STATA SE 13.1 help encode command

[36] STATA SE 13.1 help generate command

[37] Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: A practical guide. *Frontiers in Physiology, 10*, Article 203. https://doi.org/10.3389/fphys.2019.00203

[38] Gallardo-Gómez D, Richardson R, Dwan K. Standardized mean differences in meta-analysis: a tutorial. *Cochrane Ev Synth*. 2024; 2:e12047. doi:10.1002/cesm.12047

[39] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5 (updated August 2024). Cochrane, 2024. Available from www.training.cochrane.org/handbook.

[40] 'Standard Error (SE) Definition: Standard Deviation in Statistics Explained'. Investopedia, https://www.investopedia.com/terms/s/standard-error.asp.

[41]'Standard Deviation'. Wikipedia, 4 Ιούνιος 2024. Wikipedia, https://en.wikipedia.org/w/index.php?title=Standard_deviation&oldid=122716 1300.

[42] Wikipedia contributors. (2025, April 27). Bernoulli distribution. In *Wikipedia, The Free Encyclopedia*. Retrieved 08:57, June 3, 2025, from https://en.wikipedia.org/w/index.php?title=Bernoulli_distribution&oldid=1287698176

[43] 'SMD'.Cochrane Handbook

https://handbook-5-1.cochrane.org/chapter_9/9_2_3_2_the_standardized_mean_difference.htm

[44] 'Effect Size'. Statistics Solutions,

https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/effect-size/

[45] Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2022). Fixed-Effect vs Random-Effects Models for Meta-Analysis: 3 Points to Consider. *Global spine journal*, *12*(7), 1624–1626. https://doi.org/10.1177/21925682221110527

[46] 'Condifence Intervals'.nlm.nih.https://www.nlm.nih.gov/oet/ed/stats/02-950.html

[47] Wikipedia contributors. (2025, June 2). P-value. In *Wikipedia, The Free Encyclopedia*. Retrieved 09:06, June 3, 2025, from https://en.wikipedia.org/w/index.php?title=P-value&oldid=1293555710

[48] 'Bernoulli distribution'. Statistics by Jim. https://statisticsbyjim.com/probability/bernoulli-distribution/

[49] Richardson, M., Garner, P., & Donegan, S. (2019). Interpretation of subgroup analyses in systematic reviews: A tutorial. *Clinical Epidemiology and Global Health,* *7*(2), 192–198. https://doi.org/10.1016/j.cegh.2018.05.005

[50] T*he American Heritage New Dictionary of Cultural Literacy* (third ed.). Houghton Mifflin Company. 2005. Retrieved 4 October 2008. Phyla in the plant kingdom are frequently called divisions.

[51] Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. H., ... & Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology, 12*(9), 635–645. https://doi.org/10.1038/nrmicro3330

[52] CD Genomics. (n.d.). *Absolute abundance vs relative abundance in the microbiome*. Retrieved from https://www.cd-genomics.com/resource-absolute-abundance-relative-abundance.html

[53] Glen, S. (n.d.). *Mean Difference: Definition, Formula and Examples. Statistics How To.* https://www.statisticshowto.com/probability-and-statistics/hypothesis-testing/mean-difference/

6. APPENDIX

6.1 Mean difference (MD) meta-analysis

We performed a mean difference meta-analysis on the alpha diversity data. The following forest plots were a result of the STATA SE 13.1 metan command.

