# LPS quantification of oral samples from the RHINESSA study population

**Kristina Utslottøy Mathiesen**

*This thesis is submitted in partial fulfillment of the requirements for the degree of Master in Biomedical Sciences*

Department of Biomedicine

BRuSH, Department of Clinical Science

University of Bergen

Spring 2023

# 1. Acknowledgments

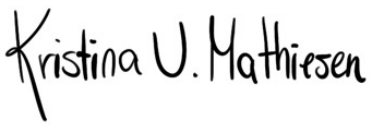
I would like to express my sincere gratitude to my supervisors, Maryia Khomich and Christine Drengenes, for their guidance and support throughout my thesis. Their expertise and willingness to share their knowledge have been invaluable in shaping my research project. I am grateful for their timely feedback and constructive criticism, which have helped me improve the quality of my work.

I also wish to extend my gratitude to the members of the BRuSH research group at the Department of Clinical Science (University of Bergen) for creating a supportive and stimulating research environment. I am particularly grateful to Randi Jacobsen Bertelsen (PI) for her exceptional leadership, guidance, and support throughout the entire research process. I also extend my thanks to Hesham Amin for his technical assistance and advice, and to my lab partner and fellow master student Tuva Norderud Jensen for her collaboration and support.

I would also like to acknowledge the research team behind the RHINESSA study for their important contribution to the understanding of environmental impacts on human health and disease.

I am deeply grateful to my family and friends for their unwavering love, support, and understanding during this time, especially towards the submission of my thesis. Their encouragement and motivation have been crucial in helping me overcome the obstacles of the research process. In particular, I would like to thank my husband for his steadfast belief in me and his support in managing time and stressful scenarios. I could not have accomplished this without you.

Finally, I would like to acknowledge the participants of this study who generously gave their time and effort. Their willingness to participate has made this work possible and I am grateful for their contribution.



Kristina Utslottøy Mathiesen

Bergen, June 2023

# Table of contents

# 2. Abbreviations

ASV – Amplicon Sequence Variant

CPI – Community Periodontal Index

ECRHS – European Community Respiratory Health Survey

EFW – Endotoxin Free Water

GCF – Gingival Crevicular Fluid

LAL – Limulus Amebocyte Lysate Assay

LPS – Lipopolysaccharide

MD2 – Myeloid Differentiation factor 2

MYD88 – Myeloid Differentiation primary response protein 88

OD – Optical Density

OM – Outer Membrane

PAMPs – Pathogen-Associated Molecular Patterns

PRRs – Pattern-Recognition Receptors

RHINE – Respiratory Health in Northern Europe

RHINESSA – Respiratory Health in Northern Europe, Spain, and Australia

TLRs – Toll-Like Receptors

TNF – Tumor Necrosis Factor

TRIF – TIR domain-containing adaptor inducing IFNbeta

# 3. Summary

The human immune system defends the body against threats through a complex network of organs, cells, and proteins. It has two subsystems: the innate immune system provides immediate defense, while the adaptive immune system responds specifically to encountered pathogens. Among these pathogens, bacteria hold particular significance. Bacteria are diverse single-celled organisms characterized by their relatively simple cellular structure. They can be broadly classified as Gram-negative or Gram-positive based on the composition of their cell walls. Gram-negative bacteria have a complex structure with an outer membrane and lipopolysaccharides (LPS). The structural configuration of the lipid A domain within LPS significantly influences the interactions with the host immune system, with certain lipid A variants demonstrating higher potency compared to others.

Given the substantial connection between the oral microbiome and the overall systemic health of individuals, one of the primary objectives of this master's project was to investigate the levels of LPS in saliva samples collected from adult participants from the community-based generation study Respiratory Health in Northern Europe, Spain, and Australia (RHINESSA), Bergen study center. Limulus Amoebocyte Lysate (LAL) assay is a commonly used technique to determine bulk LPS concentration in samples. However, the extent to which the LAL assay can accurately detect the various lipid A structures remains uncertain, and regrettably, the assay lacks the capability to discriminate between specific lipid A variants. We aimed to evaluate the suitability of the LAL assay for measuring LPS levels in samples obtained through two distinct collection methods (method 1: use of a commercial kit (Norgen); method 2: no kit) as part of a pilot project. Furthermore, we sought to explore the association between gingival and salivary LPS-producing microbiota assessed by 16S rRNA sequencing and salivary LPS activity measured by LAL assay.

For this master project, the samples were either collected on-site specifically for the pilot project or obtained from the larger RHINESSA study. The RHINESSA study is a longitudinal, multi-generational research initiative conducted in Europe and Australia, focusing on environmental and genetic factors associated with lung health.

Two different LAL assay kits were employed to measure LPS concentrations. The pilot project revealed that both LAL assays was unsuitable for samples collected using the commercial kit. Consequently, when conducting further investigations on the RHINESSA

samples, the samples collected using the kit were excluded from analysis due to resource constraints and interference.

DNA extraction was performed on the RHINESSA saliva samples, followed by sequencing and bacterial community profiling based on the 16S rRNA gene. The DNA and LPS concentrations of the samples exhibited non-Gaussian distributions, and Mann-Whitney tests unveiled no significant differences between genders or known factors related to oral diseases and LPS concentration. However, male participants exhibited higher LPS levels than females, which could potentially be attributed to factors such as tobacco use, Community Periodontal Index (CPI) score, and oral hygiene practices. Analysis of the bacterial taxonomy and lipid A annotation of gene sequences obtained from the saliva samples revealed a low ratio between bacteria producing hexa-acylated lipid A and those producing penta-acylated lipid A, possibly due to a healthy study population. Similar results were obtained from the analysis of previously collected gene sequences from gingival crevicular fluid (GCF) samples obtained from the same participants. The bacterial composition and abundance in both saliva and GCF exhibited similarities to previous studies in the field.

Results from correlation analysis demonstrated a positive correlation between the abundance of penta-acylated bacteria and the concentration of LPS in the samples, while a negative correlation was observed between the abundance of hexa-acylated bacteria and LPS concentration. This finding, in conjunction with the previously identified low hexa:penta ratio, suggests that the potential pro-inflammatory effects of hexa-acylated LPS-producing oral bacteria may be counteracted by the prevalence of more abundant penta-acylated LPS producers.

Overall, the analysis of bacterial microbiota utilizing techniques such as 16S rRNA sequencing and the LAL assay for measuring LPS concentration can provide valuable insights into microbial ecology and facilitate a deeper understanding of the complex interplay between environmental and genetic factors that influence human health.

# 4. Introduction

## 4.1 The immune system

The human body is equipped with a complex immune system that safeguards it against pathogenic invaders, toxins, and cellular anomalies that may threaten health. This multifaceted defense mechanism is comprised of various organs, including but not limited to the liver, spleen, and skin, in addition to diverse cells and proteins such as T cells, B cells, and cytokines [1]. Each of these immune cells plays a distinctive role and responds differentially to antigens, which are non-self-molecules that the immune system recognizes as potential threats. The binding of antigens to immune cells activates a complex cascade of reactions that is specific to the type of antigen and the immune cell involved. CD8+ T cells, for instance, are referred to as "killer T cells" due to their ability to directly target and destroy virus-infected and cancerous cells [2].

### *4.1.1 The innate and adaptive immune system*

The innate immune system, a crucial subsystem of the human immune system, operates in a non-specific manner. It is inherited, present from birth and does not require prior exposure to pathogens for activation. Through the use of immune cells like natural killer cells and phagocytes, the innate immune system provides an immediate and generalized defense against invading pathogens regardless of entry route.

In contrast, the adaptive (acquired) immune system has a highly specific response to pathogens encountered by the body. When the antigens of a pathogen bind to specific receptors present on T-cells and B-cells, these cells undergo activation and proliferation. The process of clonal expansion, which typically takes 3-5 days, leads to the generation of a sufficient number of cells to effectively eliminate the pathogen from the body. B-cells contribute to the immune response by producing and secreting antibodies that circulate in the bloodstream, binding to foreign antigens and rendering them inactive. On the other hand, T-cells possess the ability to directly kill cells that bind to their receptors or activate other cell types, thereby triggering a more robust immune response [3].

### *4.1.2 PAMPs and PRRs*

Pathogen-associated molecular patterns (PAMPs) are small microbial molecules that share several general patterns that alert immune cells to destroy them. Lipoteichoic acids in Gram-

positive bacteria, lipopolysaccharides (LPS) in Gram-negative bacteria, and double-stranded RNA (ds-RNA) produced by viruses are all examples of PAMPs. PAMPs play a key role in the innate immune system and are recognized by toll-like receptors (TLRs) and other pattern-recognition receptors (PRRs).

PRRs are prominently expressed on various effector cells within the innate immune system, including B-cells, dendritic cells, and respiratory cells. These receptors can be categorized into three distinct groups based on their functional roles: endocytic, signaling, and secreted. Endocytic PRRs primarily facilitate the internalization and engulfment of pathogens, residing on the surface of phagocytes. Signaling PRRs initiate specific pathways that lead to the synthesis of several inflammatory cytokines and antimicrobial peptides when activated. Lastly, secreted PRRs enhances the process of phagocytosis, serving as opsonin [3].

### 4.2 Bacterial cells

Bacteria are prokaryotic microorganisms characterized by their single-celled nature and inhabit a wide array of environments across the Earth. They exhibit remarkable adaptability, enabling their survival in even the most extreme conditions, making them the most abundant form of life on the planet. While the majority of bacteria contribute positively to human well-being and function as beneficial ecological agents, certain bacterial strains can pose threats and give rise to diseases [4].

Prokaryotic cells exhibit a simple structure characterized by their small size and lack of intracellular organelles or a nucleus to store genetic material. Bacteria, specifically, possess a single circular chromosome that contains all their genetic information, and this straightforward design enables them to rapidly grow and multiply. In terms of morphology, bacteria are highly diverse but can be broadly categorized into one of three main shapes: spherical (coccus), curved (vibrio), or rod-shaped (bacillus) [4].

#### *4.2.1 Gram-negative and Gram-positive bacteria*

Bacteria can be classified into two distinct groups based on their reaction to the Gram staining technique. This technique involves the fixation of bacteria in suspension onto a glass slide, followed by exposure to a blue or violet-colored dye known as hexamethyl pararosaniline chloride (also called crystal violet). The dye adheres to the peptidoglycan layer present in the cell wall. Subsequently, the slide is subjected to alcohol, which leads to the decolorization of Gram-negative bacteria while Gram-positive bacteria retain the dye.

The Gram-positive bacteria possess a thick peptidoglycan layer intertwined with teichoic acids and some proteins within their cell walls (Figure 1). In contrast, Gram-negative bacteria have a more intricate, multilayered cell wall composed of an outer membrane (OM) and a slim peptidoglycan layer. The OM is composed of phospholipids and lipopolysaccharides (LPS). The LPS hinders the attachment of the Gram stain to the cell's peptidoglycan layer, leading to the failure of the cell to retain the dye's color [5].

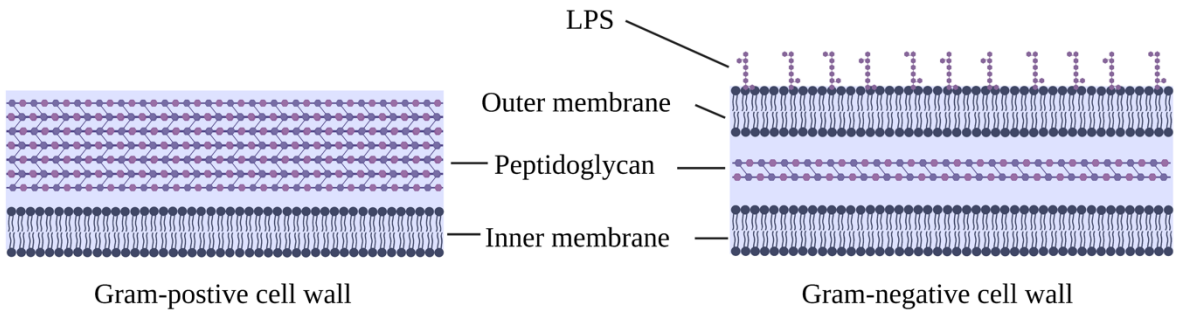



**Figure 1: The cell walls of Gram-positive and Gram-negative bacteria. The Gram-positive cell has a thick layer of peptidoglycan making it retain the color from the dye. The Gram-negative cell has an outer membrane containing lipopolysaccharides that hinders attachment of the dye. Created using BioRender.com, inspired by Pajerski et. al [6]**

### *4.2.2 Lipopolysaccharides (LPS)*

LPS, commonly referred to as endotoxins, are strong stimulators of innate immunity in eukaryotes and are essential for bacterial survival. As a key structural element of the OM, LPS serves as a robust defense barrier against small, hydrophobic molecules that can easily penetrate phospholipid bilayers. This structural characteristic grants Gram-negative bacteria inherent resistance to numerous antimicrobial agents and holds significant importance in the interactions between the bacteria and their host organisms [7].

LPS consists of three structural domains: highly conserved lipid A, core oligosaccharide, and highly variable O-antigen (Figure 2). Lipid A forms the outer leaflet of the OM and is responsible for human immune system reactions [7, 8].

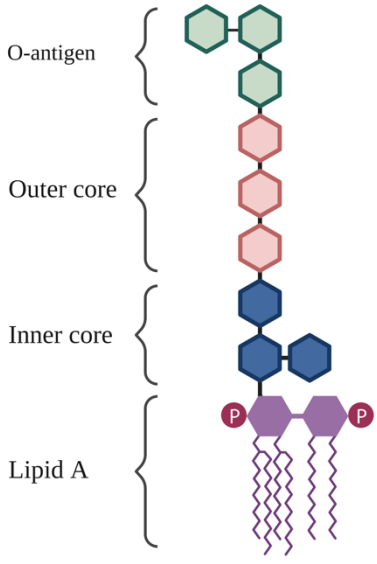



**Figure 2: Simplified lipopolysaccharide structure. Created using Biorender.com, inspired by figure 1 in Lin et.al [8]**

### 4.3 Lipid A

Lipid A in the cell's outer membrane is required for bacterial growth in almost all Gram-negative bacteria with *Neisseria meningitis* being an exception. The biosynthetic pathway for lipid A is highly conserved, and any mutations in the genes required for the initial steps in the pathway are lethal. The characterization of the lipid A biosynthetic pathway was discovered in the 1990s, challenging the previous notion that lipid A was a static structure. The characterization established that lipid A could indeed undergo post-synthesis modifications, allowing Gram-negative bacteria to adapt to unpredictable and hostile environments [1].

Many Gram-negative bacteria allow for modification of LPS during and after trafficking to the cell surface. Such modifications often involve the addition or removal of acyl chains and phosphate groups in the lipid A domain which directly affect the interactions with the host [1]. An example of such bacterial adaptability is the Gram-negative bacterium *Heliobacter pylori,* which causes stomach ulcers in humans. *H. pylori* expresses a unique type of lipid A which is significantly less stimulatory to the immune system than other lipid A structures, enabling the bacterium to persist for prolonged periods within the human body [9].

#### *4.3.1 Lipid A activation of the immune system*

LPS in the Gram-negative bacteria's OM is recognized by the PRR system Toll-like receptor 4-Myeloid Differentiation factor 2 (TLR4-MD2) complex on innate immune cells. The TLR4-MD2 receptor is synthesized in the endoplasmic reticulum and is localized in the plasma membrane or intracellular compartments such as the endosome. Recognition and binding of LPS occur by the LPS-binding protein (LPB) transferring the LPS to the CD14 co-receptor which is found either in soluble form in serum or bound to the cell surface (Figure 3). CD14 presents LPS to the TLR4-MD2 complex which dimerizes upon binding to LPS, triggering signaling through one of the two pathways: TIR domain-containing adaptor inducing IFNbeta (TRIF), or myeloid differentiation primary response protein 88 (MYD88) [9]. In some cases, such as for some *Yersinia* species [10] and *Aeromonas hydrophila* [11], LPS binding to the TLR4-MD2 receptor can activate both signaling pathways.

Between the two signaling pathways, MYD88 causes the most severe reactions in the host. Dimerization of TLR4-MD2 recruits MYD88 and MAL (TIRAP) adaptor proteins which initiates a signaling cascade that activates the production of pro-inflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNFα). The adaptor proteins additionally activate transcription factor NF-κB, which has a key role in the activation of the immune system [9]. NF-κB is rapid-acting and does not require protein synthesis for activation, leading to rapid changes in gene expression, vasodilation and recruitment of immune cells [12].

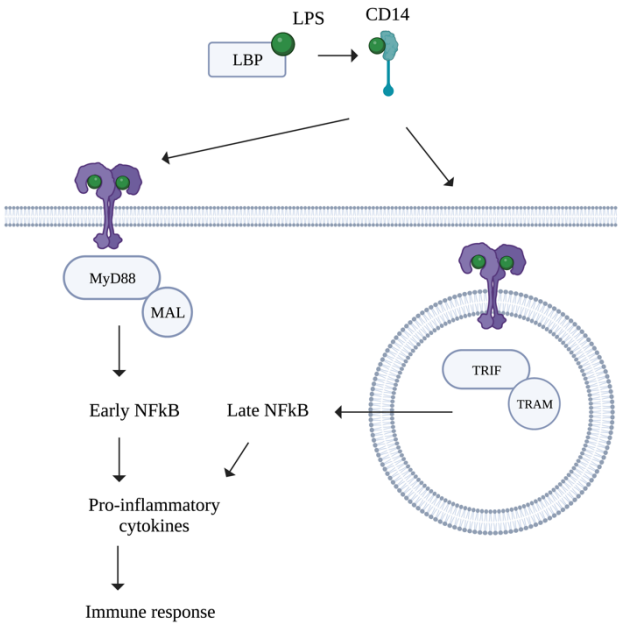



**Figure 3: Lipid A activation of the immune system. LBP transfer LPS to CD14, which present it to the TLR4-MD2 complex either in the cell wall or in the endosome. The complex dimerizes and triggers signaling which ends in activation of the immune system. Created using BioRender.com, inspired by Maeshima et.al [9].**

On the other hand, the TRIF pathway is slower and less inflammatory than MYD88 and is characterized by production of interferon-ß (IFNß). The pathway occurs in endosomes where the TLR4-MD2 receptor is expressed. Upon LPS binding and dimerization of TLR4-MD8, the TIR domain of TLR4 uses TRAM to recruit TRIF to the complex. The signaling complex activates several transcription factors including late NF-κB leading to maturation of myeloid dendritic cells, and production of cytokines and interferons. Understanding the mechanisms behind LPS recognition and signaling pathways can support the development of therapies to combat bacterial infections [9].

### *4.3.2 Variations in Lipid A acyl chains*

Most of what is known about the lipid A basic structure and synthesis has been studied in *Escherichia coli*. In *E. coli,* lipid A is composed of a di-glucosamine backbone, six fatty acyl chains, and 1- and 4′-phosphate groups, and is hexa-acylated (as depicted in Figure 4). Four of the six acyl chains are directly attached to the glucosamine head group, while the remaining two are linked as secondary chains to the hydroxyl groups of the 2′- and 3′-linked acyl chains. Due to its hexa-acylated structure, *E. coli* lipid A acts as a potent antagonist to human immune

system receptors. This hexa-acylated lipid A is recognized by the TLR4-MD2 complex, leading to the activation of the MYD88 pathway and subsequent production of NF-κB, thereby causing robust immune reactions [9, 13]. Among the various strains of *E. coli*, variations can be observed in the components of LPS, such as the O-antigen and the core oligosaccharide [14]. These structural differences influence the activation of downstream pathways following the binding of LPS to the TLR4-MD2 complex. It is important to note that these variations in LPS structure extend to other bacterial species as well, giving rise to diverse immune responses. The specific immune response elicited by LPS depends on multiple factors, including the host immune system and environmental conditions [9].

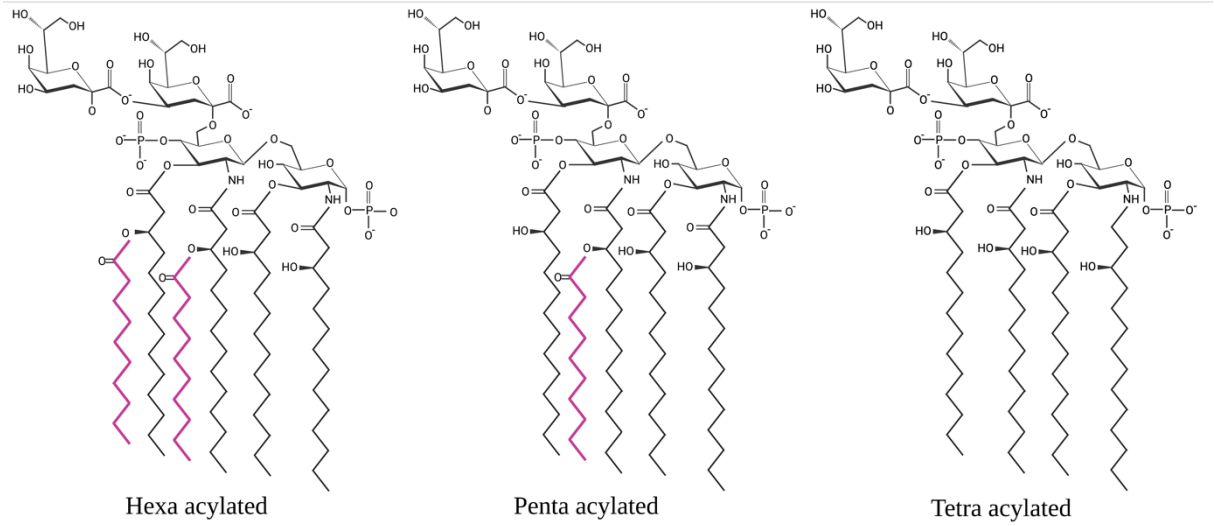



**Figure 4: The different structures of Lipid A. Hexa-acylated lipid A has six acyl chains attached to the backbone, while penta-acylated has five and tetra-acylated has four. Created using BioRender.com, inspired by Raetz et al [13].**

The penta-acylated lipid A structure, found in pathogens such as *P. aeruginosa*, exhibits structural similarity to hexa-acylated lipid A, but with only five acyl chains attached to the glucosamine group (Figure 4). Studies have demonstrated that this lipid A structure exhibits less binding affinity to the TLR4-MD2 receptor complex, resulting in reduced potency to the immune system. Consequently, pathogens bearing penta-acylated lipid A can evade host pathogen detection systems, leading to increased proliferation time. Nevertheless, penta-acylated lipid A is still recognized by the TLR4-MD2 receptor complex, and triggers signaling via the TRIF pathway, leading to inflammation in the host [9].

