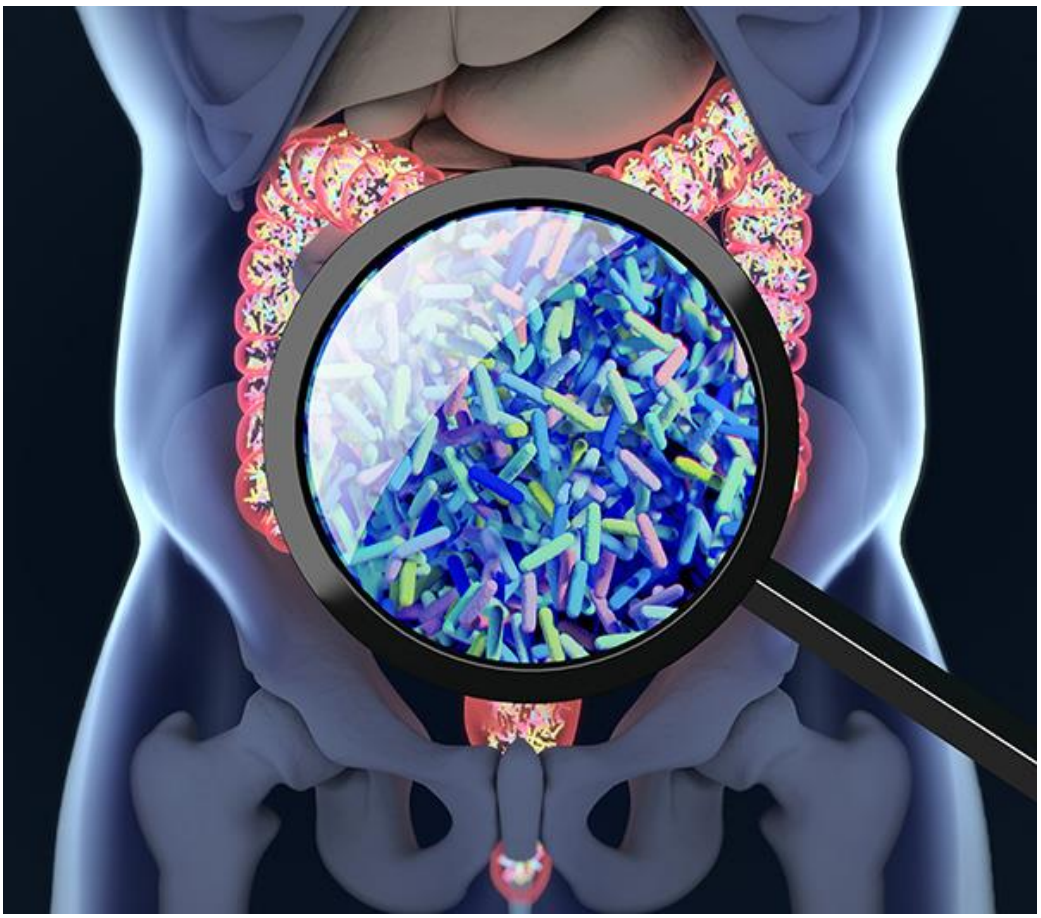
Ulcerative colitis and its association with the microbiome across age groups

MARIE VIBEKE VESTERGAARD

06-06-2021



Source: ChrisChrisW/Getty Images

AARHUS
UNIVERSITY

Acknowledgements

I would first like to thank my external supervisor Ph.D. and CEO Louise Bruun Thingholm for providing the settings for this interesting project. She has arranged the cooperation with MD, Ph.D. fellow Mikkel Malham Knudsen at Hvidovre hospital and provided the everyday setting of the master thesis. This includes a physical workplace and an environment of scientific sparring with her and the two other master students Ditte Torlyn and Anne Mathiesen. This has been very inspiring, motivating and rewarding to me. Louise Thingholm has been an excellent supervisor with many inputs and ideas and has always been very dedicated and helpful. I could not have asked for a better guidance.

I would like to thank my two amazing internal supervisors associate prof. Thomas Bataillon and Palle Villesen. I feel very lucky to have two so committed and passionate supervisors to guide me. They have throughout the project been very helpful. With a lot of patience, they have been open for discussing many subjects with me and provided a lot of helpful inputs and feedback.