6.1.1 Shannon MD meta-analysis done with all the comparisons.

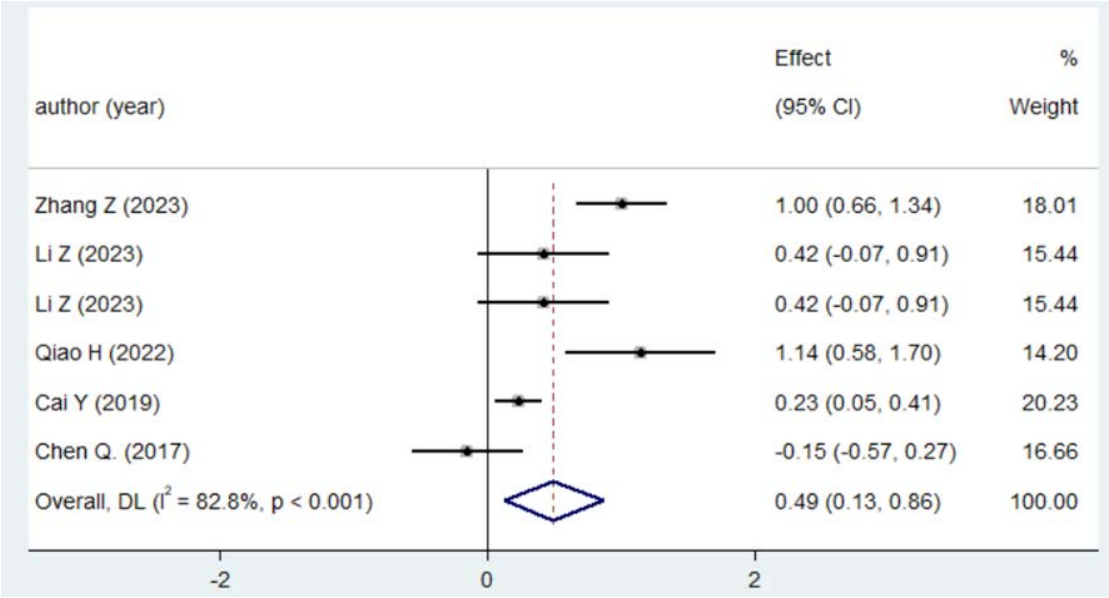


Figure 24 – *Forest Plot of the* ***Shannon*** *alpha diversity index while performing mean difference meta-analysis for the* ***Healthy vs IBD condition.***

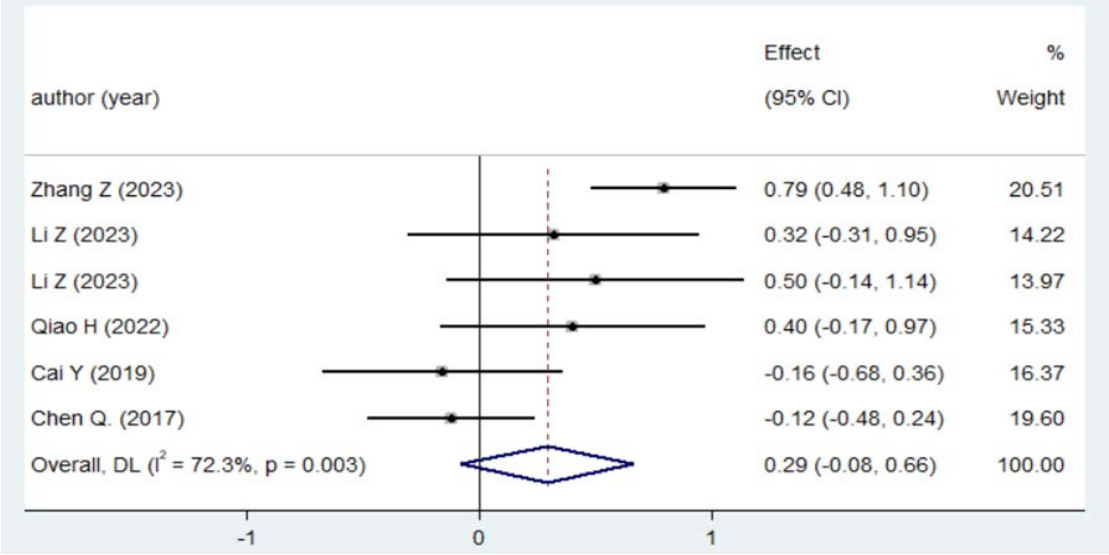


Figure 25 – *Forest Plot of the* ***Shannon*** *alpha diversity index while performing mean difference meta-analysis for the* ***Healthy vs IBD+inulin condition.***