In contrast, tetra-acylated lipid A carries only four acyl chains (Figure 4) and it is speculated if it is unrecognized by the TLR4-MD2 complex, leading to decreased production of TNF-α [9]. Studies on *Porphyromonas gingivalis* have revealed lipid A heterogeneity comprising both

penta- and tetra-acylated lipid A, a finding that has complicated the understanding of the role of lipid A in *P. gingivalis* pathogenesis. Interestingly, while binding of tetra-acylated lipid A expressed by *P. gingivalis* has been shown to upregulate pro-inflammatory genes, binding of penta-acylated lipid A from *P. gingivalis* leads to down-regulation of the same genes, highlighting the significance of lipid A heterogeneity in *P. gingivalis*-mediated inflammation [15].

The Raetz pathway, which is responsible for the biosynthesis of lipid A, involves a series of enzymatic reactions wherein nine enzymes add various moieties to the lipid A precursor, culminating in the addition of fatty acid chains and phosphate groups to the diglucosamide backbone. Of particular note are the genes encoding the final two enzymes in the pathway, *LpxL* (Enzyme Commission (EC) 23.1.241) and *LpxM* (EC 23.1.243). In most Gram-negative bacteria, only the *LpxL* gene is present, resulting in the production of penta-acylated lipid A in the last step of the pathway. Conversely, bacteria harboring the *LpxM* gene can synthesize hexa-acylated lipid A, which confers increased potency compared to their penta-acylated counterparts [13]. *In silico* prediction methods based on the presence or absence of *LpxL* and *LpxM* from the whole genome sequenced bacteria makes it possible to predict the bacteria's potential for synthesizing penta- or hexa- acylated lipid A variants [16].

### 4.4 The human microbiota

While certain microorganisms have the ability to cause diseases and elicit immune responses in humans, it is important to recognize that a significant number of microorganisms can be beneficial to human health. The human microbiota [17], an intricate and diverse community of microorganisms, encompasses a wide range of organisms including bacteria, viruses, eukaryotes, and archaea. These microorganisms collectively contribute to vital functions that help maintain human health and well-being [18].

The human microbiota can be categorized into two distinct components: the core microbiota and the variable microbiota. The variable microbiota refers to the microbial communities that differ between individuals and are influenced by a multitude of factors, including sex, ethnicity, lifestyle, diet, and physiological variances. These factors contribute to the uniqueness of an individual's microbiota composition. In contrast, the core microbiota encompasses microorganisms that are consistently present across all humans, indicating their evolutionary conservation throughout time [18].

The human microbiome, characterized as the entirety of genetic material derived from microorganisms residing within the human body, far surpasses the size of the human genome. It has been estimated that the microorganisms in the human body make up to 100 trillion cells, more than ten times the number of human cells. Qin et al. (2010) characterized 3.3 million non-redundant microbial genes from fecal samples, making the microbial gene set in our gut 150 times larger than the human gene complement. Bacterial genes make up 99% of the gene catalog, and the diversity in the gut microbiota between the participants varies [19]. In fact, studies have shown that individuals can differ by up to 80-90% in terms of their gut microbiome or skin microbiome composition, while they are 99.9% identical to one another in terms of their human genome [20].

Studying and defining the core microbiome is of great importance, as it can provide insights into the fundamental elements that contribute to a healthy microbiome and aid in identifying potential therapies for microbiome-related diseases. Moreover, the identification of factors that disrupt the balance of the human microbiome, known as dysbiosis, can help to explain the mechanisms underlying the pathogenesis of numerous diseases, including inflammatory bowel disease, diabetes, and obesity [18].

Analyzing and understanding the human microbiome is a complex process, and numerous initiatives have been launched to investigate the roles of different microorganisms in the human body and their impact on human health. Among the more prominent endeavors are the Human Microbiome Project (HMP) [21] and the NIH Common Fund Human Microbiome Project [22]. These projects aim to advance our understanding of the human microbiome by exploring its composition, diversity, and function in various body sites and under different health conditions.

## 4.5 Oral microbiota

### *4.5.1 The oral microbiome*

The oral microbiome, regarded as the second largest microbial community in the human body, is defined as “the collective genetic material of microorganisms inhabiting the oral cavity” [23]. This intricate ecosystem comprises a wide array of microorganisms, including over 700 bacterial species, as well as viruses, fungi, and protozoa. Within the oral cavity, there are two types of surfaces that are susceptible to bacterial colonization: the hard tissues, such as the teeth, and the soft tissues such as the oral mucosa. The combined structures of the teeth,

tongue, cheeks, gingival sulcus, tonsils, and palates collectively create an environment where microorganisms can flourish.

Most bacteria find the oral environment favorable, with a mean temperature of 37°C and saliva pH of approximately 7, with regular water and nutrient supply. The oral microbiota is intimately connected with the host in the form of a biofilm, where the bacteria stick to the surfaces of the mouth. The oral microbiota protects the oral cavity, maintains homeostasis, and promotes health. In turn, dysbiosis may lead to progression towards disease. In general, the microorganisms in the human body play a critical role in the general health of a person by food digestion, energy generation, metabolic regulation, and maintenance of the immune system. Gaining insight into the impact of the oral microbiome on our health and how the bacterial composition contributes to oral and systemic health can improve our understanding of human body functioning and disease treatment [23].

### *4.5.2 Development*

The oral cavity of a newborn is usually sterile up until the first feeding where the mouth is inoculated to microorganisms from the mothers' skin and breast milk. Pioneer species that first colonize the mouth after birth include *Streptococcus*, *Lactobacillus*, *Actinomyces* and *Neisseria*. Following the eruption of teeth, more surfaces for bacterial colonization become available, and colonization of periodontal microbes begins in the gingival sulcus. Species diversity develops over time as new bacteria are introduced and the oral cavity changes with the person's age [23].

### *4.5.3 Composition*

The mouth is constantly exposed to various environmental factors that present challenges, such as food, chewing forces, saliva production and introduction of foreign microbes. The oral core microbiome is shared among most humans and consists of predominant species while the variable oral microbiome changes in response to lifestyle and oral hygiene. Normal, healthy oral microbiome has strong diversity and consists of both Gram-positive and Gram-negative bacteria, along with protozoa, fungi, and viruses. Bacteria are by far the most dominant group of microorganisms in the oral cavity and some of the most common bacterial genera are:

Gram-positive:

1. Cocci – *Streptococcus*, *Abiotrophia*
2. Rods – *Actinomyces*, *Corynebacterium*, *Lactobacillus*, *Propionibacterium*

Gram-negative:

1. Cocci – *Neisseria, Moraxella*
2. Rods – *Campylobacter*, *Fusobacterium*, *Haemophilus*, *Treponema*, *Desulfobacter* [23]

Variations in bacterial composition can be observed in samples collected from different surfaces and niches within the oral cavity of the same person. This can be attributed to the differences in local environments. For instance, gingival crevicular fluid (GCF) is more localized to the site, while saliva can provide a more representative view of the total oral microbiome [24].

#### *4.5.4 Oral and systemic health*

The oral environment, characterized by its dynamic nature, provides an environment where imbalances within the microbial community can occur, contributing to the development of dental and periodontal diseases. According to the World Health Organization (WHO), oral diseases affect a staggering 3.5 billion individuals worldwide, making them the most prevalent health conditions globally, even in high-income countries. Among the most common oral diseases are untreated caries, gingivitis, periodontitis, and edentulism (total tooth loss). Alongside genetic predispositions, factors such as inadequate oral hygiene practices, tobacco and alcohol use, and the consumption of sugar-sweetened foods and beverages are recognized as risk factors for the development of oral diseases [25].

In addition to painful symptoms in the mouth, oral disease has been linked to systematic diseases like diabetes mellitus, chronic obstructive lung disease, cardiovascular disease and also certain cancers [25]. Bacteria from the oral cavity can enter the body through ingestion, aspiration, and systemic dissemination. Pathogens that colonize the surfaces in the mouth can for instance lead to lung infections when inhaled [26]. Furthermore, despite environmental differences in temperature, pH, and oxygen levels, it has been reported that over half of the microbial species colonizing the gut lining are also present in the oral cavity [27].

### 4.6 Analyzing bacterial microbiomes and LPS

It has become increasingly evident in the past decades that the microbiome plays a critical role in the maintenance of human health and the development of various disease states. Analyzing bacterial microbiomes involves several steps, including sample collection, DNA

extraction, sequencing, and subsequent data analysis. The bacterial microbiome may be analyzed by targeted sequencing of a marker gene (e.g., the bacterial 16SrRNA gene) or sequencing of the total genetic material (e.g., metagenomics). These methods allow for the evaluation of bacterial diversity and abundance within microbiomes, as well as the functional potential of bacterial communities. The outcomes of these analytical approaches can provide valuable insights into microbial ecology and may have significant implications for human health and disease [28].

### *4.6.1 Targeted sequencing (amplicon sequencing)*

Targeted sequencing, specifically the polymerase chain reaction (PCR) amplification and sequencing of the 16S ribosomal gene, is a widely used approach for studying bacterial communities. The 16S ribosomal gene, which is approximately 1500 base pairs in length, is highly conserved and exists in almost all prokaryotic cells. As the gene contains variable regions (V1-V9), it can serve as a barcode for specific bacterial taxonomies, enabling the determination of bacterial genus and sometimes even species in a sample. The resulting data from 16S sequencing provides crucial information regarding the composition of bacterial communities, including the type of bacteria present in the sample and the relative abundance of each member [28]. Bioinformatic processing of the sequencing output is commonly performed using packages such as QIIME (https://qiime2.org/) [29] and Mothur (https://mothur.org/) [30]. These packages serve as wrappers for different bioinformatic tools (e.g., DADA2) used for processing the raw 16S amplicon sequencing output [28].

DNA extraction is the initial and most crucial step in targeted sequencing, as it introduces the most significant bias. Complete cell lysis is imperative to extract DNA successfully, and commonly, the peptidoglycan layer of the cell wall is targeted for lysis. As described in section 4.2.1 and depicted in the figure – thickness of the peptidoglycan layer varies between Gram-negative and Gram-positive bacteria, and lysis methods must be tailored to the bacterial communities found in the sample. Mechanical, enzymatic, and chemical lysis are the most common lysis methods employed, with careful consideration given to not impact DNA integrity during extraction. Though the ideal lysis method should be customized for each sample, this is rarely feasible [31].

Library preparation is also a critical step in amplicon-based marker gene sequencing and involves adding index sequences for sample multiplexing and amplifying the target marker gene for sequencing. The 16S ribosomal gene is a suitable marker gene as it is present in all

bacteria but absent in eukaryotic cells, enabling the capture of the full bacterial community without amplifying human DNA. Each amplicon is assigned a label sequence that links back to the corresponding sample, allowing for comparison of bacterial communities between samples. The index sequence can be added during the PCR amplification or after amplification of the target sequence [28]. The PCR step involves adding primers targeting conserved regions of the 16S rRNA gene to each sample, followed by three repeated steps of denaturation, primer annealing, and extension to amplify the target DNA exponentially. The resulting library of indexed amplicons is then sequenced using high-throughput sequencing technologies to generate data on the composition and abundance of bacterial communities [28].

Targeted sequencing has proven to be a valuable method for studying bacterial communities, providing insights into microbial ecology and its relevance to human health and disease. However, proper DNA extraction, optimization of the lysis method, selection of appropriate primers and indexes, and minimizing bias are crucial for obtaining accurate results.

### *4.6.2 Limulus Amebocyte Lysate assay*

The Limulus Amebocyte Lysate (LAL) assay, also known as the Limulus test or the endotoxin test, is a well-known *in vitro* assay used to detect bacterial endotoxins in various samples. It is extensively utilized in the pharmaceutical industry to ensure the safety of drugs and medical devices, as well as in environmental testing and water quality monitoring. The methodology was initially discovered by Frederick Bang during his studies of an infectious disease in the 1950s. He observed that Gram-negative bacteria caused clotting of the American Horseshoe Crab (*Limulus polyphemus*). Further research revealed that the endotoxin in the bacterial outer membrane activated the *Limulus* clotting factor C in the amebocytes of the crab's major circulating blood cells. Bang and Jack Levin then developed the LAL gel-clot assay by extracting the clotting factors through lysis. When the lysate came in contact with Gram-negative bacteria, it reacted with endotoxins, resulting in the formation of a solid gel clot that was visible to the researcher. The assay was highly specific for endotoxins in various samples, despite some challenges. It was time-consuming, highly skill-dependent, and difficult to replicate manually. To overcome these issues, a faster and more sensitive quantitative chromogenic LAL assay was developed in the 1980s. In this method, a chromogenic reagent is added to the LAL reagent to create a synthetic chromogenic substrate that changes color upon binding of endotoxin in the sample. The color development can be measured using spectrophotometry at 405 nm [32].

The most used LAL assay today utilizes a chromogenic reagent consisting of a peptide conjugated to p-nitroaniline, which enables spectrophotometric measurement (Figure 5). When running the assay, Factor C within the LAL reagent is responsible for activating Factor B upon encountering LPS. Subsequently, Factor B cleaves a pro-clotting enzyme, leading to its conversion into an active form. The active clotting enzyme then catalyzes the hydrolysis of the peptide-p-nitroaniline bonds, resulting in the release of a yellow colorant. The intensity of the yellow color directly correlates with the amount of Factor C in the sample, enabling precise quantification of the endotoxin content [33].

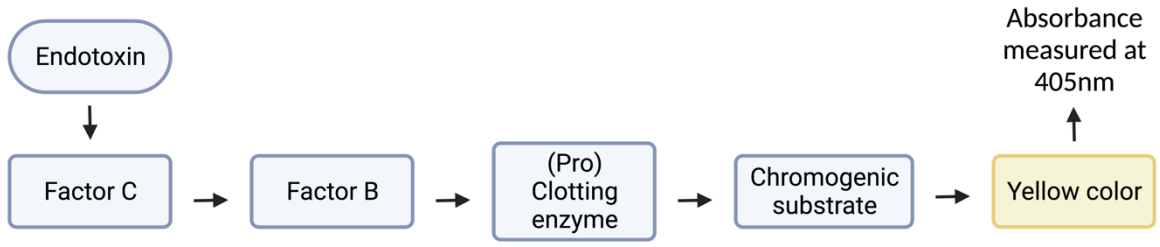



**Figure 5: Schematic simplification of the enzymatic cascade of the chromogenic LAL assay, based on information from Piehler et al (2020) [33]. Endotoxin comes in contact with factor C leading to activation of factor C and subsequently activation of Factor B Factor B in turn activates the clotting enzyme which cleaves the bonds between peptide and the coloring agent resulting in color formation which can be measured by spectrophotometry. Created using BioRender.com.**

The chromogenic LAL assay is highly sensitive (0.005 EU/ml as the lowest measurable LPS concentration) and specific for the detection of endotoxins. The colorimetric readout allows for quantitative measurements, which is important for ensuring the safety of pharmaceutical products and medical devices. Additionally, this method is faster and more convenient than the traditional gel-clot LAL assay, as it does not require manual interpretation of results. However, it should be noted that the chromogenic LAL assay may still have limitations and potential interferences, and careful validation is necessary for its use in specific applications [33]. An example of this is that LAL assay measures bulk endotoxin levels and does not distinguish between the different variations of lipid A. This can be an issue as the different lipid A structures varies enough to react differently with the reagent of the LAL assay. In addition, some strains can be below the detection limit of the assay even though they could impact inflammatory effects once they come into circulation in the body [34].

The diagnostic application of LAL assay has not yet received approval due to methodological limitations and inconclusive findings regarding the clinical significance of

endotoxemia in blood. However, the assay continues to hold value in research settings for the exclusion of endotoxin in samples. Additionally, it serves as a valuable tool in the development of medical devices and drugs [35].

### 4.7 The RHINESSA study

Respiratory Health in Northern Europe, Spain, and Australia (RHINESSA) [36] is a multi-generational, multi-center research cohort that aims to understand the impact of lifestyle and environmental factors on the health of the subjects from a general population. The project is international in scope, with study centers in seven countries: Denmark, Norway, Sweden, Spain, Australia, Iceland, and Estonia. The study builds on two large-scale research projects, RHINE (Respiratory Health in Northern Europe, https://rhine.w.uib.no/) and ECRHS (European Community Respiratory Health Survey, https://www.ecrhs.org/), which were established in the early 1990s to investigate respiratory health, allergies, eczema, and related diseases in adults.

The ECRHS study [37] was initiated in 1991, by inviting random individuals from well-defined populations in Europe, born between 1945 and 1973, to participate in a questionnaire study. Subgroups of participants from 30 study centers were then invited to participate in clinical studies every 10 years. The original ECRHS study involved over 23000 people who completed the questionnaires and over 6700 of them underwent clinical testing.

The RHINE study is questionnaire-based only and followed-up all the participants from the original ECRHS study population in seven study centers. From 2002 to 2004, over 16000 people responded to extensive postal questionnaires, resulting in a response rate of 75%. In 2010-2012, approximately 13000 people participated in the third round of the study.

In 2012, the offspring of the original RHINE participants (ECRHS for Spain and Melbourne) were invited to participate in the RHINESSA study. The new generation of participants included about 10000 people who completed the questionnaire and 1600 of them underwent clinical testing. Regular follow-up of RHINESSA participants is planned to take place every 10 years, with new generations of offspring being included.

The questionnaire in the RHINESSA study includes questions on education, place of upbringing, tobacco use, diet, exercise, sleep, respiratory diseases, and allergies. Women's questionnaires also include questions on pregnancies, menstrual cycle, and gynecological diseases. Clinical tests measure anthropometry, lung function, blood pressure, and heart rate,

among other things. Blood, skin swabs, saliva, and urine are some of the biological samples that have been collected from the participants.

The RHINESSA study will follow multiple generations over time, with similar questions and clinical tests being administered at each follow-up. This approach has the potential to shed light on how lifestyle and environmental factors impact the health of both parents and their offspring and may help to identify susceptible time windows where intervention could be effective [36].

# 5. Aims

LAL Chromogenic Endpoint Assay was used for quantification of LPS in saliva samples collected from the RHINESSA adult study population, Bergen study center. Although the LAL assay is an effective method for detecting LPS, it does not discriminate between the different structural forms of lipid A, a part of the LPS molecule that triggers an inflammatory response in the host. Hence, it remains uncertain whether the LAL assay is equally sensitive to the various types of lipid A.

The primary objectives of this thesis are threefold:

1. To determine whether the two modes of collecting saliva samples ("kit" or "no-kit") affect LAL measurements.

2. To measure and investigate the LPS concentration in RHINESSA saliva samples (n=79).

3. To explore the LAL assay's ability to detect bacteria that produce potent hexa- and less inflammatory penta-lipid A variants collected from the saliva samples of the adult RHINESSA participants.

# 6. Material and methods

## 6.1 Saliva collection

### *6.1.1 Saliva collection in the RHINESSA study*

A total of 87 saliva samples from adult participants (mean age 27.2 years, 24.29% male), Bergen study center, were collected both at baseline in 2014-2015 and at follow-up in 2020-2021. The baseline samples were collected in a clinical setting under the supervision of a clinician or a researcher as part of a larger study. A list of guidelines was provided to the participants before collection to ensure the accuracy of the results. These guidelines included refraining from smoking, avoiding heavy meals for at least one hour before sample collection, and abstaining from asthma medication for 12 hours prior to sampling. Detailed instructions can be found on page 3 of the Standard Operating Procedure (SOP) [38].

In accordance with SOP (page 47), saliva samples were collected by having the subjects hold a sterile falcon tube and tilt their head down to allow saliva to drip into the container. To maintain the integrity of the samples, the inside and top of the tube were not allowed to be touched during collection. A minimum volume of 2 ml of saliva was required for the sample to be accepted. The laboratory procedure outlined in the SOP involved mixing the sample with an equal amount of PBS (at least 3 ml), transferring it to 2 ml tubes, and storing it at -80℃ [38].

#### *6.1.1.1 Follow-up year 2020-2021*

During the 2020-2021 follow up of the RHINESSA study participants in Bergen, the sampling procedure was affected by the ongoing Covid-19 pandemic, leading to some of the clinical samples being collected in the participant's home, rather than at the clinical facility. The samples collected at home included saliva samples, and they were collected using Norgen Saliva DNA Collection and Preservation Devices (Product #RU49000, Norgen Biotek Corporation, Thorold, ON, Canada).

Together with the saliva home collection kit, the participants were provided with the same clear instructions as for the baseline collection. After collecting 2 ml saliva the participants added the preservative that came with the kit and mixed the sample by shaking the container. The container was sent to the research facility by post. Along with the sample, the participants

sent a form in which they stated the date and hour of sampling and when and what they had eaten or been drinking prior to the saliva collection [38].