Mikkel Knudsen from Hvidovre hospital has provided the dataset for the project and the biological questions to be answered. I would like to thank Mikkel Knudsen for trusting me with his project and for the great discussions we have had about the biological relevance of the project.

Lastly, I would like to thank prof. Dr. Astrid Dempfle from the university of Kiel for the great inputs on statistical modelling of longitudinal data.

Abstract

Ulcerative colitis (UC) is a chronic inflammatory disease that involves the rectum and colon and is one of the two main disease types of inflammatory bowel disease. Despite the identification of several genetic and environmental risk factors the cause remains unknown. The interplay between the immune system and the gut microbiome seems to play an important role, as both immune defects and microbiome dysbiosis are associated with the disease. Patients have a very diverse disease course, with fluctuations between relapse and remission, and paediatric patients often have a more severe disease course with higher need for surgery, but the cause for a more severe disease phenotype in children remains unknown.

For this study, stool samples were collected longitudinally from 53 patients and subjected to microbiome profiling using whole metagenome sequencing. Up to five samples were collected across 1 year from each patient. The dataset also includes information on disease scores, age, BMI, gender, and treatment up to sampling.

The aim of this project is twofold; I) to investigate microbiome features that might explain the difference in disease severity between paediatric and adult patients. This includes investigation of diversity measures, microbiome stability, abundance of single species and abundance of microbiome metabolic pathways and enzymes, and II) determine implied causality and detect microbiome species and functions that are important for the patients' relapse/remission status in the future. This is to find microbiome features that precedes a change in disease score and thereby infer a possible casual association.

The associations are investigated using mixed effects models, designed specifically for each aim. Both linear, logistic, and negative binomial mixed effects models are used in this project, where some also include zero-inflation. The overall microbiome composition is investigated using PERMANOVA.

Two different transformation methods are often used in microbiome studies, which are the relative abundance and the central log ratio transformation. In this project I have evaluated both methods and decided to go further with the relative abundance of various reasons.

I discovered that the alpha diversity measure richness is significantly reduced with a higher disease score, and that the microbiome composition changes with disease score. Further, the stability of the microbiome is decreased in paediatric compared to adult patients and in patients with a former diagnosis compared to newly diagnosed patients.

I detected 13 bacterial species, that associates differently with the disease score in children and adults. I additionally detected 24 species robustly associated with disease score in children and adults. In the functional analyses, I found the folate pathway to be significantly negatively associated with disease score. This result is interesting considering that UC patients have decreased levels of folate in serum.

In implied causality analysis, I detected one pathway and two enzymes that associated differently with the remission/relapse status three months later. These should be investigated further to evaluate their importance for the disease course and their potential protective or harmful effects.

More research is needed to determine the effect and the direction of the association of the significant features. They could in the future potentially be used to improve the disease course.

This study shows that features of the microbiome associate differently with disease score between adults and children with UC, and thereby highlights the importance of age in such studies and the potential role of the microbiome for the severe phenotype in children. It also shows that the choice of statistical method and disease score measure greatly influence the results and must be considered carefully.

Table of content

List of abbreviations

5-ASA	5-aminosalisylic acid
AR1	Autoregressive order 1
BH	Benjamini-Hochberg
BMI	Body Mass Index
CD	Crohn's disease
CDCV	Common Disease, Common Variant
CDRV	Common Disease, Rare Variant
clr	Central log ratio
EC	Enzyme Commission
F-cal	Faecal calprotectin
FDR	False discovery rate
GWAS	Genome-wide association study
IBD	Inflammatory bowel disease
Ig	Immunoglobulin
IL	Interleukin
IWLS	Iterative weighted least squares
KO	KEGG Orthologs
LMM	Linear mixed effects model
LogMM	Logistic mixed effects model
LPS	Lipopolysaccharide
NB	Negative binomial
NBZIMM	Negative binomial zero-inflated mixed effects model
NMDS	Non-metric multidimensional scaling
OR	Odds ratio
PERMANOVA	Permutational multivariate analysis of variance
PUCAI	Paediatric Ulcerative Colitis Activity Index
QC	Quality control
SCCAI	Simple Clinical Colitis Activity Index
SCFA	Short-chain fatty acid
SD	Standard deviation
SE	Standard error
UC	Ulcerative colitis
VIF	Variance inflation factor
WMS	Whole metagenome sequencing
ZIGMM	Zero-inflated gaussian mixed effects model