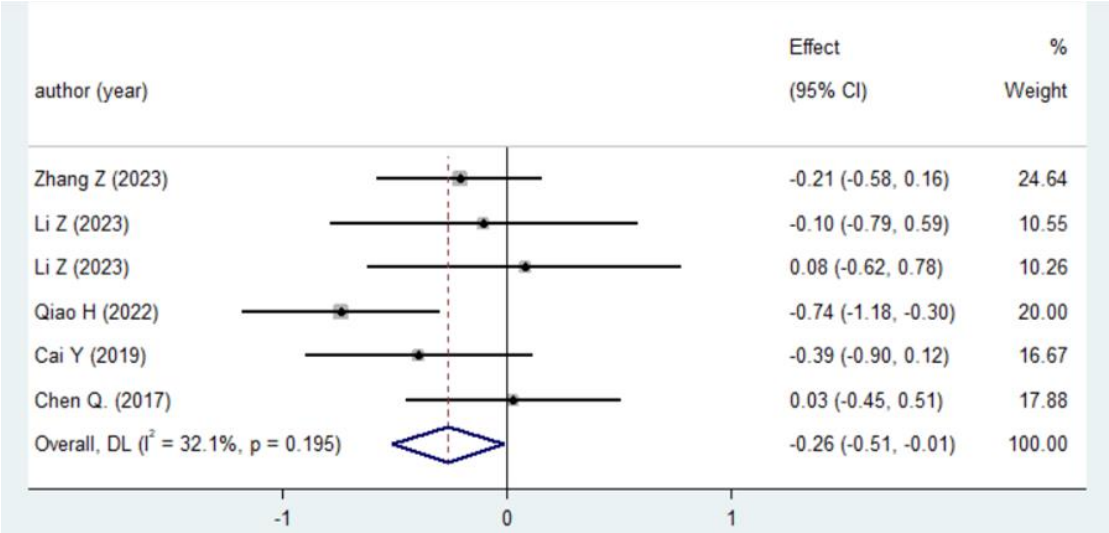


Figure 26 – *Forest Plot of the* ***Shannon*** *alpha diversity index while performing mean difference meta-analysis for the* ***IBD vs IBD+inulin condition.***

6.1.2 Chao MD meta-analysis done with all the comparisons.

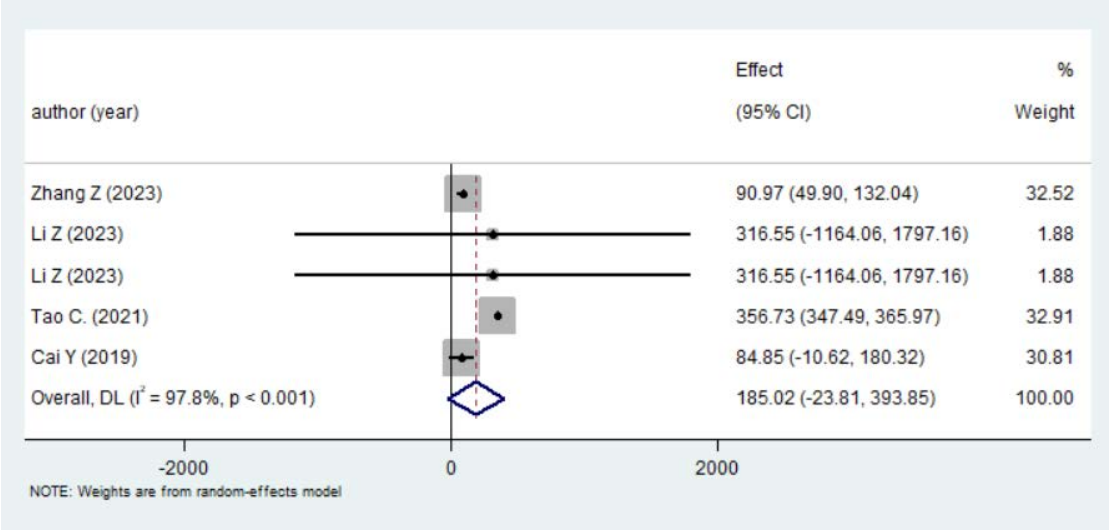


Figure 27 – *Forest Plot of the* ***Chao*** *alpha diversity index while performing mean difference meta-analysis for the* ***Healthy vs IBD condition.***