### *6.1.2 Saliva collection for the pilot project*

To compare the two methods of collection used for the baseline and follow-up sampling, saliva samples were collected from 32 healthy volunteers in September-October 2022. Saliva from 16 participants was collected using the baseline (no-kit) method first, while the remaining 16 saliva samples were collected using the follow-up (kit) method first. The no-kit samples were collected in sterile, non-pyrogenic 50 ml Falcon tubes. The first 23 samples were collected in FisherBrand 50 ml Centrifuge Tubes (Cat.: 05-539-13), while samples 24-32 were collected in Sarstedt Tube 50 ml (Ref.: 62.547.254) due to a shortage of the former. The kit samples were collected using the same type of kit as in the RHINESSA study (Norgen Saliva DNA Collection Product #RU49000, Norgen Biotek Corporation, Thorold, ON, Canada).

Apart from a shorter fasting period (30 minutes prior to sampling), the participants received the same instructions prior to sample collection as in the RHINESSA SOP [38]. The participants were closely observed throughout the sampling process and provided with guidance to ensure passive flow of saliva into the collection container, avoiding any forceful actions. They were instructed to maintain consistency between the two collection methods, refraining from applying additional force during the latter method.

To account for possible interference in the samples, the participants were all asked the same questions regarding food intake, drinking, tobacco use, and asthma medication use in the hours prior to sampling. In adherence to the SOP outlined in the RHINESSA study (p. 47), the laboratory procedure for processing the no-kit saliva samples involved adding 3 ml of phosphate-buffered saline (PBS, prepared in-house) within one hour of collection. The sample was then mixed by inversion 10 times and divided into 3 containers. These containers were sterile, DNA-, RNA-, and pyrogen-free 2.0 ml Eppendorf® Biopur® Safe-Lock microtubes (Merck, Cat# Z317217). The containers were frozen immediately and stored at -80 ℃. The kit samples were frozen directly in the collection tubes at -80 ℃ [38].

## 6.2 DNA extraction

DNA was extracted from 78 participants due to lack of material in one sample. The extraction was done using the FastDNA SPIN kit from MP Biomedicals (Cat.no: 6540-600) following protocol based on previous experiments at our lab [39]. Following the optimization

of the DNA extraction protocol for saliva samples collected using both the kit and no-kit methods, the protocol overview is presented as follows (Figure 6):

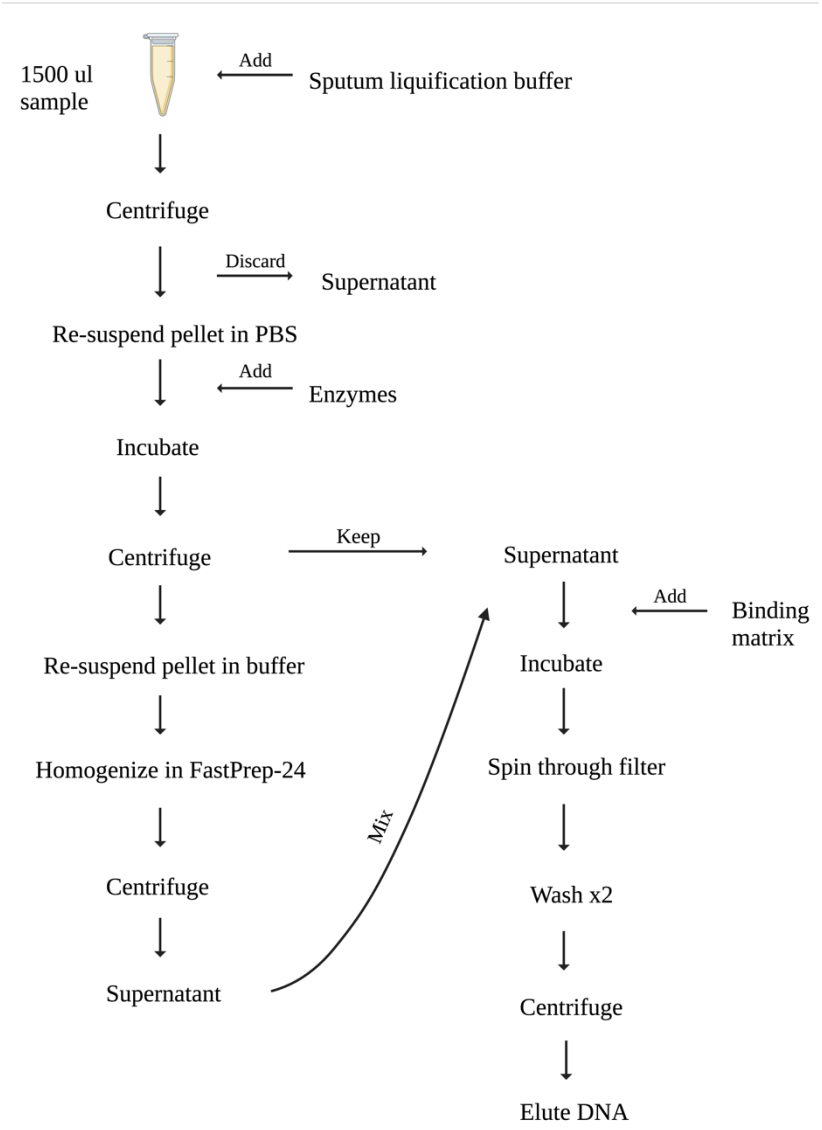



**Figure 6: Overview of the DNA extraction protocol based on previous work at our lab [39]. Created using BioRender.com.**

### *6.2.1 DNA extraction protocol*

The following protocol was originally made by Hoang et.al (2019) [39] and later optimized for our sample types and volume. The samples were thawed at room temperature (RT), vortexed for 10 s and spun down before dividing 1500 µl of the sample into two aliquots of 750 µl in sterile 2 ml tubes (Sarstedt, cat.no.: 72.695.200). A working solution of Sputum Liquification Buffer (Norgen, product no.: 28289) was made based on the number of samples being processed:

**Table 1: Amounts of Sputum Liquification Buffer to be added to DNA- and endotoxin-free water for making of working solution.**

| Number of samples | 3 | 7-8 | 9-12 |
|---|---|---|---|
| Sputum Liquification Buffer (ml) | 0.75 | 1.125 | 1.5 |
| DNA- and endotoxin-free H2O (ml) | 9.25 | 13.875 | 18.5 |
| Total (ml) | 10 | 15 | 20 |

750 µl of the Sputasol working solution was added to each of the two sample aliquots and vortexed for 5 s. The tubes were incubated on a thermoshaker at 37°C with shaking at 1000 rpm for 15 min followed by centrifugation at 15800 x g for 8 min. The supernatant was transferred to a separate tube and stored at -80°C. The three pellets from one sample were resuspended and combined in a total of 250 µl PBS. The sample was vortexed for 15 s and spun down briefly.

An enzyme cocktail solution was prepared according to the number of samples being processed (Table 2) and 50 µl was added to each sample. The entire volume was pipetted up and down 10 times. The samples were incubated on a thermoshaker at 37°C with shaking at 350 rpm for 1 h followed by centrifugation at 15800 x g for 15 min.

**Table 2: Amounts of enzymes to add to each sample.**

| Components | Volume/Sample (µl) |
|---|---|
| Lysozyme (10 mg/ml, Sigma-Aldrich, Cat.no.: L3790-10) | 5 |
| Mutanolysin (25 000 U/ml, Sigma-Aldrich, Cat.no.: M9901-50KU) | 3 |
| Lysostaphin (4 000 U/ml in NaAC, Sigma-Aldrich, Cat.no.: L4402-5MG) | 1.5 |
| TE5 buffer (10 mM Tris HCl, 5mM EDTA, pH 8.0, made in-house) | 40.5 |
| Total | 50 |

The supernatant of each sample was transferred to a separate 2 ml tube and stored at 4°C. The pellet was resuspended in 800 µl CLS-TC buffer (FastDNA SPIN kit) by pipetting up and down 10 times. The suspension was transferred to a Lysing Matrix A Tube (FastDNA SPIN kit) and homogenized using FastPrep-24 (MP Biomedical) at 6 m/s for 40 s. The samples were then centrifuged at 14000 x g for 10 min. 650 µl of the supernatant were transferred to the supernatant stored at 4°C. After combining the supernatants, the total volume was approximately 950 µl per sample. An equal amount (950 µl) Binding Matrix (FastDNA SPIN kit) was added to each sample followed by incubation on a rotator for 5 min at low speed. After incubation 700 µl of the suspension was transferred to a SPIN filter (FastDNA SPIN kit) and centrifuged at 14000 x g for 1 min. Flow through was discarded and the process was repeated until the whole sample had been spun through the filter. The pellet in the SPIN filter was washed by resuspending it in SEWS-M (FastDNA SPIN kit) and the sample was centrifuged at 14000 x g for 1 min. Flow through was discarded and the washing was done once more. The catch tube was replaced with a new catch tube and 100 µl DES (FastDNA SPIN kit) was used to resuspend pellet by pipetting up and down several times. Following, the sample was incubated at 55°C for 5 min. To elute the DNA, the sample was centrifuged at 14000 x g for 1 min. The DNA was stored at 80°C until further use.

All centrifugation was done at room temperature.

### 6.3 NanoDrop

For DNA purity assessment, a NanoDrop 1000 spectrophotometer (ThermoFisher Scientific) was employed, utilizing 2 µl of the sample. The 'Nucleic acid' application module was selected, enabling absorbance measurements at wavelengths of 230 nm, 260 nm, and 280 nm. Following the completion of each sample measurement, both the upper and lower pedestals were carefully wiped using a dry laboratory wipe. To ensure accuracy and prevent cross-contamination, 2 µl aliquots of distilled water were used for cleaning the pedestals before the first measurement of the day, as well as for additional cleaning when needed [40].

### 6.4 Qubit

Qubit 3.0 Fluorometer (Life technologies, ThermoFisher, cat.no.: Q33216)) along with Qubit dsDNA HS Assay kits (Life technologies, ThermoFisher, cat.no.: Q32851) and Qubit dsDNA BR Assay kits (Life technologies, ThermoFisher, cat.no.; Q32850) was used to measure DNA concentration. The fluorometer was calibrated prior to using each kit, either by standard 1 (0 ng/µl) and 2 (10 µl/ng) from Qubit dsDNA HS Assay kits or by standard 1 (0 ng/µl) and 2

(100 ng/µl) from Qubit dsDNA BR Assay. Thin-wall, clear 0.5 ml Qubit assay tubes (Life technologies, cat.no.: Q32856) were used to mix the sample with reagents. The final volume in all tubes was 200 µl, containing 2 µl sample and 198 µl working solution. All measurements were done at room temperature [41].

## 6.5 LAL assay

LAL assay was performed using two kits as listed below (Table 3: Materials used for LAL assay.. The main measurements were made using the Hycult LAL Chromogenic Endpoint assay while the Lonza LAL Kinetic-QCL Assay was used for troubleshooting in the pilot project when comparing kit/no-kit samples.

### *6.5.1 Material*

All materials used for the LAL assay were sterile, non-pyrogenic, DNA- and RNA-free. Absorbance was measured using Synergy H1 Hybrid Multi-Mode Reader (BioTek).

**Table 3: Materials used for LAL assay.**

| Material | Manufacturer | Category number | LOT |
|---|---|---|---|
| LAL Chromogenic Endpoint Assay | Hycult Biotech | HIT302 | 34079K0722 |
| LAL Kinetic-QCL Assay | Lonza | 50-650U | 0001035661 |
| 10 µl Sterile Finntip Flex Filter | ThermoFisher Scientific | 94056980 | 22027B2 |
| 200 µl Sterile Finntip Flex Filter | ThermoFisher Scientific | 94056380 | 20363B2, 21148A1 |
| 1000 µl Sterile Finntip Flex Filter | ThermoFisher Scientific | 94056710 | 1051521 |
| 2.0 ml SafeSeal reaction tube Biosphere® Plus | Sarstedt | 72.695.200 | 1082821 |
| 1.5 ml SafeSeal reaction tube Biosphere® Plus | Sarstedt | 72.706.200 | 2083221 |

### *6.5.2 Dilution optimization*

To efficiently use time and sample material, the optimal dilution for the samples was determined by utilizing both a subset of the pilot project samples in addition to saliva samples collected throughout the optimization stage. The optimization process was done by selecting random samples, performing serial dilutions, and running them through the Hycult kit protocol [42].

The serial dilution was done in the following way: 100 µl of the selected sample was transferred to a sterile, non-pyrogenic 2 ml tube (Sarstedt, cat.no.: 72.695.200) containing 900 µl distilled, endotoxin free water (EFW, Sigma Aldrich, cat.no.: 102535346). The mixture was then vigorously mixed for 15 s. To further dilute the sample, 100 µl of the mixture was transferred to a new sterile tube containing 900 µl of distilled water. This process was further repeated two times, leading to a final dilution of 1/10000 (Figure 7).

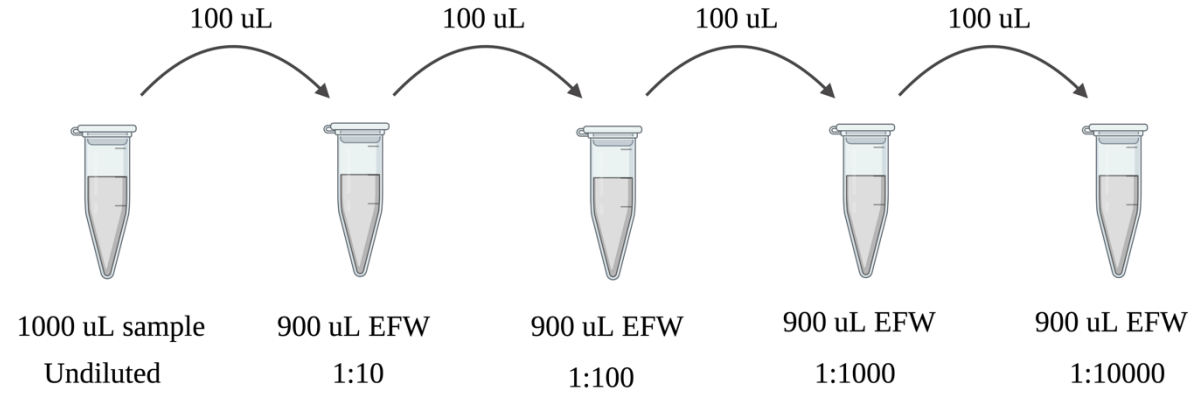



**Figure 7: Serial dilution of samples. 100 µl of undiluted sample was transferred to 900 µl EFW, diluting the sample 1:10. The sample was then further diluted up to 1:10000 following the same principle. Created using BioRender.com.**

To obtain consistent and accurate results, the dilution experiment was conducted multiple times using samples from various individuals. To address the challenges and find interference encountered during the process, several adjustments were made to the original protocol. These adjustments included refining the pipette technique, prolonging the mixing of the LAL reagent, and testing for interference in the PBS, H2O, and buffer from the home-kit. Spiking was performed in order to determine the source of interference in the samples and was done using the 50 EU/ml standard solution from the kit not being used for the LAL assay. This means that the 50 EU/ml standard from the Lonza kit was used to spike samples measured with the Hycult assay, and vice versa.

### *6.5.3 Hycult LAL Chromogenic Endpoint Assay protocol*

This protocol was made based on the LAL Chromogenic Endpoint Assay manual from the manufacturer of the LAL kit (Hycult) [42].

The samples were thawed at RT before starting the procedure. Reagents were incubated at RT and stored at 2-8°C after use, for as long as they were stable. The LAL reagent was prepared by adding 3.8 ml EFW to the vial. The top of the vial was removed after the addition of EFW, and the vial was covered with the inner side of a parafilm. To dissolve the lyophilized reagent, the vial was gently swirled until the solution was colorless. The vial was kept out of the light and stored at 2-8°C until use (maximum 3 h).

The standard solution was reconstituted by adding 1.2 ml EFW, incubated at RT for at least 5 min before vortexing for 30 s. After reconstitution the standard had a concentration of 50 EU/ml. The stop solution was prepared by mixing a specified amount of 2.5x concentrated stop solution with distilled water. The exact quantity of 2.5x concentrated stop solution to be added was determined by consulting the table provided (Table 4).

**Table 4: Amounts of concentrated stop solution to be added to distilled water.**

| **Plates** | 1 (96 wells) | 2 (192 wells) | 3 (288 wells) |
|---|---|---|---|
| **Concentrated stop solution (ml)** | 4 | 8 | 12 |
| **Distilled water (ml)** | 6 | 12 | 18 |

For standard series the reconstituted standard was diluted 2 times to a concentration of 25 EU/ml by adding 50 µl standard solution to 50 µl EFW and vortexed for 30 s. To make duplicate standard curves the first 16 wells in the plate were filled with 50 µl EFW. 33 µl of the diluted standard was transferred to well A1, and further diluted 1:1.5 by mixing thoroughly and pipetting 33 µl over to well B1 and so on until well G1 (see Figure 8). 33 µl was discarded from well G1 as H2 was used as control value. The same procedure was repeated for well A2-H2.

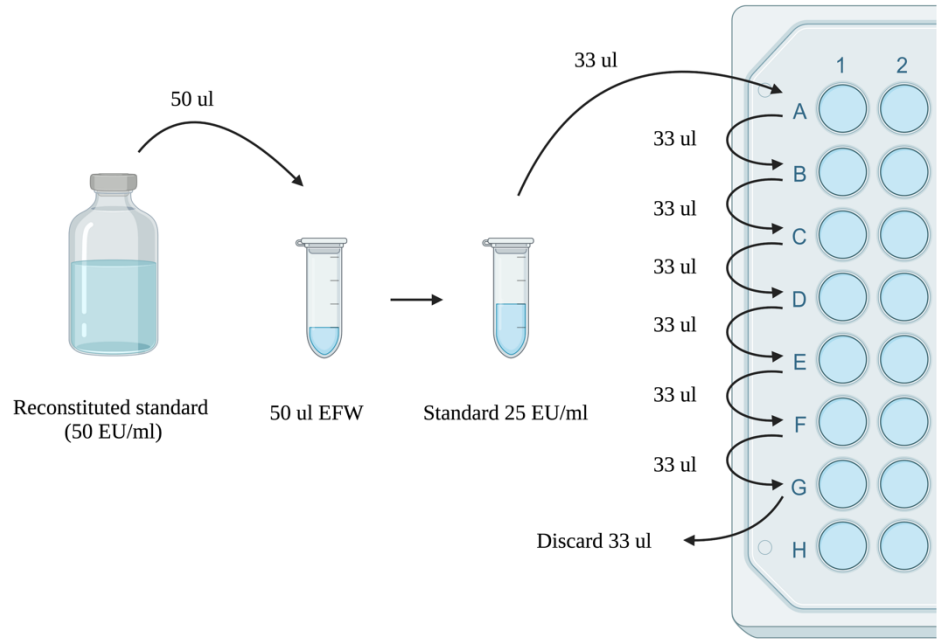



**Figure 8: Diluting the standard solution and making the standard curve. Done in duplicates. Wells were filled with 50 µl EFW before adding the diluted standard. Final concentration in wells A1-H1: 10 EU/ml; 4 EU/ml; 1.6. EU/ml; 0.64 EU/ml; 0.26 EU/ml; 0.10 EU/ml; 0.04 EU/ml; 0 EU/ml. Figure inspired by Hycult LAL assay HIT302 protocol p.8. Created using BioRender.com.**

Based on dilution optimalization, the no-kit samples were diluted 1:10000. 50 µl in duplicate from each diluted sample and control was transferred to the assigned wells using clean pipette tips for each transfer. EFW, H2O and PBS were used as negative controls and blanks, respectively, as they were added to samples or reagents during the experiment. The negative controls were in duplicates, with one of the duplicates being added LAL reagent to check for endotoxin interference and the other duplicate being used as a blank.

50 µl of the reconstituted LAL reagent were added to all wells except blanks which were added 50 µl EFW. The plate was incubated at RT for 20 min before measuring the absorbance at 405 nm using Synergy H1 Hybrid Multi-Mode Reader (BioTek). Additional incubation was done in 5-min intervals until the standard concentrations 10 and 4 EU/ml differed less than 10%. The reaction was stopped by adding 50 µl 1x stop solution keeping the same sequence and timing in mind as used when adding LAL reagent to the wells. The final measurement was done at 405 nm.

### *6.5.4 The Lonza LAL Kinetic-QCL Assay*

Incongruent results from both the dilution experiment and testing of samples in the pilot project led to utilizing another type of LAL assay to see if the results improved. The Lonza

LAL Kinetic-QCL Assay was used for this, following a protocol made by Aarhus University [43]. The protocol was followed step by step using a standard curve (0.012-25 EU/ml) to calculate endotoxin concentration. All reagents were brought to room temperature before use. The reagents were reconstituted with the specific amount of LAL Reagent Water (LRW, Kinetic-QCL Assay Kit) specified in the protocol. Standard series was made by a serial dilution, transferring 100 µl reconstituted standard between wells A1-E2. The wells were pre-filled with 100 µl LRW. 100 µl blanks, controls, and samples were dispensed into appropriate wells and pre-incubated at 37°C in the plate reader for 10 min. 100 µl reconstituted LAL reagent was added and the plate was read at 1-min intervals at 405 nm over a period of 41 min. Reconstituted standard (50 EU/ml) from the Hycult LAL kit served as corresponding positive controls: 5 EU/ml, 2.5 EU/ml, 1.25 EU/ml, and 0.75 EU/ml. EFW, PBS and home-kit buffer were spiked with positive controls to check for interference.

### 6.6 Lipid A annotation

Lipid A annotation was done *in silico* based on the amplicon sequenced data from both the GCF and saliva samples of the RHINESSA participants. Annotation of oral bacteria to hexa-, penta- or tetra-acylated LPS-producers was done according to Brix et al. [16]. It is based on the presence or absence of the genes in the Raetz pathway in the whole genome sequenced bacteria. Bacteria containing all nine genes, including *LpxM*, were assigned to have the ability to produce hexa-acylated LPS, while the bacteria with the first eight genes in the pathway, and not *LpxM*, were annotated to have the ability for penta-acylated LPS production. Bacteria able to produce LPS, but not carrying the *LpxL* or *LpxM* genes were annotated as tetra-acylated [16]. The annotation was done using bacterial genera when available, and family or order if genus was unknown.

### 6.7 Illumina MiSeq sequencing

The GCF samples collected at baseline were prepared and sequenced at the UNC Microbiome Core Facility at the University of North Carolina, USA. DNA was extracted from the GCF samples and based on all five paper points from the lower jaw. For compatibility with the Illumina MiSeq sequencing platform, the primers targeting the V1-V2 region of the 16S rRNA gene contained overhang adapters attached to the 5’ end of the primer. Amplicon sequence variants (ASVs) were assigned taxonomy against the Human Oral Microbiome Database v.15.1 (https://www.homd.org/) [44]. The bioinformatics settings are described in detail in Khomich et al. (2023) [45].

For the RHINESSA saliva samples, PCR and library preparation was done by another master student at our lab, following the in-house protocols [39]. The 16S amplicon sequencing (V3-V4 region) was done at The Mohn Cancer Research Laboratory (MCRL) at Haukeland University Hospital, Norway. The choice of the targeted region for MiSeq sequencing was based on the Illumina protocol described in Klindworth et. al [46]. Bioinformatic steps were performed as described for GCF samples, but with a distinction in the reference database used for taxonomic assignment. Specifically, the Greengenes database v.13.8 (https://greengenes.lbl.gov/Download/) [47] was utilized for this purpose.

### 6.8 Statistics and computer analysis

Standard curves and statistical analysis from the LAL assays were made using Gen5 software (BioTek, v.2.00.18). All samples were measured in duplicates and average values were used. Standard curves were made using non-linear 4-point regression as stated in the kit protocols. Data from each run were regarded as valid only if the OD was under 0.7, the Coefficient of Variation (CV) of all replicates was <15% and fit of the standard curve was $r < 0.985$.

GraphPad Prism 9 (https://www.graphpad.com/) was used for statistical tests and making graphs. Based on sample size the Shapiro-Wilks test (α: 0.05) was used to measure normality. P-value < 0.05: rejection of the null hypothesis of normal distribution. Mann-Whitney (α: 0.05) test of significant difference was used to compare groups as the data was non-Gaussian distributed. P-value < 0.05: Rejection of the null hypothesis of significant difference.

Descriptive statistics are presented as mean (± standard deviation (SD)) and median (interquartile range (IQR)) for continuous variables (compared by Shapiro-Wilk test) and as frequency (percentage) for categorical variables.

%Recovery was calculated using this formula: $\frac{Measured\ concentration}{Known\ concentration} * 100$

Correlation analysis to examine the relationship between LPS concentration and bacterial genera with different lipid A variants was done using centered log-ratio (CLR) transformed ASVs, which were initially aggregated by genus. Spearman correlation was selected as the appropriate model for this analysis, considering the non-Gaussian distribution of the data.

# 7. Results

## 7.1 Dilution project and optimization

This project was initiated to find the optimal dilution for the sample types and to evaluate the suitability for the LAL Assay for the intended samples.

### *7.1.1 Different mixing techniques*

To optimize and eliminate any interference, dilution experiments were carried out on saliva samples collected using both methods of collection. The samples were collected specifically for this purpose and named by letters A-C (kit) and G-I (no-kit). The experiments were repeated multiple times with variations to the protocol. Initially, the samples were analyzed following the standard protocol, where mixing during serial dilution was achieved by pipetting up and down several times. Subsequently, mixing was performed using a vortex mixer for five seconds. Finally, the sample was mixed for a longer duration before dilution, and the LAL reagent was prepared right before use, with longer mixing time. The samples were tested at undiluted (UD), 1:10, 1:100, and 1:1000 dilutions, and the absorbance readings at 405 nm were classified as out of range ("OR") when above 0.7 optical density (OD) and "OR (low)" when below 0.1 OD. It is important to note that when the absorbance readings are out of range, the LPS concentration is not reliable, particularly for higher values on the standard curve.

The standard curves obtained from the dilution experiments were non-linear and had R2 values >0.85, as demonstrated by the standard curve for one of the first LAL Assay performed in this project (Figure 9).

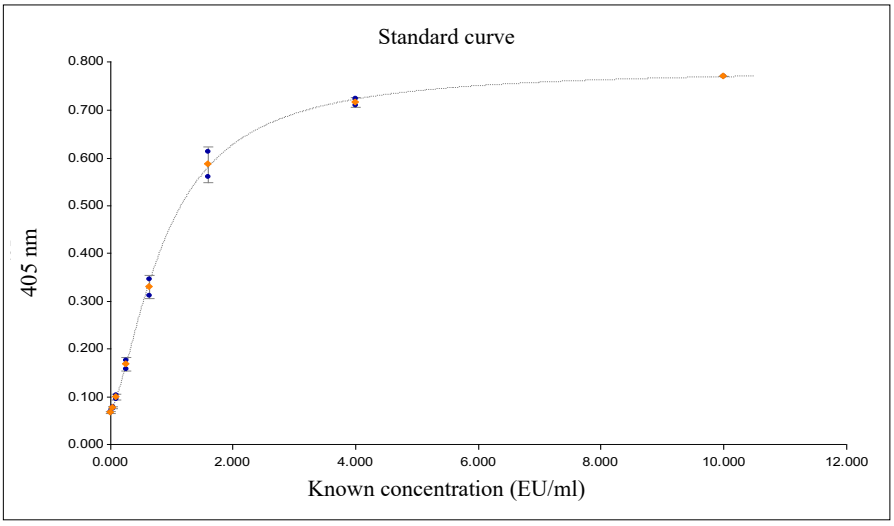



**Figure 9: Standard curve for the first LAL analysis of the experiment. Standard 1 (10 EU/ml) and 2 (4 EU/ml) had OD values 0.771 and 0.717, respectively. Standard 8 (0 EU/ml) with OD value 0.069 was used as a blank. The curve fit was R2=0.999.**

Table 5 presents the OD values obtained for each sample at different dilutions after the addition of stop solution to the wells. Notably, the no-kit samples (A-C) exhibited generally higher OD values and LPS concentration compared to the kit samples (G-I). When undiluted, the three kit samples showed OR OD values, and approximately 0.2 OD for the 1:10 dilution. Among the kit samples, the G and H samples exhibited low OD values across all dilutions, whereas the I sample exhibited higher OD values upon further dilution. Specifically, dilutions of 1:100 and 1:1000 showed significantly higher OD values for sample I compared to the other two kit samples, which exhibited values below the standard curve minimum value (OD 0.1). The LPS concentration was calculated using the OD measured closest to the middle of the standard curve and multiplying by the dilution factor. Consequently, the I sample had a much higher calculated LPS concentration than the other kit samples, as the OD used for calculation was multiplied with a higher dilution factor.

The no-kit samples generally decreased in OD value as the dilution factor increased. The A sample had the lowest LPS concentration of the no-kit samples. There was some variation between the different mixing techniques, with the two last techniques being more alike for most of the samples, with the exception of sample C.

**Table 5: OD values and LPS concentration for samples A-C and G-I. For OD values: The darker blue color, the higher OD value. The LPS concentration was calculated using the standard curve function and the OD values that fell within the low- to mid-range (i.e., OD 0.1-0.7) of the standard curve (e.g., OD value 0.27 for A no-kit when mixed by**

**pipetting). In cases where multiple OD values were available, the value closest to the middle value of the standard curve (OD 0.4) was used for LPS calculations. Finally, the LPS concentration was multiplied by the respective sample's dilution factor to yield the final concentration. OR = values outside the standard curve highest value.**

| Sample | 405 nm | | | | Mixing technique | LPS concentration (EU/ml) | SD |
|---|---|---|---|---|---|---|---|
| | Sample dilution | | | | | | |
| | 0 | 1:10 | 1:100 | 1:1000 | | | |
| A no-kit | OR | OR | OR | 0.27 | Mixing by pipetting | 445 | 42 |
| | OR | OR | 0.37 | 0.08 | Mixing by vortexing | 167.6 | 7.5 |
| | OR | OR | 0.54 | 0.14 | Mixing sample and LAL reagent longer | 192.5 | 9.8 |
| B no-kit | OR | OR | OR | 0.31 | Mixing by pipetting | 501 | 90 |
| | OR | OR | 0.55 | 0.07 | Mixing by vortexing | 362 | 50.7 |
| | OR | OR | OR | 0.16 | Mixing sample and LAL reagent longer | 393 | 5.7 |
| C no-kit | OR | OR | OR | 0.32 | Mixing by pipetting | 498 | 140 |
| | OR | OR | 0.57 | 0.05 | Mixing by vortexing | 391 | 44 |
| | OR | OR | OR | 0.24 | Mixing sample and LAL reagent longer | 736 | 67 |
| G kit | OR | 0.14 | 0.09 | 0.08 | Mixing by pipetting | 1.8 | 0.1 |
| | OR | 0.2 | 0.1 | 0.05 | Mixing by vortexing | 6.7 | 0.1 |
| | OR | 0.21 | 0.09 | 0.09 | Mixing sample and LAL reagent longer | 5.8 | 0.07 |
| H kit | OR | 0.14 | 0.09 | 0.09 | Mixing by pipetting | 1.9 | 0.06 |
| | OR | 0.27 | 0.08 | 0.06 | Mixing by vortexing | 7.1 | 0.07 |
| | OR | 0.21 | 0.08 | 0.07 | Mixing sample and LAL reagent longer | 6.3 | 0.2 |
| I kit | OR | 0.14 | 0.69 | 0.37 | Mixing by pipetting | 665 | 92.6 |
| | OR | 0.28 | 0.31 | 0.64 | Mixing by vortexing | 134.7 | 15.5 |
| | OR | 0.23 | 0.47 | 0.78 | Mixing sample and LAL reagent longer | 143.2 | 19.9 |

### *7.1.2 Using kit and no-kit samples from the same participants*

To account for random differences in LPS concentrations between the no-kit and kit samples, new LAL experiments were carried out using both types of samples taken concurrently from the same person. These samples were randomly selected from the 32 samples collected for the pilot project. All samples were mixed with vortex during serial dilution and the LAL reagent was mixed right before its addition to the wells. This was a consequence of the earlier

results mentioned above. Generally, the trend of kit samples having much lower signal and LPS concentration than no-kit samples continued. The kit samples were inconsistent, some rising and some falling in OD with higher dilutions (Table 6).

Considerable inter-individual variation in LPS concentrations was observed (Table 6). In particular, for participant number 2 the no-kit sample (diluted 1:1000) measured 0.66 OD and had an LPS concentration of 3666 EU/ml. The corresponding kit sample at 1:10 dilution measured 0.21 OD and had an LPS concentration of 10.5 EU/ml. For participants 10, 21 and 22 the no-kit samples were OR for all dilutions, thus rendering it impossible to determine their LPS concentration. In contrast, the respective kit samples displayed LPS concentrations of 2717 EU/ml (0.55 OD), 162.8 EU/ml (0.43 OD), and 183.2 EU/ml (0.39 OD), respectively. LPS concentration could not be determined for kit samples number 19 and 28, as they both had too low OD signal. The no-kit samples number 19 and 28, on the other hand, exhibited high OD signals at a dilution of 1:1000, with calculated LPS concentrations of 7560 EU/ml and 2743 EU/ml, respectively (Table 6).

**Table 6: LAL analysis of both no-kit and kit samples for participants number 2, 10, 19, 21, 22 and 28. All samples measured at dilution 0, 1:10, 1:100 and 1:1000. LPS concentration determined by OD within range and closest to mid value 0.4 and multiplied with the respective dilution factor. OR = values outside the standard curve highest value.**

| Sample | 405 nm (OD) | | | | LPS concentration (EU/ml) |
|---|---|---|---|---|---|
| | Sample dilution | | | | |
| | 0 | 1:10 | 1:100 | 1:1000 | |
| 2 no-kit | OR | OR | OR | 0.66 | 3666 |
| 2 kit | OR | 0.21 | 0.08 | 0.07 | 10.5 |
| 10 no-kit | OR | OR | OR | OR | N/A |
| 10 kit | OR | 0.26 | 0.21 | 0.55 | 2717 |
| 19 no-kit | OR | OR | OR | 0.58 | 7560 |
| 19 kit | OR | OR (low) | OR (low) | OR (low) | N/A |
| 21 no-kit | OR | OR | OR | OR | N/A |
| 21 kit | OR | 0.29 | 0.43 | OR | 162.8 |
| 22 no-kit | OR | OR | OR | OR | N/A |
| 22 kit | OR | 0.16 | 0.39 | 0.81 | 183.2 |
| 28 no-kit | OR | OR | OR | 0.63 | 2743 |
| 28 kit | OR | OR (low) | OR (low) | OR (low) | N/A |

## 7.3 Pilot project

### *7.3.1 LPS concentration of pilot project samples*

Due to inter-individual variation in the LPS levels and potential interference in the kit samples, it was challenging to determine an optimal dilution for all samples. In the case of no-kit samples, they were initially evaluated at a dilution of 1:1000, and if they yielded out of range (OR) results, further dilution was performed. The kit samples collected for the pilot project were included for analysis to possibly verify the interference. All the kit samples were first analyzed at a dilution of 1:10, followed by further or lesser dilution if necessary.

The LPS concentrations for the no-kit samples varied from 106 EU/ml to 46000 EU/ml (Figure 10A), with a mean value of 6453 (± 10644) and a median value of 2970 EU/ml. In contrast, the LPS concentrations for the kit samples ranged from 5.73 EU/ml to 19.14 EU/ml with a mean value of 8.17 EU/ml (± SD 2.46) and a median value of 7.55 EU/ml (Figure 10B).

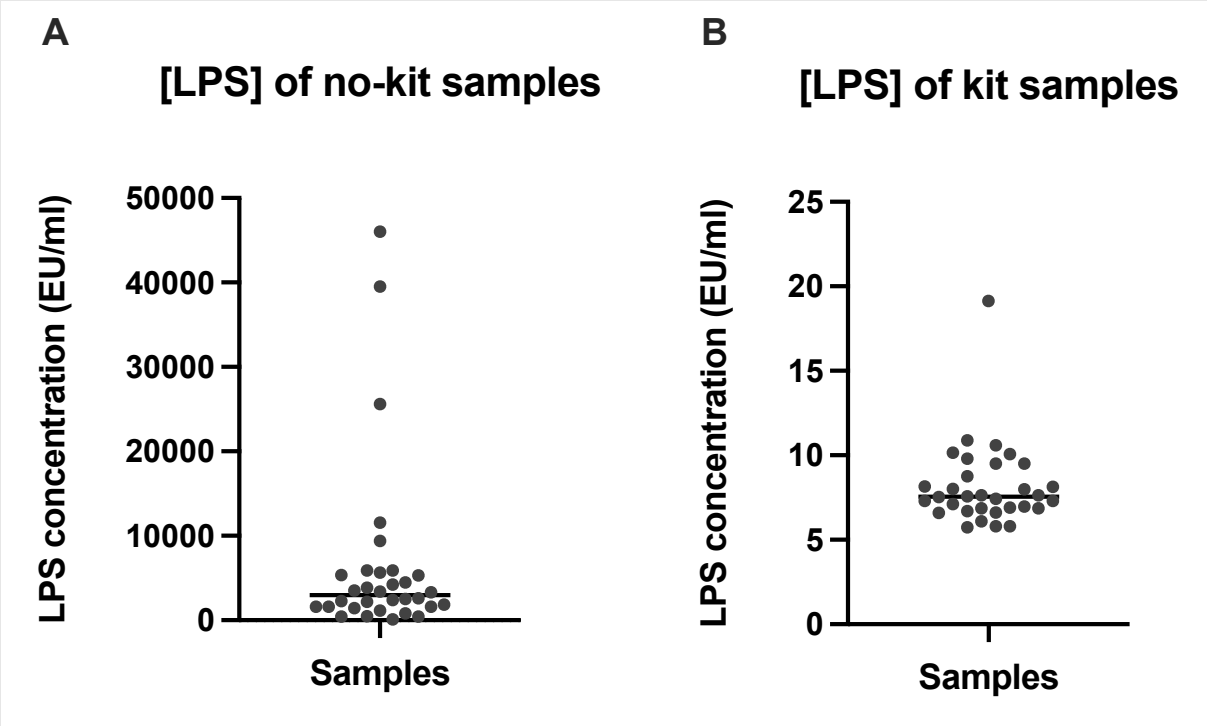



**Figure 10: A) Measured LPS concentration of no-kit samples (1-32) from the pilot project. Median value 2970 EU/ml. B) Measured LPS concentration of kit samples (1-32) from the pilot project. Median value 7.55 EU/ml.**

### *7.3.2 DNA concentration and purity of pilot project samples*

The DNA concentrations in the no-kit samples were found to range from 0.21-159.6 ng/µl, with a mean value of 48 (± 40.81) and median 41.25 ng/µl. The kit samples showed a range of

0.28-207.6 ng/µl, with a mean value of 65.21 (± 53.87) and median 51.7 ng/µl (Figure 11). The DNA concentrations were determined using the Qubit assay, as outlined in the Methods section. However, due to a lack of Qubit Broad Range reagents during the pilot project, NanoDrop was used to determine the DNA concentrations for samples that were OR (high) on Qubit. This was the case for a total of 31 samples, including both kit and no-kit samples.

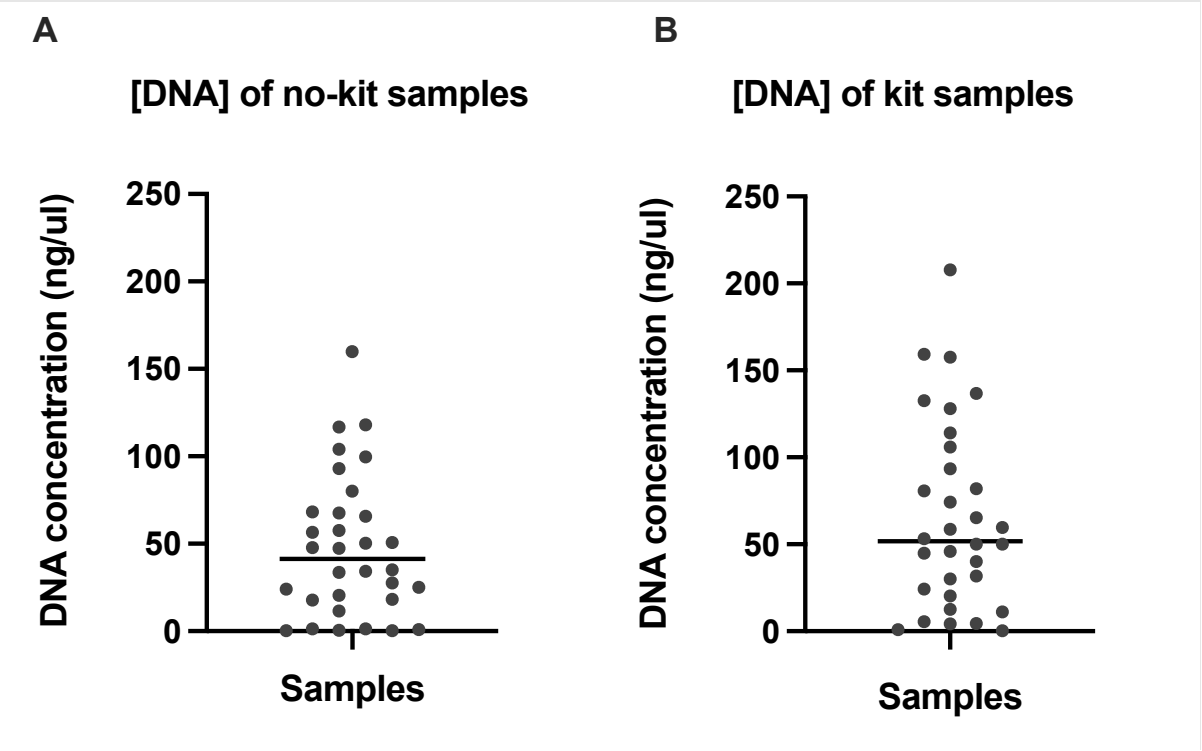



**Figure 11: DNA concentration measured in pilot project samples 1-32. A) [DNA] of no-kit samples, median value 41.45 ng/µl. B) [DNA] of kit samples, median value 51.7 ng/µl.**

The purity of DNA was assessed through the measurement of the 260/280 and 260/230 ratios using NanoDrop, as detailed in the Methods chapter. The median 260/280 ratio for no-kit samples was 1.77 (IQR: 1.72-1.85), and for kit samples it was 1.78 (1.72-1.83). Regarding the 260/230 ratio, no-kit samples had a median value of 0.43 (0.1-0.64), whereas kit samples had a median value of 0.52 (0.22-0.83).

### *7.3.3 Normalization*

Due to sample size the normality of the LAL results for both the no-kit and kit samples were assessed using the Shapiro-Wilk test (α: 0.05). Including all LPS values in the test, neither the no-kit nor the kit samples' LAL results passed the Shapiro-Wilk test of normality, with both p-values smaller than 0.0001 (W = 0.55 and 0.71, respectively). Removing the outliers, specifically no-kit samples number 12 (39500 EU/ml), 13 (25600 EU/ml), and 21 (46000

EU/ml), and kit sample number 12 (19141 EU/ml), resulted in a p-value of 0.003 (W=0.87) for the no-kit samples and p-value = 0.04 (W=0.92) for the kit samples, still not passing the test of normality. To visualize this, the normal quantile-quantile (QQ) plots for all values are depicted in Figure 12. Both sample types do not follow the expected diagonal line in the QQ plot, indicating that they do not conform to a Gaussian distribution.