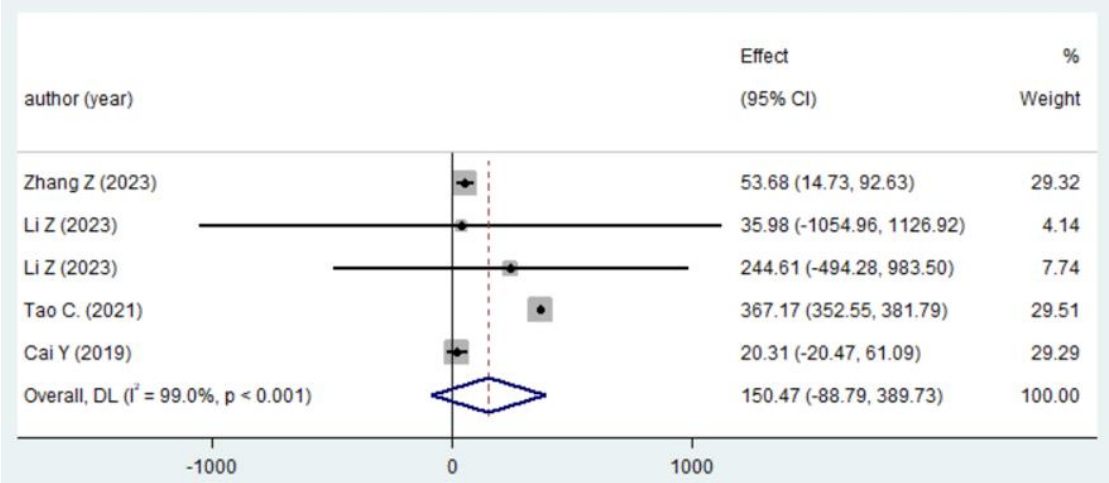


Figure 28 – *Forest Plot of the* ***Chao*** *alpha diversity index while performing mean difference meta-analysis for the* ***Healthy vs IBD+inulin condition****.*

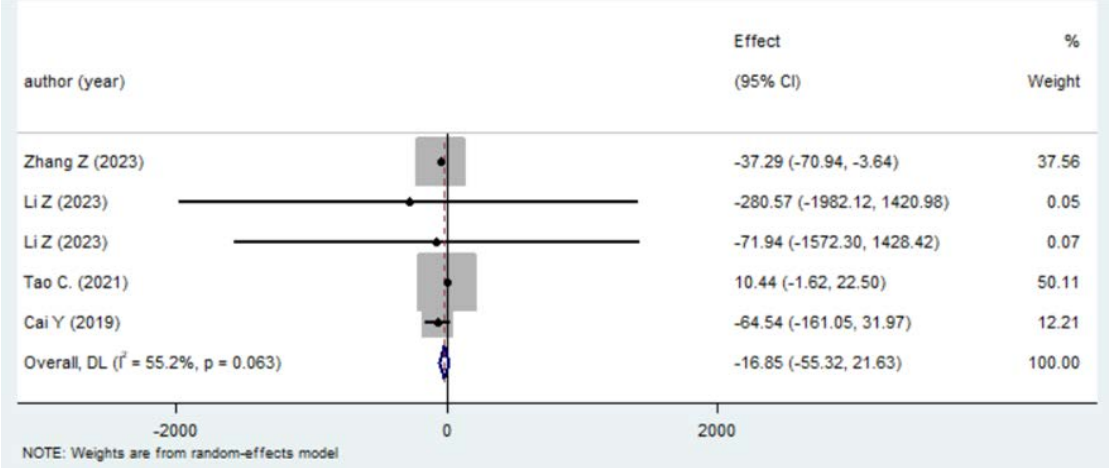


Figure 29 – *Forest Plot of the* ***Chao*** *alpha diversity index while performing mean difference meta-analysis for the* ***IBD vs IBD+inulin condition.***

6.2 Multivariate meta-analysis

We performed multivariate meta-analysis using SMD (Standardized Mean Difference) on the alpha diversity data and mean difference on the relative abundance data, because they were already normalised. The following tables of data were used on STATA SE 13.1 in order for the multivariate meta-analysis to be performed.