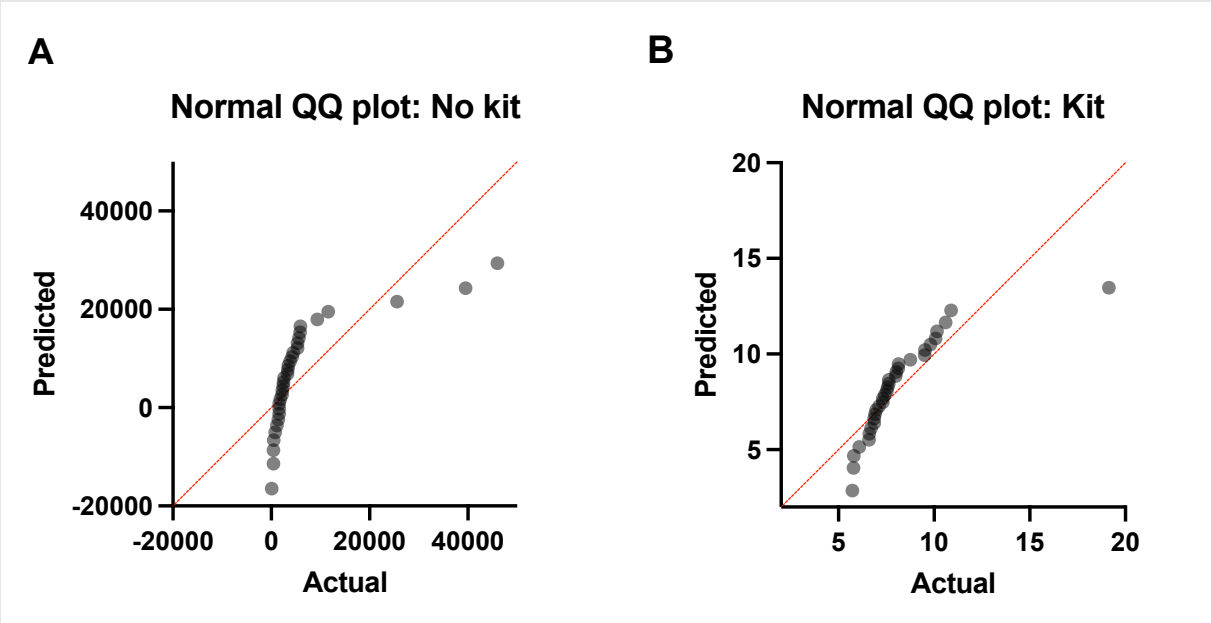



**Figure 12: Normal QQ plots including predicted and actual values for all samples. A) No kit samples. B) Kit samples. Both sample types deviate from the line which indicates non-Gaussian distribution.**

### 7.4 Spiking of samples

To investigate the source of interference, a spiking experiment was performed by adding different dilutions of *E. coli* standards to undiluted home-kit buffer, PBS and water (EFW, from the kit). The Hycult LAL assay was used to measure LPS concentrations, adding known amounts of endotoxin using the *E. coli* standard from the Lonza kit.

The LPS concentration in the home-kit buffer was undetectable, as the OD value of the blank buffer was higher than that of the samples. Consequently, the recovery rate was zero, indicating that none of the added endotoxin was detected in the samples. In contrast, the measured LPS concentrations in the PBS and EFW showed less deviation from the known amount of added endotoxin, and the recovery rate ranged from 80.1-135.32%, which is an acceptable range (Table 7).

**Table 7: Spiking of samples using the Hycult LAL assay. 5 EU/ml, 2.5 EU/ml and 1.25 EU/ml *E. coli* samples were used to spike buffer, PBS and EFW. Due to dilution in the wells, the known LPS concentration in the samples were 2.5, 1.25 and 0.75 EU/ml.**

| Sample type | Known concentration (EU/ml) | Mean LPS concentration (EU/ml) | %Recovery |
|---|---|---|---|
| Home-kit buffer | 2.5 | -0.105 | 0 |
| | 1.25 | -0.099 | 0 |
| | 0.75 | -0.145 | 0 |
| PBS | 2.5 | 3.321 | 135.32 |
| | 1.25 | 1.247 | 104.70 |
| | 0.75 | 0.565 | 83.60 |
| EFW | 2.5 | 2.236 | 89.42 |
| | 1.25 | 1.001 | 80.10 |
| | 0.75 | 0.712 | 94.94 |

Together with the results from spiking of undiluted samples, the data obtained from the dilution process and pilot study indicated a notable deficiency in the recovery rate of the home-kit buffer. Consequently, a further spiking experiment was conducted, which involved diluting home-kit buffer, PBS, and EFW up to 1:10000 and spiking them with 0.75 EU/ml. The dilutions of PBS and EFW exhibited acceptable recovery rates ranging from 83.47-95.20% (Table 8). However, for the home-kit buffer, only the 1:1000 and 1:10000 dilutions demonstrated acceptable recovery rates of 128.8% and 105.07%, respectively, while the 1:2 and 1:10 dilutions did not. The 1:2 dilution had a recovery rate of only 21.87%, with the mean LPS concentration almost outside the standard curve. Similarly, the buffer 1:10 dilution was outside the standard curve, and the LPS concentration and %recovery could not be determined (Table 8).

**Table 8: Dilution, mean LPS concentration and %recovery for Norgen buffer, PBS and EFW spiked with 0.75 EU/ml. Known concentration of the samples were 0.375 due to further dilution in the wells.**

| Sample type | Dilution | Mean LPS concentration (EU/ml) | %Recovery |
|---|---|---|---|
| Home-kit buffer | 1:2 | 0.082 | 21.87 |
| | 1:10 | N/A | N/A |
| | 1:100 | 0.124 | 33.07 |
| | 1:1000 | 0.483 | 128.80 |
| | 1:10000 | 0.394 | 105.07 |
| PBS | 1:2 | 0.356 | 95.07 |
| | 1:10 | 0.363 | 96.80 |
| | 1:100 | 0.364 | 97.20 |
| | 1:1000 | 0.331 | 88.40 |
| | 1:10000 | 0.313 | 83.47 |
| EFW | 1:2 | 0.357 | 95.20 |
| | 1:10 | 0.348 | 92.80 |
| | 1:100 | 0.332 | 88.53 |
| | 1:1000 | 0.331 | 88.40 |
| | 1:10000 | 0.310 | 82.67 |

### 7.5 Lonza kit

In order to determine if our samples could be better measured using a different LAL assay, an alternative LAL kit was utilized to measure LPS concentrations in select samples. The Lonza LAL kinetic-QCL Assay was used for this purpose. The standard curve, illustrated in Figure 13, exhibited an R value of 0.981, with maximum V values of 170, 121, 60, and 1.2 for Standards 1 (25 EU/ml), 4 (3.125 EU/ml), 8 (0.195 EU/ml), and 12 (0.012 EU/ml), respectively.

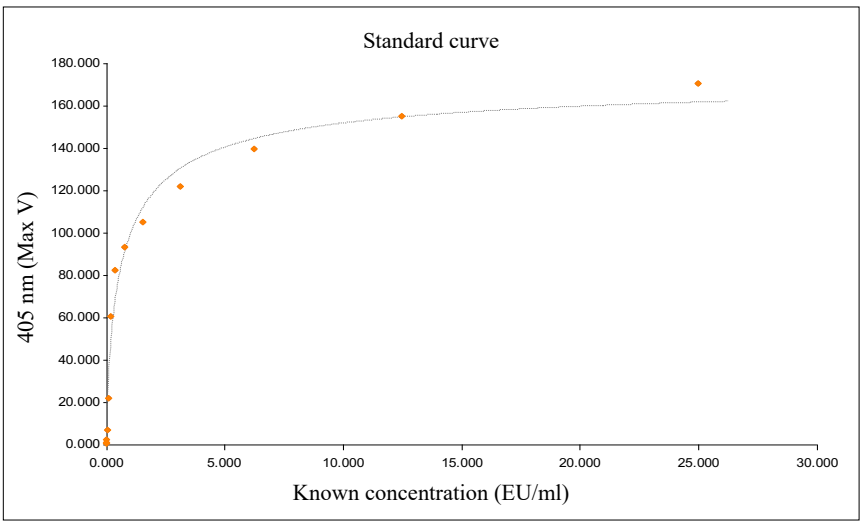



**Figure 13: Standard curve obtained using standards 0.012-25 EU/ml from the Lonza LAL Kinetic QCL assay. x: The known concentrations of the standards. y: 405 nm measurement at Max V**

LPS concentration was measured using randomly selected samples from the pilot study. The samples were serial diluted up to 1:10000. The positive control (10 EU/ml) had mean concentration 8.17 EU/ml. The trend of variability in both signal and concentration observed in the previous experiments with the kit samples persisted (Table 9). The LPS concentration for the no-kit samples was relatively similar between the 1:100 and 1:1000 dilutions, while 1:10 and 1:10000 dilutions exhibited greater variability.

**Table 9: LPS concentration of randomly selected samples measured at different dilutions. OR = values outside the standard curve highest value. OR (low) = values outside the standard curve lowest value.**

| Sample | LPS concentration at Max V | | | |
|---|---|---|---|---|
| | Dilution | | | |
| | 1:10 | 1:100 | 1:1000 | 1:10000 |
| 2 no-kit | OR | 444.10 | 521 | 605 |
| 2 kit | 16.64 | 3.85 | 26.5 | OR (low) |
| 6 no-kit | OR | 331.80 | 384.50 | 500 |
| 6 kit | 70.12 | 3.95 | 52.5 | OR (low) |
| 20 no-kit | 121.89 | 70.95 | 77 | OR (low) |
| 20 kit | OR | 4.2 | 48.5 | OR (low) |
| 30 no-kit | OR | 2046.75 | 1769.5 | 1935 |
| 30 kit | 0.86 | 5.15 | 689.5 | 3205 |

A spiking experiment in undiluted home-kit buffer, PBS, and water (LAL Reagent water, LRW, from the kit) was also carried out using the Lonza kit. Different dilutions of the Hycult LAL Assay standard solution were used to spike the samples. The results revealed that the spiked buffer demonstrated very low recovery rates, with the maximum %recovery being 1.13% and the LPS concentration recorded at 0.009. On the other hand, the spiked PBS and LRW demonstrated better %recovery, with a mean %recovery of 117.9 and 86.6, respectively (Table 10).

**Table 10: Spiking of samples using the Lonza LAL assay. OD measured at 405 nm (Max V), mean LPS concentration and %recovery of spiked undiluted Norgen buffer, PBS and LRW.**

| Sample type | Known concentration (EU/ml) | 405 nm (Max V) | Mean LPS concentration (EU/ml) (±SD) | %Recovery |
|---|---|---|---|---|
| Norgen buffer | 5 | -0.5 | 0.008 (0) | 0.16 |
| | 2.5 | -0.3 | 0.008 (0) | 0.32 |
| | 1.25 | -0.3 | 0.008 (0) | 0.64 |
| | 0.75 | -0.05 | 0.009 (0.001) | 1.13 |
| PBS | 5 | 145.3 | 6.71 (2.15) | 134.11 |
| | 2.5 | 125.5 | 2.54 (0.08) | 101.56 |
| | 1.25 | 109.4 | 1.42 (0.01) | 113.68 |
| | 0.75 | 95.4 | 0.92 (0.25) | 122.20 |
| LRW | 5 | 136.3 | 4.10 (0.83) | 81.94 |
| | 2.5 | 116.9 | 1.84 (0.03) | 73.72 |
| | 1.25 | 100.0 | 1.04 (0.03) | 83.16 |
| | 0.75 | 91.45 | 0.81 (0.19) | 107.4 |

## 7.6 RHINESSA

The kit samples from the RHINESSA study were excluded from the LAL analysis due to interference in the samples.

### *7.6.1 Study population characteristics*

Initially, the experiment included 87 participants based on the samples' availability. Eight samples were excluded from statistical analysis either due to recent antibiotic use (n=2) or lack of information on the age group (adolescence, n=6). Consequently, 79 participants were

included in the statistical analysis, comprising 16 men (20%) and 63 women (80%) (Table 11). The female and male participants were about the same age (median (IQR): 27 (23-34), 28 (26-30), respectively). The age range of the participants was between 18-45 and 18-33 years for women and men, respectively. The male participants were taller ($p < 0.0001$) and had higher weight ($p < 0.0001$) than their female counterparts ((median (IQR): 180 (177.3-188.3) cm and 88 (74-100) kg, 168 (165-171) cm and 64 (60-73) kg, respectively) Notably, snus usage was more prevalent among men (18.75%) than women (9.52%), whereas female participants were more prone to smoking than male participants (7.94% vs 0%, respectively). The Community Periodontal Index (CPI) was measured and reported by the fieldworkers, while toothbrushing frequency and gum bleeding were self-reported. Of the participants, 8 lacked CPI measurements, and only the maximum CPI score (highest score recorded across all the measured teeth) was used, which ranged from 0 to 3. Of the participants, 88.6% brushed their teeth twice daily or more frequently, and 63.29% reported to rarely or never experience bleeding from the gums during tooth brushing. Of those participants who used asthma medication, two had experienced asthma attacks in the last 12 months. In addition, two participants who did not use asthma medication reported having had asthma attacks (Table 11).

**Table 11: Study population characteristics obtained from participants in the RHINESSA study (n=79) included for statistical analysis and further work in this project.**

| Variable | All (*n* = 79) | Women (*n* = 63) | Men (*n* = 16) |
|---|---|---|---|
| Age, years | | | |
| Mean (SD) | 28.2 (6.43) | 28.5 (6.83) | 27.0 (4.70) |
| Range, min-max | 18-45 | 18-45 | 18-33 |
| Height, cm | | | |
| Mean (SD) | 171.3 (8.10) | 168.2 (5.0) | 182.9 (6.83) |
| Range, min-max | 156-197 | 156-182 | 175-197 |
| Weight, kg | | | |
| Mean (SD) | 71.4 (15.01) | 67.2 (12.43) | 86.8 (14.39) |
| Range, min-max | 50-107 | 50-107 | 61-106 |
| Smoking habits, *n* (%) | | | |
| Never[a] | 63 (79.75) | 48 (76.19) | 15 (93.75) |
| Previous | 11 (13.92) | 10 (15.87) | 1 (6.25) |
| Current | 5 (6.33) | 5 (7.94) | 0 (0.00) |
| Snus[b] (snuff) habits, *n* (%) | | | |
| Never[a] | 60 (75.95) | 51 (80.95) | 9 (56.25) |

| | | | |
|---|---|---|---|
| Previous | 7 (8.86) | 4 (6.35) | 3 (18.75) |
| Current | 9 (11.39) | 6 (9.52) | 3 (18.75) |
| Toothbrushing habits, *n* (%) | | | |
| Never to 1 time per day | 8 (10.13) | 7 (11.11) | 1 (6.25) |
| Twice per day | 66 (81.01) | 50 (79.37) | 15 (93.75) |
| More than twice per day | 6 (7.59) | 6 (9.52) | 0 (0.00) |
| Gum bleeding[b] while brushing teeth, *n* (%) | | | |
| Never | 12 (15.19) | 10 (15.63) | 2 (12.50 |
| Rarely | 39 (49.37) | 31 (49.21) | 8 (50.00) |
| Sometimes | 20 (25.32) | 16 (25.00) | 4 (25.00) |
| Often | 5 (6.33) | 4 (6.34) | 1 (6.25 |
| Always | 2 (2.53) | 2 (3.17) | 0 (0.00) |
| CPI[c], *n* (%) | | | |
| 0 | 61 (77.22) | 50 (79.37) | 11 (68.75) |
| 1 | 2 (2.53) | 2 (3.17) | 0 (0.00) |
| 2 | 5 (6.33) | 2 (3.17) | 3 (18.75) |
| 3 | 3 (3.80) | 3 (4.76) | 0 (0.00) |
| Asthma attacks the last year, *n* (%) | | | |
| Yes | 4 (5.06) | 3 (4.69) | 1 (5.88) |
| Current use of any asthma medication, *n* (%) | | | |
| Yes | 4 (5.06) | 3 (4.69) | 1 (5.88) |

[a]Never and Current smoker/snuff user were defined as having answered no and yes, respectively, to the question "Do you smoke?"/ "Do you use snuff?". Previous smoker/snuff user was defined as answering yes to the question "Did you smoke previously?"/" Did you use snuff previously?". [b]Missing data for Snus (n=3); CPI (n=8); Gum bleeding (n=1). [c]No participants with score above 3.

### *7.6.2 LPS and DNA concentration of RHINESSA samples*

The male participants exhibited a median DNA concentration of 6.43 ng/µl, ranging from 0.02-31.9 ng/µl. The DNA concentration of the samples collected from female participants ranged from 0.05-63.9 ng/µl, with median value 8.1 ng/µl (Figure 14A). Notably, three participants (one male and two females) had DNA concentrations below the detection limit when utilizing 2 µl sample volume for Qubit measurements. The LPS concentration exhibited a higher mean value in males (11709 (±12298) EU/ml, median: 7140 EU/ml) as compared to females (mean: 5069 (±6919) EU/ml, median: 2669 EU/ml) (Figure 14B). The significance of this difference is discussed below. The median 260/280 ratio for all samples was 1.77 (1.46-1.85), while the median 260/230 ratio was 0.12 (0.05-0.26).

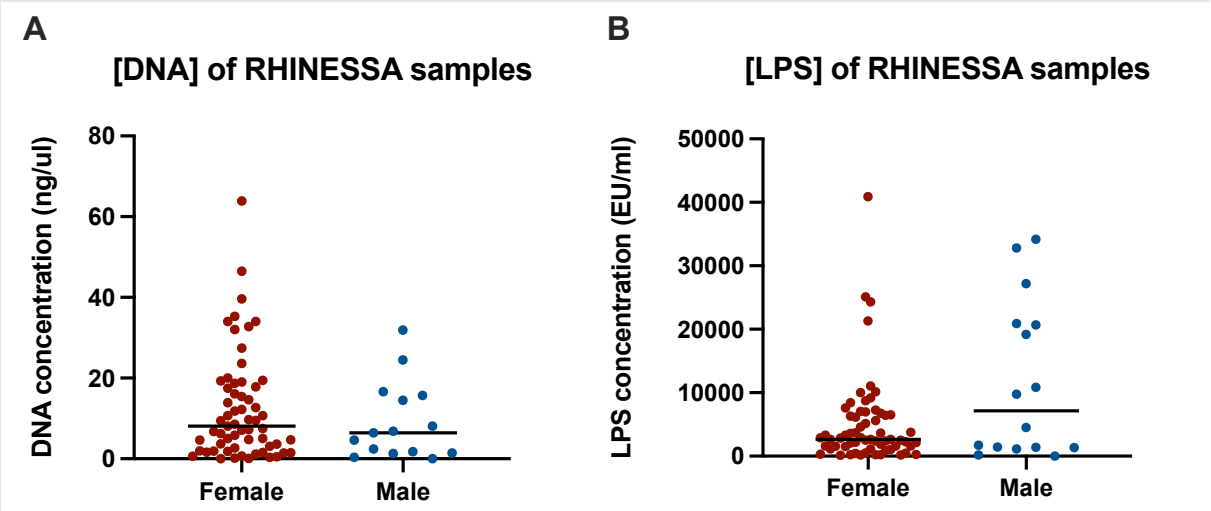



**Figure 14: A) [DNA] measured in RHINESSA kit and no-kit samples from 79 participants. Female median value: 8.1 ng/µl. Male median value: 6.43 ng/µl. B) [LPS] measured in the same samples. Female median value: 2660 EU/ml. Male median value 7140 EU/ml.**

Shapiro-Wilk test results indicated that the data were non-Gaussian distributed for both DNA and LPS concentration, with W values = 0.86 and 0.96, and p-values = 0.01 and 0.04, respectively. Normalization for LPS concentration for both genders is shown in Figure 15 (QQ plot). Furthermore, the Mann-Whitney (MW) test revealed a non-significant difference in LPS concentration between male and female participants, with a p-value of 0.3. MW test of DNA concentration between the genders also had p-value = 0.3.

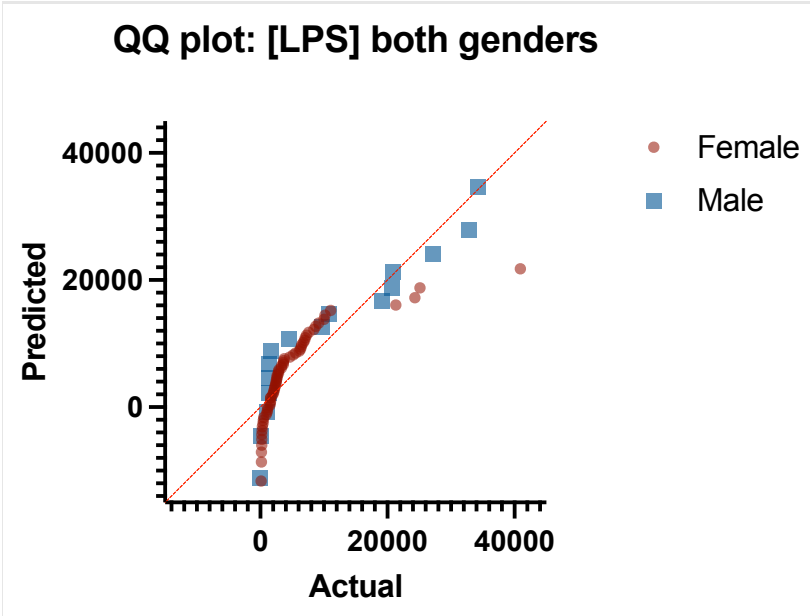



**Figure 15: Normal QQ plot of LPS concentration for both genders. None of the genders follow the expected diagonal line in the QQ plot, indicating that they do not conform to a Gaussian distribution.**

Statistical analysis showed that none of the known factors of oral disease led to significantly higher or lower LPS concentration. Participants with CPI>0 (n=9) had a median LPS concentration of 4610 (IQR: 361-22320) EU/ml, which was higher than participants with CPI=0 (median LPS concentration 2540 (1370-6875), but the difference was not statistically significant (MW p-value 0.7). Similarly, participants who reported gum bleeding rarely/never had a mean LPS concentration of 2660 (1160-7610) EU/ml while participants who often/sometimes/always (n=27) had a median LPS concentration of 2890 (1440-6970) EU/ml, but the difference was not statistically significant (MW P value 0.7) (Figure 16).

Furthermore, tobacco users (n=14) had a higher median LPS value of 4380 (1530-12795) EU/ml than non-tobacco users (n=57) with a median LPS concentration of 2850 (1134-6875) EU/ml, but the difference was not statistically significant (MW P-value 0.3) (Figure 16). When examining the LPS concentrations among various categories of tobacco users, it was found that current smokers exhibited a median LPS concentration of 2200 (904-8085) EU/ml, whereas current snus users demonstrated a higher median LPS concentration of 6420 (1990-20800) EU/ml.

Participants who reported use of antibiotics in the last 12 months (n=27) had lower median LPS concentration of 2520 (1138-6300) EU/ml compared to those who did not use antibiotics (median LPS concentration 3480 (1460-9970) EU/ml), but the difference was not statistically significant (MW p-value 0.3). Participants who brushed their teeth one or less than once per day (n=8) had median LPS concentration of 2420 (658-8475) EU/ml while the participants who brushed their teeth from 2 to more than two times a day had median LPS concentration of 2890 (1440-7280) EU/ml (MW P-value 0.6) (Figure 16).

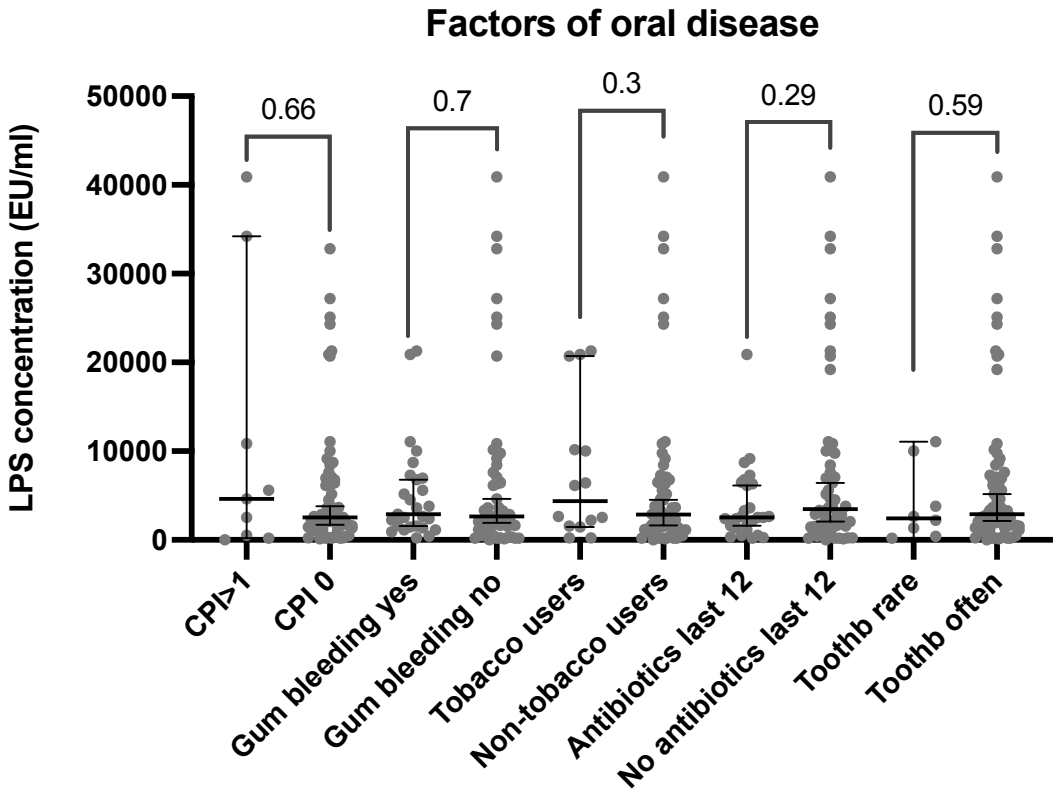



**Figure 16: Different factors and contributors to oral disease. Shown with median value as line with 95% CI. Points indicate measured LPS concentration per sample. P-value from MW test was calculated between the corresponding groups, such as CPI>1 vs CPI 0 and gum bleeding yes and no. Gum bleeding yes = Gum bleeding often/sometimes/always. Gum bleeding no = rare/never. Tobacco users = current smokers or snuffers. Toothb rare = Toothbrushing 1 or less than one time per day. Toothb often = toothbrushing from 2 to more than two times per day.**

## 7.7 Lipid A annotation of saliva and GCF samples

ASVs were obtained from 78 saliva samples and 79 GCF samples collected at baseline level (2014-2015) in the RHINESSA Bergen study.

### *7.7.1 Bacterial diversity and taxa distribution*

In the saliva samples, analysis revealed the presence of 16 distinct bacterial phyla. The most prevalent phyla observed were Firmicutes (59.19%), Actinobacteria (20.82%), Bacteroidetes (9.69%), and Proteobacteria (5.41%), demonstrating their dominance in the saliva microbial composition (Figure 17). Within the phyla, a total of 139 genera were identified, with notable abundance exhibited by key genera such as *Streptococcus* (41.34%, phylum Firmicutes), *Actinomyces* (9.62%, phylum Actinobacteria), *Rothia* (8.12%, phylum Actinobacteria), and *Prevotella* (7.25%, phylum Bacteroidetes) (Figure 17).

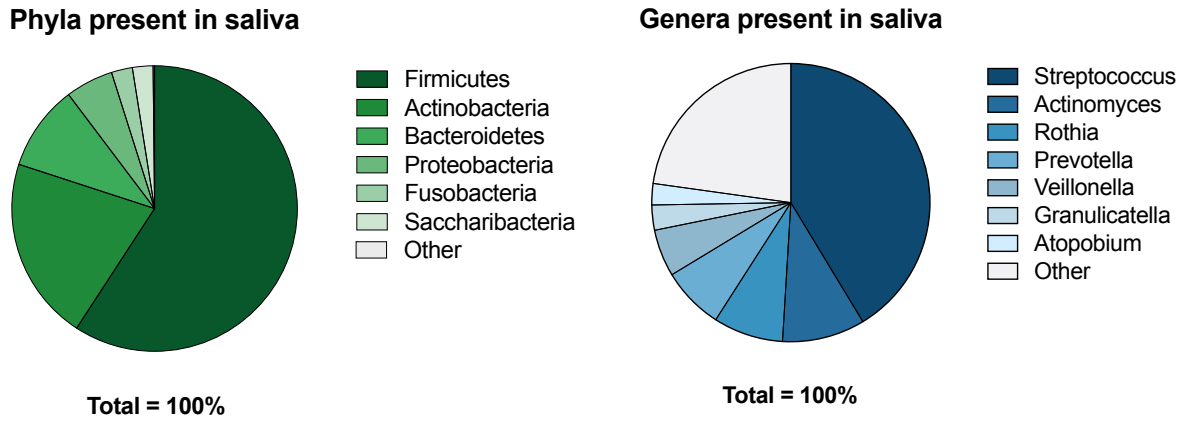



**Figure 17: All phyla and genera found in 78 RHINESSA saliva samples collected at baseline (year 2014-2015). Firmicutes (59.19%), Actinobacteria (20.82%), Bacteroidetes (9.69%) and Proteobacteria were the dominating phyla out of the 16 identified. *Streptococcus* (41.34%), *Actinomyces* (9.62%), *Rothia* (8.12%), and *Prevotella* (7.25%) were the dominating genera of the 139 identified.**

In GCF, ten different bacterial phyla were detected, with Bacteroidetes (25.19%), Firmicutes (24%), Fusobacteria (19.36%), Proteobacteria (14.62%), and Actinobacteria (7.8%) being the most prevalent (Figure 18). Among the 154 genera identified, *Fusobacterium* (16.1%, phylum Fusobacteria), *Streptococcus* (10.85%, phylum Firmicutes), *Prevotella* (7.95%, phylum Bacteroidetes), and *Capnocytophaga* (6.56%, phylum Bacteroidetes) dominated (Figure 18).

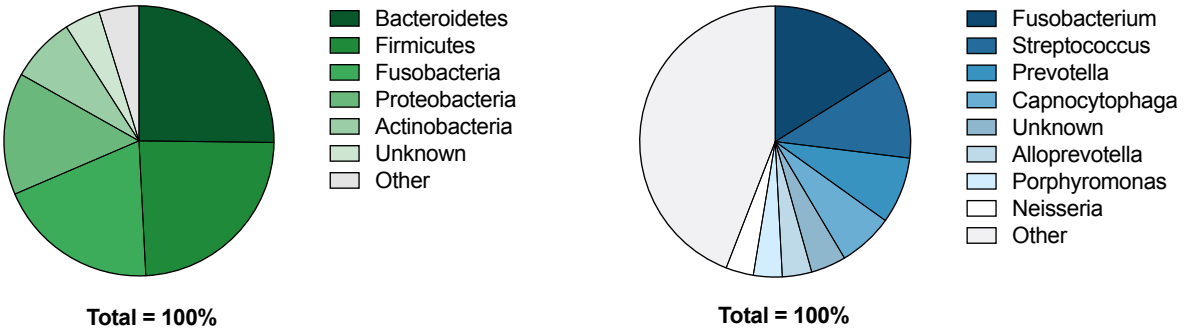



**Figure 18: All phyla and genera found in 79 RHINESSA GCF samples collected at baseline (year 2014-2015). Bacteroidetes (25.19%), Firmicutes (24%), Fusobacteria (19.36%) and Proteobacteria (14.62%) were the dominant phyla of 10 identified. *Fusobacterium* (16.1%), *Streptococcus* (10.85%) and *Prevotella* (7.95%) were the dominant genera of 154 identified.**

### *7.7.2 Annotation*

Among the identified bacteria in saliva, 73.37% were annotated as Gram-positive (G+) whereas 26.52% were annotated as Gram-negative (G-). (Figure 19). Additionally, 0.11% of the bacteria were either of unknown species or had unknown response to Gram-staining. Within the Gram-negative bacteria, 80.84% were identified as penta-acylated, 10.13% as hexa-acylated, and 9.02% exhibited unknown lipid A variants or lacked LPS in their OM (Figure 19). No bacteria were identified as solely tetra-acylated.

At the genus level, the most prevalent hexa-acylated bacteria was *Haemophilus*, which was present in 77 samples, and *Aggregibacter* which was present in 56 (72%) samples. For penta-acylated bacteria the most prevalent in saliva was *Veillonella* which was in all samples and *Leptotrichia, Porphyromonas,* and *Prevotella* which were present in 77 samples. The most prevalent Gram-positive bacteria were *Streptococcus* and *Rothia*, present in all samples.

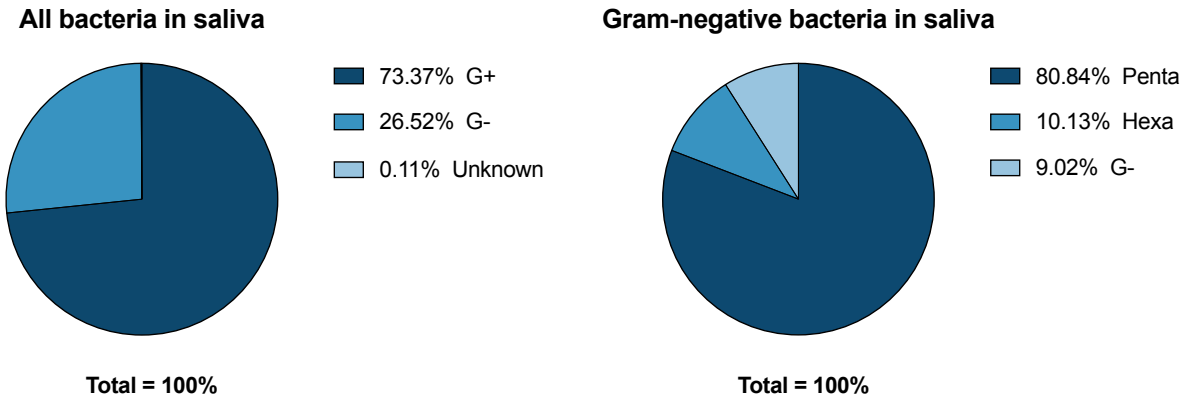



**Figure 19: All bacteria found in 78 RHINESSA saliva samples collected at baseline (year 2014-2015). 73.37% of the found bacteria were Gram-positive (G+) and 26.52% Gram-negative (G-). Of the G- bacteria found in the samples, 80.84% were penta-acylated while 10.13% were hexa-acylated. 9.02% were G- bacteria with unknown LPS variants.**

In GCF, 71.04% of the bacterial genera present were annotated as Gram-positive (G+), 28.6% were annotated as Gram-negative (G-), and 0.37% were unknown bacteria (Figure 20). Of the 78 Gram-negative bacteria, 89.71% exhibited penta-acylated lipid A, 3.71% exhibited hexa-acylated lipid A, and the remaining 6.57% were assigned as Gram-negative bacteria, but either had uncertainty regarding their lipid A structure or lacked lipopolysaccharide (LPS) in their cell wall (Figure 20). No bacteria were found to have only tetra-acylated LPS.

The dominant bacterial genus capable of producing hexa-acylated lipid A was taxonomically annotated to *Haemophilus*, which was present in all 79 samples. *Aggregatibacter*

was present in 75 (95%) samples, while *Proteus* was present in 59 (75%) samples. *Fusobacterium*, *Prevotella*, and *Capnocytophaga* were the most commonly occurring bacterial genera producing penta-acylated lipid A and were present in all 79 samples. The dominant Gram-positive genera included *Streptococcus*, *Actinomyces*, and *Rothia*, also detected in all samples.

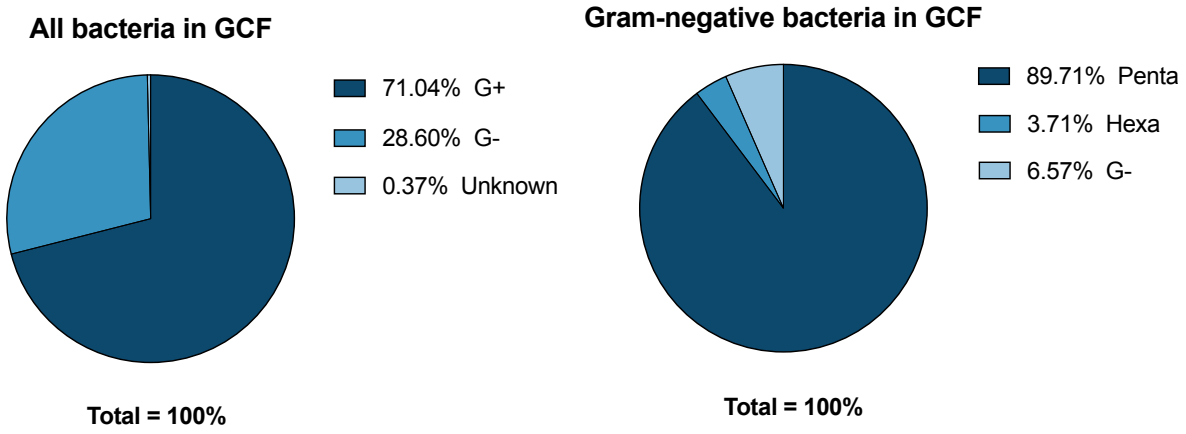



**Figure 20: All bacteria found in 79 RHINESSA GCF samples collected at baseline (year 2014-2015). A total of 154 bacteria genera were found. In terms of abundance, 71.04% of the bacteria were Gram-positive (G+) and 28.6% were Gram-negative (G-). Of the G- bacteria found in the samples, 89.71% were penta-acylated while 3.71% were hexa-acylated. 6.57% were annotated as G- bacteria with unknown LPS variants.**

### *7.7.3 Correlation between LPS concentration and Lipid A variants*

When dividing the samples into quantiles based on the LPS concentration, the proportion of Gram-negative (G-) bacteria increased with the amount of LPS present. This happened in both sample types, as illustrated for GCF samples (Figure 21). A more detailed examination of the composition of Gram-negative bacteria in the samples revealed differences in the amounts of hexa- and penta-acylated lipid A-producing bacteria (see Figure 22).

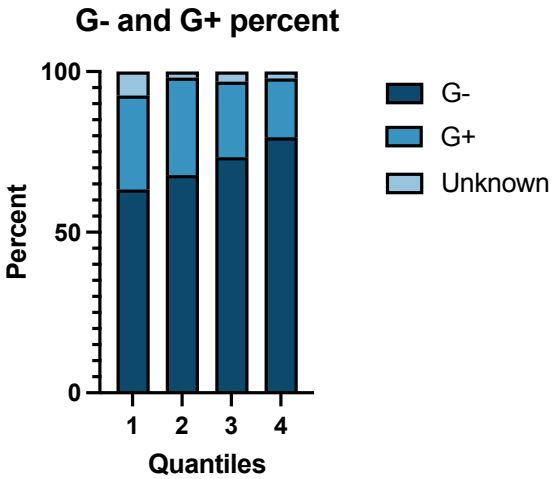



**Figure 21: Percentages of Gram-negative (G-), Gram-positive (G+) and unknown bacteria in GCF. Samples from the RHINESSA study. The samples were divided into quantiles based on LPS concentration.**

In saliva samples, the proportion of penta-acylated lipid A-producing bacteria in the quantiles ranged from 78.37% to 85.61%, while the proportion of hexa-acylated lipid A-producing bacteria ranged from 7.09% to 12.7%. In GCF the range for penta- and hexa-acylated lipid A variants were 86.58% to 92.11% and 2.22% to 4.99%, respectively. All quantiles in both sample types contained some Gram-negative bacteria with unknown lipid A variants (Figure 22).

To explore the association between LPS concentration and the bacterial community (based on ASVs), we employed a Spearman correlation test on the CLR transformed ASVs obtained from the saliva and GCF RHINESSA Bergen samples. In both sample types, the ASV sum of penta-acylated bacteria correlated positively with LPS concentration. The values were r = 0.358 in GCF and r = 0.082 in saliva. The ASV total sum of hexa-acylated bacteria were negative in both sample types with values r = -0.013 for GCF and r = -0.155.

In saliva samples, the Gram-positive bacteria *Bulleidia* and *Atopobium* exhibited the highest positive correlation with LPS concentrations, with r = 0.427 and r = 0.399, respectively. Among the Gram-negative bacteria, penta-acylated *Selenomonas* demonstrated the highest positive correlation with r = 0.398. Conversely, penta-acylated *Prevotella* (r = -0.323) and hexa-acylated *Aggregatibacter* (r = -0.290) exhibited the most negative correlation values.

In GCF samples, the highest positive correlations were observed for *Desulfobulbus* (penta-acylated, r = 0.276), *Capnocytophaga* (penta-acylated, r = 0.273), and *Stomatobaculum* (Gram-positive, r = 0.272). Conversely, the most negative correlations were observed for *Acinetobacter* (penta-acylated, r = -0.299), *Kocuria* (Gram-positive, r = -0.283), and *Aerococcus* (Gram-positive, r = -0.274). Notably, for hexa-acylated bacteria, *Enterobacter* exhibited the highest positive correlation with r = 0.223. With the exception of *Aggregatibacter* at r = -0.122, all other hexa-acylated bacteria showed positive correlation values. Among penta-acylated genera, 76% of the bacteria had positive correlation values.

It is important to highlight that several of the mentioned bacteria exhibit low abundance in the respective sample types. For instance, *Stomatobaculum*, which exhibited a correlation value of 0.272, constituted only 0.004% of the total bacterial community in the GCF samples. Similarly, in GCF, the abundance of *Desulfobulbus* was 0.01%, *Acinetobacter* was 0.04%, *Kocuria* was 0.02%, *Aerococcus* was 0.02%, and *Enterobacter* was 0.002%. The bacteria with highest abundance were *Capnocytophaga* with 5.56%. In saliva, the abundance of *Bulleidia* (r = 0.427)

was 0.8%, *Selenomonas* was 0.48%, and *Aggregatibacter* was 0.12%. *Prevotella* had the highest abundance with 7.5%.