6.2.1 Alpha Diversity SMD multivariate meta-analysis data

We categorized the alpha diversity data by index and used the variables n for sample size, x for mean, se for standard error and sd for standard deviation, as seen in Table 13.

pmid	author	year	index	n0	x0	se0	sd0	n1	x1	se1	sd1	n2	x2	se2	sd2
38024237	Zhang Z	2023	Shannon	4	3.56	0.10	0.2	4	2.56	0.14	0.28	5	2.77	0.13	0.28
37659809	Li Z	2023	Shannon	6	4.13	0.14	0.35	6	3.71	0.21	0.50	5	3.81	0.29	0.64
37659809	Li Z	2023	Shannon	6	4.13	0.14	0.35	6	3.71	0.21	0.50	6	3.63	0.30	0.72
35415348	Qiao H	2022	Shannon	8	3.99	0.24	0.68	8	2.85	0.16	0.44	8	3.59	0.16	0.46
30589960	Cai Y	2019	Shannon	6	6.99	0.07	0.18	6	6.76	0.05	0.13	6	7.15	0.25	0.62
29232851	Chen Q.	2017	Shannon	13	5.85	0.10	0.35	10	6	0.19	0.6	12	5.97	0.16	0.54
38024237	Zhang Z	2023	Chao	5	305	16.42	36.71	5	214.03	13.02	29.11	4	251.32	11.20	22.39
37659809	Li Z	2023	Chao	6	5014.39	251.80	616.77	6	4697.84	712.23	1744.6	5	4978.41	496.40	1109.99
37659809	Li Z	2023	Chao	6	5014.39	251.80	616.77	6	4697.84	712.23	1744.6	6	4769.78	280.58	687.26
32813896	Tao C.	2021	Chao	3	414.93	4.48	7.75	3	58.20	1.492	2.58	3	47.76	5.97	10.34
30589960	Cai Y	2019	Chao	6	900.61	13.80	33.81	6	815.76	46.71	114.42	6	880.3	15.57	38.14

Table 13 – *Alpha diversity data for multivariate meta-analysis.*

6.2.2 Relative abundance SMD multivariate meta-analysis using Bernoulli formula.

We categorized the relative abundance data by taxa and used the variables n for sample size, x for the bacterial phyla's percentages, se for standard error and sd for standard deviation, as seen in Table 14.

A/A	PMID	author	year	taxa	x0	n0	se0	sd0	x1	n1	se1	sd1	x2	n2	se2	sd2
19	38024237	Zhang Z	2023	Firmicutes	46.3	4	0.249	0.499	79.5	4	0.202	0.404	52.5	5	0.223	0.499
50	37659809	Li Z	2023	Firmicutes	27.77	6	0.183	0.448	63.67	6	0.196	0.481	39.21	5	0.218	0.488
50	37659809	Li Z	2023	Firmicutes	27.77	6	0.183	0.448	63.67	6	0.196	0.481	35.2	6	0.195	0.478
234	35527507	Huang W	2022	Firmicutes	71.65	5	0.202	0.451	58.52	5	0.220	0.493	66.61	5	0.211	0.472
591	30589960	Cai Y	2019	Firmicutes	57.5	6	0.202	0.494	32.7	6	0.192	0.469	51.22	6	0.204	0.500
19	38024237	Zhang Z	2023	Bacteroidota	34.4	4	0.238	0.475	10.8	4	0.155	0.310	43	5	0.221	0.495
50	37659809	Li Z	2023	Bacteroidota	60.74	6	0.199	0.488	24.08	6	0.175	0.428	42.95	5	0.221	0.495
50	37659809	Li Z	2023	Bacteroidota	60.74	6	0.199	0.488	24.08	6	0.175	0.428	49.78	6	0.204	0.500
234	35527507	Huang W	2022	Bacteroidota	22.84	5	0.188	0.420	8.35	5	0.124	0.277	9.83	5	0.133	0.298
591	30589960	Cai Y	2019	Bacteroidota	36.46	6	0.196	0.481	58.09	6	0.201	0.493	41.87	6	0.201	0.493
50	37659809	Li Z	2023	Proteobacteria	1.33	6	0.047	0.115	2.17	6	0.059	0.146	8.21	5	0.123	0.274
50	37659809	Li Z	2023	Proteobacteria	1.33	6	0.047	0.115	2.17	6	0.059	0.146	5.83	6	0.096	0.234
591	30589960	Cai Y	2019	Proteobacteria	4.5	6	0.085	8.228	7.43	6	0.107	0.262	4.51	6	0.085	0.208

Table 14 – *Relative abundance data for multivariate meta-analysis*