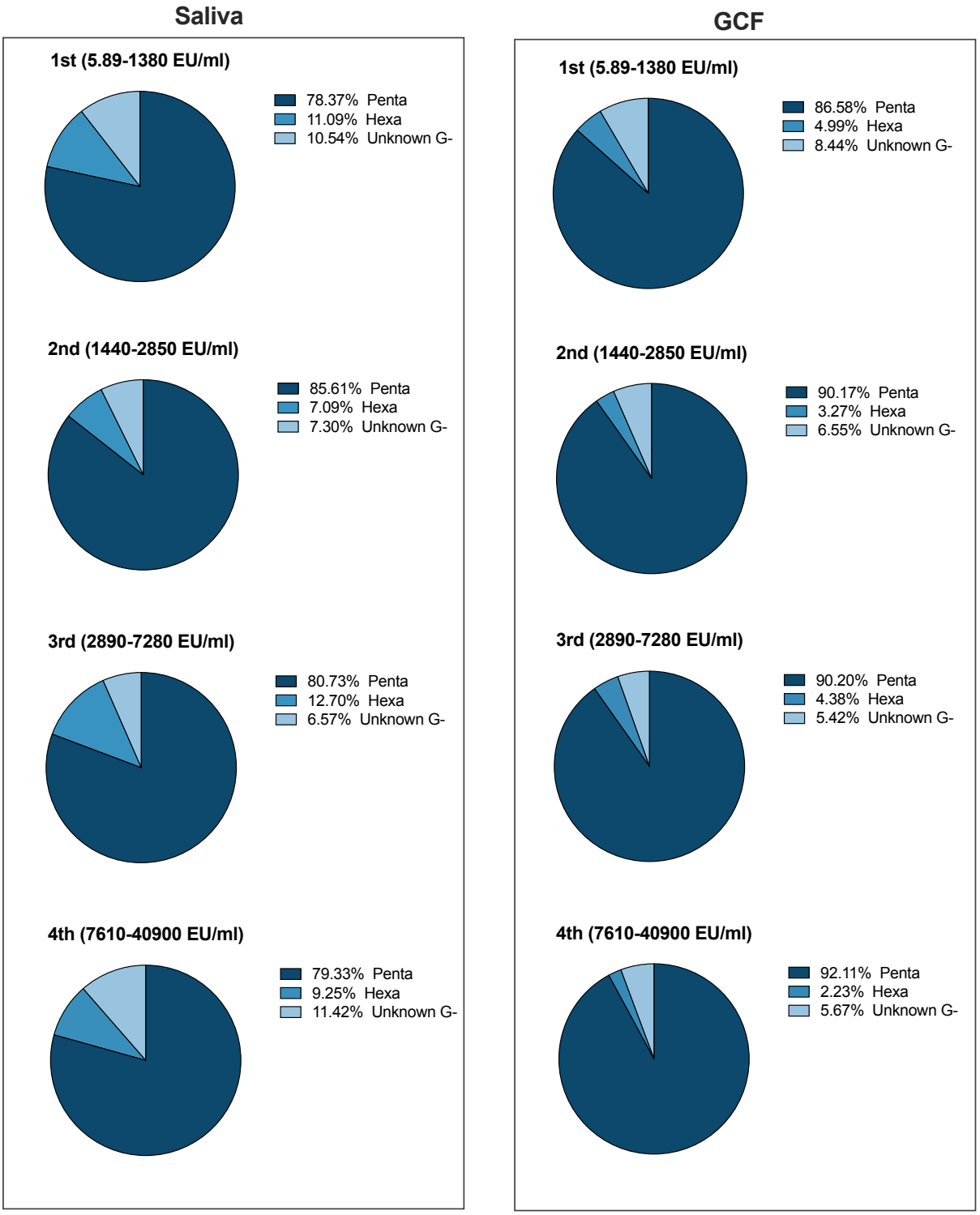



**Figure 22: Percentages of penta- and hexa-lipid A-producing bacteria, calculated as part of total Gram-negative bacteria in saliva and GCF samples, respectively. Based on LPS concentration the samples were divided into quantiles; 1st quantile: 5.89-1380 EU/ml, 2nd quantile: 1440-2850 EU/ml, 3rd quantile: 2890-7280 EU/ml, 4th quantile: 7610-40900 EU/ml.**

# 8. Discussion

The human microbiota is a complex and dynamic ecosystem that varies among individuals and is intricately linked to the microbial communities in our surroundings. While many of the microorganisms inhabiting the human body coexist with their host in a mutually beneficial relationship, others can cause disease. The immune system is vital for protection against pathogens, and studying the microbiome offers insights into human body functioning and disease mechanisms. The immune system recognizes LPS in Gram-negative bacteria's outer membrane, and the structure of the lipid A domain in LPS influences the strength of the immune response.

In this project we explored the oral LPS concentration in the saliva samples of the adult RHINESSA participants from the Bergen center, determined whether two methods of saliva collection interfered with the results, and annotated LPS-producers to hexa- or penta-acylated. To the best of our knowledge, no previous studies have assessed the LAL assay to samples collected with the saliva collection kit used in RHINESSA, but as the method is quick and sensitive, we wanted to investigate how it worked on our samples.

## 8.1 Pilot project sample collection and study design

In order to investigate if two different methods of saliva collection would affect the results, a pilot project was conducted as part of this thesis. Following the standardized procedure for saliva sample collection known as RHINESSA SOP, saliva was collected from 32 participants concurrently using both methods. Participants were instructed to follow the same instructions as those in the RHINESSA study, except for fasting requirements which were relaxed due to difficulties in finding candidates willing to fast or come in early in the morning. Saliva collection mostly took place between 10:00-14:00, and most participants had already consumed at least one meal before collection.

Since the main aim of this pilot project was to compare the two collection methods, we assessed samples in terms of intrasubject variability rather than intersubjective variability, making prior food consumption less important. Nevertheless, participants were asked about their eating and drinking habits, as well as whether they were taking asthma medication and when they last brushed their teeth. This information was not used in any other context, but it could explain why some participants found the sampling process difficult, as many factors can cause dry mouth [48]. Difficulties in saliva production may have led some participants to force

salivary production, which due to higher viscosity may have affected the processing of the samples.

The saliva collection had a cross-over study design, whereby half of the participants used the no-kit collection method first, while the remaining half used the kit collection method first. This design aimed to minimize potential discrepancies in saliva composition at the outset and towards the end of the collection process, thereby preventing possible variation in bacterial composition and abundance across the two samples, which might have occurred due to salivary flow stimulation, environmental changes, or other factors that could induce stress [49]. Although some alterations in pH and temperature might have occurred during the collection process, the sampling duration was limited, which reduced the likelihood of notable changes.

The collected saliva samples were processed according to the RHINESSA SOP, which involved either adding PBS and dividing the samples into 2 ml containers or freezing them immediately in the collection tubes. The no-kit samples were thoroughly mixed with PBS before being divided into three 2 ml containers, but due to high viscosity and the presence of mucus in certain samples, there was a potential variation in the saliva-to-PBS ratio, which could have influenced the bacterial abundance and composition observed across the containers. The addition of Sputum liquification buffer could have mitigated this issue; however, it was not used to adhere to the RHINESSA SOP. Furthermore, there may have been differences in bacterial concentration between kit and no-kit samples from the same person due to possible variations in the amount of saliva collected. The SOP stated to collect at least 2 ml of saliva using the no-kit method and to add 3 ml PBS, regardless of the amount of saliva. Although the goal was to collect 2 ml of saliva using both methods, some subjects produced more than 2 ml, especially with the no-kit method, and it was difficult to determine the exact amount collected due to the choice of collection tubes. When using the kit method of collection, it was easier to be precise about the 2 ml of saliva collected as the collection tube was smaller and marked with a line for 2 ml. In hindsight, to standardize the dilution of saliva 2 ml of collected saliva could have been measured using pipettes and added to 2 ml of either PBS or home-kit buffer before freezing. However, this approach was not implemented to comply with the RHINESSA SOP. Nonetheless, the approximate dilution of saliva was taken into account during the computer analysis, hence the results should not be substantially affected.

### 8.2 Dilution and interference

To optimize time and resources, sample dilution was deemed necessary to use the LAL assay efficiently, as per the manufacturer's recommendation. From the serially diluted samples collected continuously through the process, the results initially varied a lot both between each sample, but also between modes of collection. Generally, the kit samples showed lower OD values and LPS concentration than the no-kit samples, and the pattern in OD signal was not the same for each of the sample types, as the no-kit samples decreased in OD while the kit samples both decreased and increased in OD. Different mixing and pipetting techniques were performed to optimize the result and avoid variations in technique between the sample types, but eventually, it didn’t lead to less variation.

Kit and no-kit samples from the same person were analyzed to control for tentative difference in LPS concentration due to individual variability, but the results had the same trend with kit samples having much lower OD and LPS concentration. It should be noted that differences in OD values between mixing methods do not necessarily indicate superiority of one method over the other, given that the standard curve and incubation time may vary between the methods. Nonetheless, the differences in OD and LPS concentration in samples from the same person indicated interference. Initially we hoped there would be less differences between the collection methods, and it was challenging to find a solution to the issues. To further confirm the challenges between the two modes of collection, both kit and no-kit samples were included for analysis in the pilot project.

After conducting multiple experiments to determine the correct dilution for the sample types, the OD values for the differently diluted no-kit samples indicated that the 1:1000 dilution was optimal, as evidenced by the OR OD values for the UD, 1:10, and 1:100 dilutions. The kit samples, on the other hand, exhibited significantly lower absorbance signals compared to the no-kit samples, and measurements started at 1:10 dilution (See Table 5 and Table 6).

DNA extraction and LAL assay were performed on all samples collected for the pilot project. The DNA extraction results showed generally similar DNA concentration between the kit and no-kit samples obtained from the same individual. However, the LAL assay results pointed towards interference in the samples, as was expected from the dilution experiments results. Therefore, in accordance with the manufacturer's instructions, a spiking experiment was performed to identify the interfering factor. Spiking experiments are commonly performed to

identify the potential source of interference in samples and to determine a suitable dilution at which the interfering factor is minimized or eliminated [50].

To investigate the interference in the samples, spiking was performed in home-kit buffer, PBS, EFW, and LRW instead of the original samples, given the high levels of LPS which could have complicated the determination of recovery rates. The spiking experiment revealed intriguing results: spiking undiluted home-kit buffer showed a non-existent OD signal, indicating that the buffer added to the samples collected with the kit method could interfere with the assay. On the other hand, acceptable endotoxin recovery was observed in the spiked PBS and EFW samples, indicating that they were unlikely to cause the interference in the samples. To further explore the interference from the buffer, a subsequent spiking experiment was conducted, in which buffer, PBS, and EFW were spiked in various dilutions with the same concentration of endotoxin (0.75 EU/ml) to ensure that the LPS concentration would remain within the standard curve. Recovery rates in PBS and EFW were consistent with those obtained in the previous spiking experiment and well within the acceptable range. Conversely, the buffer showed signs of interference at dilutions of 1:10 and 1:100.

To eliminate the possibility of interference from the Hycult LAL kit [42], the Lonza LAL assay kit [43], which utilizes a kinetic chromogenic method, was employed to measure both kit and no-kit samples. As expected, the trend remained consistent with the Hycult kit, wherein the no-kit samples exhibited lower OD values with dilution, while kit samples displayed variable OD signal fluctuations with dilution. Furthermore, spiking experiments were conducted with undiluted samples using the Lonza kit, yielding results that were consistent with the Hycult spiking experiments. These findings were in line with our initial suspicion that the home-kit buffer would interfere with the LAL reagent. Given that LAL reagents are typically produced using similar methods, although the specific methodology employed by manufacturers remains confidential, it was reasonable to think that the buffer would interfere with both LAL kits.

Due to the interference observed in the kit samples there was uncertainty regarding the inclusion of these samples in the next phase of the project, which involved measuring the LPS bulk concentration in RHINESSA and investigating variations of lipid A. Upon contacting the home-kit manufacturer (Norgen) the inconclusiveness of the kit samples was acknowledged. The manufacturer explained that the buffer could interfere with the LAL assay due to its potential to denature proteins involved in enzymatic reactions. As it was not possible to obtain new samples from the participants of the RHINESSA study using the no-kit collection method,

it was concluded that the kit samples would be excluded from the LAL analysis. Despite not obtaining the desired results and finding interference in the kit samples, the findings of this pilot study remain valuable, considering the number of additional samples available in the RHINESSA study that were collected using the two different methods.

### 8.3 RHINESSA findings

The results from the LAL assay conducted on the RHINESSA samples revealed a higher concentration of LPS among the male sub-population compared to the female sub-population, although the result was not statistically significant. Conversely, the DNA concentration measured during DNA extraction demonstrated more comparable values between the sexes. This outcome suggests that men have more LPS-producing Gram-negative bacteria in their saliva than women, possibly due to various factors that differ between the sexes. For example, there is a higher prevalence of men using snus (18.75% vs 6.35% in females), which is associated with both oral and systemic diseases [51, 52]. In contrast, smoking affects the oral bacterial community composition and diversity [53], and in this adult cohort, female participants were more abundant among current and previous smokers compared to the male participants (7.94% current and 15.87% previous vs. 0% current and 6.52% previous, respectively). The difference in LPS concentration between smokers and snus users may be attributed to the prolonged direct contact of snus with the oral environment and gums compared to cigarette smoke. This prolonged exposure could potentially be a crucial factor influencing the observed difference in LPS concentration between these two groups.

Factors such as gum bleeding and tooth brushing frequency are crucial when evaluating the impact of oral hygiene on oral disease. Gum bleeding serves as a reliable indicator of gingival inflammation [54], and amongst the RHINESSA study population, a higher prevalence of gum bleeding was observed among female participants, with six individuals (9.52%) reporting gum bleeding always or often, in contrast to only one male (6.52%) participant. Another crucial aspect is tooth brushing frequency, which was found to vary between genders. None of the male participants reported brushing their teeth more than twice a day, while 9.52% of the female participants reported doing so. The World Health Organization recognizes the significance of tooth brushing and proper oral hygiene practices in mitigating the risk of oral disease [25].

Another factor commonly associated with oral disease is Community Periodontal Index (CPI). CPI is used to assess the severity and prevalence of periodontitis. It measures the presence of periodontal pockets and bleeding gums and is a standardized method. Normally the

CPI score ranges from 0 to 4, with 0 indicating healthy gums and 4 indicating severe periodontitis with deep pockets around the tooth being examined [54]. In the RHINESSA study population, the highest measured CPI score was 3. The total number of male participants with CPI>0 was 17.65 %, while for women the percentage was 11.1%. Among the male participants none had a CPI score of 3, contrasting to the female population where 4.76% of the females had CPI=3.

Together these findings could explain why there were more men who had high LPS levels ($4^{th}$ quantile, 7610-40900 EU/ml) compared to the females (8 (50%) vs 11 (17%), respectively). However, due to the smaller number of male participants (16) compared to female participants (63), the difference in LPS concentration between the two sexes could be less pronounced with a larger sample size of men. Furthermore, the larger standard deviation for male participants' LPS concentration compared to females indicates more variability in the data. Nonetheless, using the median value, the data still indicates that men had higher LPS concentration than women, which could be attributed to differences in oral hygiene habits, microbial composition, and other risk factors between the sexes.

When dividing the population into groups based on known factors for oral disease, no significant difference in the median LPS concentration was observed (see Figure 16). However, there were some noteworthy differences between the groups. For instance, participants with CPI>0 had higher LPS concentration, as did those who used tobacco. Conversely, participants who had used antibiotics within the past year had slightly lower LPS concentration than those who did not. It is well-established in the literature that the use of antibiotics can alter the overall oral bacterial composition [55]. It is worth noting that the groups with factors that could indicate oral disease, such as gum bleeding and tobacco use, had a smaller population size than the comparative group. This could be attributed to the study's overall healthy population that reported no issues with oral health.

The normalization of LPS and DNA concentration in both the pilot and main projects demonstrated non-Gaussian distribution, which is anticipated given that microbiome data typically deviates from normality [56, 57]. However, it is difficult to ascertain whether the LPS concentration and distribution align with expectations, as LPS concentration measurement in saliva using the LAL assay has not been extensively researched. To illustrate this, the range of LPS concentration in this project was 5.89-40900 EU/ml, with a median value of 3300 EU/ml,

indicating a higher concentration of LPS in saliva in healthy adults than some studies [58, 59], but lower concentration than others [60].

### 8.4 General challenges and interference in LAL assay

Despite being the gold standard test for endotoxin detection, the Limulus Amebocyte Lysate (LAL) assay has some limitations. As previously mentioned, the inability to differentiate between different lipid A variants and challenges with interference and finding the appropriate sample dilution are among them.

Due to the purpose of excluding endotoxins from samples, most LAL assays have a low upper limit to the standard curve. This was particularly challenging in our project, where most samples had to be diluted at least 1:1000. This high degree of dilution can introduce bias, which may lead to inaccurate LPS concentration measurements, as each dilution step provides the opportunity for contamination. Thus, the ideal approach would be to measure LPS concentration in undiluted samples. However, in cases where this is not feasible, extra emphasis should be placed on cleanliness and materials used.

The selection of materials for the assay is also critical. For example, components of the LAL reagent may react with cotton, necessitating the use of cotton-free pipette tips [61]. Given the high sensitivity of the assay, it is critical to minimize equipment and material contamination. Despite being sterile, plastic materials frequently contain DNA and endotoxins. Additionally, endotoxins have a tendency to adhere to plastic materials, making autoclaved glassware a superior option [62]. In our study, we chose DNA- and endotoxin-free polystyrene materials for saliva collection, consistent with the RHINESSA study, and utilized the same type of material during LAL assay, taking care to mix the samples thoroughly to dislodge endotoxin from the tube wall.

Currently, LAL assays are not utilized in diagnostics due to uncertainties surrounding the impact of endotoxemia on the body, particularly in the bloodstream. Furthermore, performing LAL assays on blood plasma may be complicated by interfering factors [35]. But, despite its limitations, the LAL assay remains a widely adopted and sensitive method for ensuring the safety and quality of research samples, pharmaceuticals and medical devices, provided it is used appropriately. Moreover, it offers the benefits of being cost-effective, quick, and easily automatable, at least in comparison to targeted sequencing and other metagenomic approaches.

### 8.5 Targeted sequencing

To investigate how the LAL assay picked up the different variations of lipid A-producing bacteria, 16S gene sequences from GCF and saliva samples were used. DNA extraction is the first step in targeted sequencing and is as mentioned known to introduce the largest amount of bias. In this experiment, both mechanical and enzymatic steps were used for cell lysis to account for the different thickness and vulnerability of bacterial cell walls. The enzymatic step was performed first to maintain the integrity and yield of DNA extracted from the most easily lysed cells, as more powerful lysis methods could have fragmented DNA from the more vulnerable bacteria. Maintaining DNA integrity is crucial for an accurate representation of community composition, as fragmented DNA can increase the formation of recombinant amplification products known as chimeras during PCR, which can lead to significant errors in interpreting bacterial data. This can to some extent be fixed during the bioinformatics processing of the gene sequences, but it was taken into account during extraction to minimize the error. The enzymes lysozyme, mutanolysin, and lysostaphin have different targets for cell lysis, and were used in combination to ensure breakage of all bacterial cell walls [28].

PCR primers bias is a significant issue in post-sequencing analysis of bacterial community composition. The conserved regions in the 16S gene are not 100% similar in all bacteria, so some bacteria may not be accurately represented in the pool of 16S amplicons generated during PCR. Additionally, differences in 16S gene copy numbers across bacterial genomes can lead to further bias, with higher copy numbers overrepresented in the amplicon pool. To address these issues, degenerate primers are widely used in the design of universal PCR primers, where a mixture of primers that vary only at specific base positions that are less conserved is used. However, in this project, different primers were used for gingival and saliva samples, which can make it difficult to compare bacterial diversity and community composition between the two types of samples [28].

Despite the introduction of bias and errors during extraction, PCR, and gene sequencing, we have confidence that our implemented steps provided a reliable representation of bacterial composition and abundance in the samples used for our project. A direct comparison between the composition of saliva and gingival samples may not be entirely accurate due to different targeted variable regions of the 16S rRNA gene, but nevertheless our approach provides a close representation of the two sample types.

### 8.6 Lipid A annotation and compositional analysis

Analysis of the bacterial composition in saliva and GCF samples from the RHINESSA study population revealed the presence of 16 and 10 bacterial phyla compromising 139 and 154 genera, respectively. In terms of abundance, approximately ¼ of the bacteria were annotated as Gram-negative in both sample types. Less than 1% of the found genera were uncategorized. Within the Gram-negative bacteria, which are the focus of this study, 10.13% of the genera in saliva and 3.71% of the genera in GCF had the ability to produce pro-inflammatory hexa-acylated lipid A forms. The percentage of penta-acylated lipid A bacteria in the respective samples were 80.8 and 89.71. The low ratio of hexa- to penta-acylated LPS producers was previously reported by Brix et al. (2015) [16] in healthy individuals' lung microbiota. The same study reported that people with asthma had a different hexa:penta ratio, with more hexa-acylated LPS-producing bacteria in their lungs, possibly leading to airway inflammation. Our results were consistent with the expected findings in mainly healthy subjects, and we hypothesize that the potential pro-inflammatory effect of hexa-acylated lipid A is counterbalanced by the predominance of penta-acylated lipid A producers such as Bacteroidetes and Firmicutes.

Although no bacteria in the samples were specifically annotated as solely tetra-acylated, it is important to note that this does not exclude the possibility of bacteria potentially producing tetra-acylated lipid A. Research has shown that certain bacterial species, such as *P. gingivalis* and *H. pylori*, have the capability to produce this form of lipid A, which serves as a protective mechanism against immune system surveillance. However, identifying the specific bacteria with these characteristics within the samples in this study would necessitate more extensive data analysis on species level (e.g., shotgun metagenomics of these samples) and further research in the field, as the investigation of tetra-acylated lipid A is still an ongoing area of study.

When dividing the samples into quantiles based on their LPS concentration, an expected increase in the number of Gram-negative bacteria was observed, as LPS is primarily released by Gram-negative bacteria, as measured by the LAL assay. However, a detailed analysis of the hexa:penta ratio between the quantiles revealed only modest differences. For instance, in GCF samples, the first quantile with LPS concentrations ranging from 5.89-1380 EU/ml had 2.71% of bacteria producing hexa-acylated lipid A, while the fourth quantile with LPS concentrations of 7610-40900 EU/ml had 1.65% of bacteria producing hexa-acylated lipid A. The abundance

of penta-acylated LPS producers showed minimal variation across the quantiles, with the largest difference observed between the first quantile (90.35%) and the third quantile (95.02%). These differences could be attributed to variations in the number of bacteria with unknown lipid A status, which varied among the quantiles. Similar trends were observed in the analysis of saliva samples, where the levels of hexa- and penta-acylated bacteria were higher. However, similar to the GCF samples, there were minimal differences across the quantiles. The most notable exception was observed in the fourth quantile, where 11.17% of the bacteria were classified as either unknown genera or had unknown lipid A variants, which was 4.26% higher than the other quantiles.

Overall, the results from the saliva samples revealed more gram-negative bacteria with unknown LPS variants than the results from the GCF samples. The mean amount of unknown lipid A variants were 9.02% and 6.57%, respectively. The differences in bacterial composition and abundance between the two sample types could be due to taxonomy being based on different databases for the sample types and the use of different variable regions as primers [63]. Originally it was planned to use the same database for both sample types, but due to time and technical difficulties the saliva gene sequences were taxonomically assigned using the Greengenes database and not the HOMD like the GCF samples. Unlike the HOMD which is specifically focused on the oral microbiome, Greengenes covers a wide range of microbial taxa from various environments, including the human oral microbiome. Using the Greengenes database for salvia samples could especially explain the higher abundance of unknown gram-negative bacteria in the samples [44, 47]. The observed differences in bacterial composition and abundance between the saliva and GCF samples could also potentially arise from other factors including small variations in laboratory protocols, differences in personnel conducting the sample analysis, variations in the storage duration of the samples, and variations in the environmental conditions under which the samples were stored and analyzed.

Microbiome data exhibits a compositional nature, which presents challenges when conducting correlation analysis. Compositionality arises from the fact that the proportions of different microbial taxa sum up to 100% within a given sample, and the measurement of microbial taxa focuses on their relative abundance rather than absolute abundance. This inherent compositional structure poses implications for statistical analysis and interpretation of the data. To address the limitations associated with compositional data, a centered log-ratio (CLR) transformation was employed prior to conducting correlation analysis. The log-ratio

transformation involves calculating the logarithm of the ratios between the abundance of individual microbial taxa and the geometric mean abundance of all microbial taxa within the sample. By applying CLR transformation, the closure property of compositional data is mitigated, enabling the use of correlation analysis and any other multivariate statistical tools [64].

The correlation analysis revealed a positive correlation between the abundance of penta-acylated bacteria and the concentration of LPS, while a negative correlation was observed between the abundance of hexa-acylated bacteria and LPS concentration. These findings align with the low hexa:penta ratio observed in the samples and suggest that the pro-inflammatory effects associated with hexa-acylated bacteria are counteracted by the predominance of penta-acylated bacteria, as indicated in previous studies involving healthy populations [16, 45]. Drawing conclusions solely from the correlation of individual species with LPS concentration is challenging due to their low abundance and the lack of knowledge regarding the "normal" abundance in healthy individuals, as well as the substantial inter-individual variation. Furthermore, the relevance of specific bacteria's high or low abundance and their correlation with LPS concentration may be less significant compared to the overall microbial community composition, as abundances can vary considerably both between individuals and within individuals over time [18].

The results from LAL assay, lipid A annotation and correlation points towards that LAL assay detects both hexa- and penta-acylated lipid A forms, regardless of their relative abundance in the sample. This was observed even in samples with very low LPS and DNA concentrations, indicating that the assay is not affected by the concentration of LPS or DNA in the sample. Nonetheless, it is important to note that accurately determining the extent of this effect remains challenging. While the LAL assay is a reliable method to measure bulk LPS concentration, it cannot differentiate between the various lipid A variants [34]. Investigating how well the LAL assay detects different lipid A variants is challenging due to unknown factors such as bacterial potency, LPS content in the cell wall, environmental conditions, and inter-individual variability of the host.

### 8.7 GCF vs saliva

Saliva and GCF have different bacterial compositions. Saliva is a mixture of secretions from major and minor salivary glands [49], while GCF is an exudate from the crevice between the

tooth and the gingiva [65]. This means that saliva samples are total collective samples, while GCF samples are niche-specific.

Studies have shown that saliva has a more diverse bacterial community composition compared to GCF. The most abundant phyla in saliva are Firmicutes, Proteobacteria, Bacteroidetes, and Actinobacteria, while in GCF, some of the most abundant phyla are Firmicutes and Bacteroidetes [24, 66, 67]. Furthermore, there are also differences in the relative abundance of certain bacterial species between the two fluids. For example, *Streptococcus* and *Prevotella* are more abundant in saliva [67], while *Porphyromonas* and *Treponema* are more abundant in GCF [65].

In this project, there were observed differences in bacterial diversity and taxa distribution across the two sample types. The dominant phyla in saliva were Firmicutes (59.19%), Actinobacteria (20.82%) and Bacteroidetes (9.69%), while in GCF the most abundant phyla were Bacteroidetes (25.19%), Firmicutes (24%) and Fusobacteria (19.36%). In saliva samples *Streptococcus* was the dominant genus with 41.34% prevalence. *Streptococcus* also had a high prevalence in GCF samples, with 10.85% ranking number two in decreasing order of prevalence, beaten by *Fusobacterium* at 16.1%. These findings supported the previously reported results [65, 67]. As mentioned in section 8.6, the differences in bacterial composition and abundance between the sample types could be due to several factors.

Overall, the different bacterial compositions between saliva and GCF can provide insight into the role of oral microbiota in maintaining oral health and disease. Gaining access to samples from individuals suffering from both systemic and oral diseases is frequently arduous and time-consuming. Nonetheless, by quantifying specific markers for these diseases in saliva, the process could be streamlined. Saliva offers an easily accessible and non-invasive collection method of a wide array of immune and inflammatory markers. Repeated sampling can be conducted without causing any discomfort to the patient, allowing for the monitoring of alterations over time, and providing possible diagnostic and prognostic significance. Despite its apparent simplicity, the field of measuring inflammation markers in saliva is not yet fully understood and necessitates further research. The dependability and validity of quantifying inflammation markers in saliva are still somewhat unclear, with results showing inconsistencies [68, 69]. Studies similar to this one may provide a deeper understanding of the interplay between the host oral microbiota and measured salivary markers of inflammation.

# 9. Conclusion

The main aims of this thesis were to measure LPS concentration in saliva samples from the RHINESSA study using LAL assay, as well as investigating potential variations in the results based on two different methods of saliva collection (i.e., kit and no-kit). Additionally, the aims were to determine the different lipid A variants present in the samples and explore the relative ratios of them, while assessing whether the LAL assay could detect the different variants equally.

Using LAL assay kits on saliva samples collected with two different methods, led us to conclude that samples collected using a specific saliva collection kit from Norgen led to interference with the assay. Consequently, the results derived from the LAL assay on these specific samples could not be considered reliable, and samples collected in this manner were therefore excluded from further analysis. Conversely, the samples collected utilizing the alternative collection method were deemed suitable for inclusion and further analysis. As a larger number of samples from the RHINESSA study were collected using these two methods, this information proves valuable for future studies.

The results of the LAL assay of the RHINESSA saliva samples revealed higher LPS concentrations in male participants compared to their female counterparts, but no statistically significant difference. Taxonomic composition analysis of saliva and GCF gene sequences produced results similar to other studies, hopefully contributing to further insight into the oral microbiome composition of healthy subjects and deeper understanding of inflammatory markers in saliva and GCF. Lipid A annotation revealed a low hexa- to penta-acylated ratio of bacteria in both GCF and saliva, which could be attributed to a healthy population. The hexa:penta ratio was similar across the quantiles when dividing the samples based on LPS concentration. This along with correlation analysis pointed towards the ability of the LAL assay to pick up both lipid A variants even in low abundances and in low LPS concentration samples. However, this is complex to determine with certainty due to environmental and genetic factors influencing the potency and abundance of the specific lipid A variants.

# 10. Limitations and future perspectives

In the context of limitations and further perspectives, it is crucial to consider the potential improvements that could enhance the findings and implications of this thesis. One such aspect is the allocation of additional time and resources to explore optimal dilution for kit samples or find alternative quick methods of measuring bulk LPS concentration. This could offer the opportunity to quickly examine bulk LPS concentrations at two-time points and make a comparative analysis. The investigation of potential variations in samples taken during the pandemic would be particularly intriguing. Further exploration of the appearance of the OD signal in kit samples at higher dilutions during the spiking experiments could have provided valuable insights with additional time and resources. However, it is important to acknowledge that some samples may have had low LPS concentrations, making reliable detection at high dilutions challenging and potentially leading to false results.

Furthermore, the DNA extraction and gene sequencing processes started relatively late in the project due to delivery issues with certain reagents, which were experienced by many other researchers during this period marked by the recent pandemic and conflicts in Europe. Furthermore, the time-intensive nature of DNA extraction further delayed the arrival of gene sequencing data. Given more time, a more comprehensive bioinformatic analysis of the gene sequences could have been conducted, including compositional analyses using bacterial community- (prediction with MiRKAT) and genus-level (differential abundance analysis with ANCOM-BC) statistical methods. These analyses could have unveiled associations between oral disease factors and the abundance of hexa-acylated or penta-acylated lipid A variants in the saliva and GCF samples of participants.

Although the RHINESSA study is designed as a longitudinal investigation, our focus was on samples collected at a specific point in time rather than continuously. However, it would be valuable to conduct a long-term study measuring LPS concentrations and investigating bacterial composition in individual subjects, including parents and children. Although such an endeavor would require several years to complete, there are already available samples collected at different time points facilitating the start of this research. Additionally, investigating the bacterial composition in the saliva of healthy individuals more frequently and consistently over a shorter duration, such as multiple times per day throughout a week or once a week over several months, could provide valuable insights into the dynamics of oral environmental changes and their relationship to oral and systemic health from a broader perspective.

## 11. Funding

This work was funded by the Meltzer Research Fund 2022 grant to Maryia Khomich and European Research Council Grant No. 804199 to Randi Jacobsen Bertelsen.

# 12. References


1. Needham, B.D. and M.S. Trent, *Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis.* Nat Rev Microbiol, 2013. **11**(7): p. 467-81.

2. Barry, M., Bleackley, R., *Cytotoxic T lymphocytes: all roads lead to death.* Nature Reviwes Immunology, 2002. **2**: p. 401-409.

3. Medzhitov, R. and C. Janeway, Jr., *Innate immunity.* N Engl J Med, 2000. **343**(5): p. 338-44.

4. Kadner, R.J., Rogers, Kara., *Bacteria*, in *Britannica*. 2023, Encyclopaedia Britannica Inc. : Chicago, Illinois, USA.

5. Silhavy, T.J., D. Kahne, and S. Walker, *The bacterial cell envelope.* Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000414.

6. Pajerski, W., et al., *Attachment efficiency of gold nanoparticles by Gram-positive and Gram-negative bacterial strains governed by surface charges.* Journal of Nanoparticle Research, 2019. **21**.

7. Bertani, B. and N. Ruiz, *Function and Biogenesis of Lipopolysaccharides.* EcoSal Plus, 2018. **8**(1).

8. Lin, T.L., et al., *Like Cures Like: Pharmacological Activity of Anti-Inflammatory Lipopolysaccharides From Gut Microbiome.* Frontiers in Pharmacology, 2020. **11**.

9. Maeshima, N. and R. Fernandez, *Recognition of lipid A variants by the TLR4-MD-2 receptor complex.* Frontiers in Cellular and Infection Microbiology, 2013. **3**.

10. Rosadini, C.V., et al., *A Single Bacterial Immune Evasion Strategy Dismantles Both MyD88 and TRIF Signaling Pathways Downstream of TLR4.* Cell Host Microbe, 2015. **18**(6): p. 682-93.

11. Srivastava, N., et al., *Aeromonas hydrophila utilizes TLR4 topology for synchronous activation of MyD88 and TRIF to orchestrate anti-inflammatory responses in zebrafish.* Cell Death Discovery, 2017. **3**(1): p. 17067.

12. Liu, T., et al., *NF-κB signaling in inflammation.* Signal Transduction and Targeted Therapy, 2017. **2**(1): p. 17023.

13. Raetz, C.R.H., et al., *Lipid A Modification Systems in Gram-Negative Bacteria.* Annual Review of Biochemistry, 2007. **76**(1): p. 295-329.

14. Polissi, A. and P. Sperandeo, *The lipopolysaccharide export pathway in Escherichia coli: structure, organization and regulated assembly of the Lpt machinery.* Mar Drugs, 2014. **12**(2): p. 1023-42.

15. Herath, T.D., et al., *Tetra- and penta-acylated lipid A structures of Porphyromonas gingivalis LPS differentially activate TLR4-mediated NF-kappaB signal transduction cascade and immuno-inflammatory response in human gingival fibroblasts.* PLoS One, 2013. **8**(3): p. e58496.

16. Brix, S., et al., *Metagenomic heterogeneity explains dual immune effects of endotoxins.* J Allergy Clin Immunol, 2015. **135**(1): p. 277-80.

17. Marchesi, J.R. and J. Ravel, *The vocabulary of microbiome research: a proposal.* Microbiome, 2015. **3**(1): p. 31.

18. Sharon, I., et al., *The Core Human Microbiome: Does It Exist and How Can We Find It? A Critical Review of the Concept.* Nutrients, 2022. **14**(14).

19. Qin, J., et al., *A human gut microbial gene catalogue established by metagenomic sequencing.* Nature, 2010. **464**(7285): p. 59-65.

20. Ursell, L.K., et al., *Defining the human microbiome.* Nutr Rev, 2012. **70 Suppl 1**: p. S38-44.

21. Turnbaugh, P.J., et al., *The Human Microbiome Project.* Nature, 2007. **449**(7164): p. 804-810.

22. Proctor, L.M., et al., *The Integrative Human Microbiome Project.* Nature, 2019. **569**(7758): p. 641-648.

23. Deo, P.N. and R. Deshmukh, *Oral microbiome: Unveiling the fundamentals.* J Oral Maxillofac Pathol, 2019. **23**(1): p. 122-128.

24. Lee, Y., et al., *Efficacy of salivary versus subgingival bacterial sampling for the detection and quantification of periodontal pathogens.* J Periodontal Implant Sci, 2020. **50**(6): p. 358-367.

25. World Health Organization, *Global oral health status report: towards universal health coverage for oral health by 2030*. 2022, Geneva: World Health Organization.

26. Li, X., et al., *Systemic diseases caused by oral infection.* Clin Microbiol Rev, 2000. **13**(4): p. 547-58.

27. Kitamoto, S., et al., *The Bacterial Connection between the Oral Cavity and the Gut Diseases.* J Dent Res, 2020. **99**(9): p. 1021-1029.

28. Drengenes, C., *Methodological issues in airway microbiome studies*. 2020, University of Bergen: Bergen, Norway.

29. Bolyen, E., et al., *Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.* Nature Biotechnology, 2019. **37**(8): p. 852-857.

30. Schloss, P.D., et al., *Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities.* Appl Environ Microbiol, 2009. **75**(23): p. 7537-41.

31. Shehadul Islam, M., A. Aryasomayajula, and P.R. Selvaganapathy, *A Review on Macroscale and Microscale Cell Lysis Methods.* Micromachines, 2017. **8**(3).

32. Novitsky, T.J., *Limulus Amebocyte Lysate (LAL) detection of endotoxin in human blood.* Journal of Endotoxin Research, 1994(1): p. 254-263.

33. Piehler, M., et al., *Comparison of LAL and rFC Assays-Participation in a Proficiency Test Program between 2014 and 2019.* Microorganisms, 2020. **8**(3).

34. Mohr, A.E., et al., *Lipopolysaccharide and the gut microbiota: considering structural variation.* FEBS Lett, 2022. **596**(7): p. 849-875.

35. Hurley, J.C., et al., *Endotoxemia as a diagnostic tool for patients with suspected bacteremia caused by gram-negative organisms: a meta-analysis of 4 decades of studies.* J Clin Microbiol, 2015. **53**(4): p. 1183-91.

36. Svanes, C., et al., *Cohort profile: the multigeneration Respiratory Health in Northern Europe, Spain and Australia (RHINESSA) cohort.* BMJ Open, 2022. **12**(6): p. e059434.

37. Burney, P.G., et al., *The European Community Respiratory Health Survey.* Eur Respir J, 1994. **7**(5): p. 954-60.

38. Helse Bergen, H.U.S., *Standard Operating Procedure RHINESSA Adult, Adolecent and Child*. 2018, Haukeland Universitetssjukehus.

39. Hoang, T., Wiker, H., Eagan, TM, Drengenes, C. *16S Amplicon PCR for the V3-V4 region for the MicroCOPD samples*. 2019; Available from: https://www.protocols.io/view/16s-amplicon-pcr-for-the-v3-v4-region-for-the-micr-bp2l64zndvqe/v1/metadata

40. Thermo Fisher Scientific Inc. *NanoDop 1000 Spectophotometer V3.8 User's Manual*. 2010; Available from: https://tools.thermofisher.com/content/sfs/manuals/nd-1000-v3.8-users-manual-8%205x11.pdf

41. Thermo Fisher Scientific Inc. *Qubit 3.0 Fluorometer User guide*. 2014; Available from: https://tools.thermofisher.com/content/sfs/manuals/qubit_3_fluorometer_man.pdf.

42. HycultBiotech, *Limulus Amebocyte Lysate Chromogenic Endpoint Assay HIT 301*. 2022: Uden, Netherlands.

43. Aarhus Universitet, *Endotoxin analysis extracts by a kinetic LAL assay*. 2011: Aarhus, Denmark.

44. Chen, T., et al., *The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information.* Database, 2010. **2010**.

45. Khomic, M., et al., *Assosiation between lipid A-producing oral bacteria of different potency and fractional exhaled nitric oxide in a Norwegian poulation-based adult cohort.* Journal of translational medicine, forthcoming, 2023.

46. Klindworth, A., et al., *Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies.* Nucleic Acids Res, 2013. **41**(1): p. e1.

47. DeSantis, T.Z., et al., *Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.* Appl Environ Microbiol, 2006. **72**(7): p. 5069-72.

48. Scully, C., *Drug effects on salivary glands: dry mouth.* Oral Diseases, 2003. **9**(4): p. 165-176.

49. Iorgulescu, G., *Saliva between normal and pathological. Important factors in determining systemic and oral health.* J Med Life, 2009. **2**(3): p. 303-7.

50. Bolden, J.S., et al., *Endotoxin recovery using limulus amebocyte lysate (LAL) assay.* Biologicals, 2016. **44**(5): p. 434-40.

51. Vist, G.E., et al., *Are the health risks of moist oral snuff (snus) underestimated?* Tidsskr Nor Laegeforen, 2020. **140**(9).

52. Folkehelseinstituttet, *Health risks from snus use*. 2019. p. 243.

53. Wu, J., et al., *Cigarette smoking and the oral microbiome in a large study of American adults.* The ISME Journal, 2016. **10**(10): p. 2435-2446.

54. Petersen, P.E., Baetz, R.J.,, *Oral Health Surveys Basic Methods*. 2013, World Health Organization: Switzerland. p. 47-52.

55. Li, X., et al., *The Oral Microbiota: Community Composition, Influencing Factors, Pathogenesis, and Interventions.* Frontiers in Microbiology, 2022. **13**.

56. Wirbel, J., et al., *Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer.* Nature Medicine, 2019. **25**(4): p. 679-689.

57. Rai, S.N., et al., *Microbiome data analysis with applications to pre-clinical studies using QIIME2: Statistical considerations.* Genes & Diseases, 2021. **8**(2): p. 215-223.

58. Millns, B., M.V. Martin, and M.C. Williams, *Raised salivary endotoxin concentration as a predictor of infection in pediatric leukemia patients.* Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology, 1999. **88**(1): p. 50-55.

59. Leenstra, T.S., et al., *Oral endotoxin in healthy adults.* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 1996. **82**(6): p. 637-43.

60. Hyvarinen, K., et al., *A common periodontal pathogen has an adverse association with both acute and stable coronary artery disease.* Atherosclerosis, 2012. **223**(2): p. 478-84.

61. Finkelman, M., *(1,3)-B-D-Glucan contamination of cotton-containng material medical devices*, in *LAL update*, T.J. Novitsky, Editor. 2002.

62. U.S Department of Health and Human Services, *Guidance for Industry - Pyrogen and endotoxin testing: Qustions and Answers*. 2012.

63. López-Aladid, R., et al., *Determining the most accurate 16S rRNA hypervariable region for taxonomic identification from respiratory samples.* Scientific Reports, 2023. **13**(1): p. 3974.

64. Gloor, G.B., et al., *Microbiome Datasets Are Compositional: And This Is Not Optional.* Frontiers in Microbiology, 2017. **8**.

65. Pei, J., et al., *Microbial and metabolomic analysis of gingival crevicular fluid in general chronic periodontitis patients: lessons for a predictive, preventive, and personalized medical approach.* EPMA J, 2020. **11**(2): p. 197-215.

66. Kim, Y.T., et al., *Comparison of the oral microbial composition between healthy individuals and periodontitis patients in different oral sampling sites using 16S metagenome profiling.* J Periodontal Implant Sci, 2022. **52**(5): p. 394-410.

67. Murugesan, S., et al., *Profiling the Salivary microbiome of the Qatari population.* J Transl Med, 2020. **18**(1): p. 127.

68. Slavish, D.C., et al., *Salivary markers of inflammation in response to acute stress.* Brain Behav Immun, 2015. **44**: p. 253-69.

69. Szabo, Y.Z. and D.C. Slavish, *Measuring salivary markers of inflammation in health research: A review of methodological considerations and best practices.* Psychoneuroendocrinology, 2021. **124**: p. 105069.